1984

Aerospace Remote Sensing for Domestic Crop Area Estimation by the United States Department of Agriculture

By Charles E. Caudill

INTRODUCTION

Aerospace Remote Sensing for Domestic Crop Area Estimation

by the United States Department of Agriculture

By Charles E. Caudill

Introduction

In the United States of America (US) the mission of collecting and disseminating current statistics on the Nation's agriculture has been assigned to the Statistical Reporting Service (SRS) in the U. S. Department of Agriculture (USDA). By tradition, and in some instances by law, this mission includes responsibility for national and state level statistics. Responsibility for local area (minor civil divisions such as counties, etc.) agricultural statistics has in general been assigned to state governments. However in most states, SRS has been able to combine these two responsibilities through cooperative arrangements with state government agencies. In most cases, with the State Departments of Agriculture. SRS through its national office in Washington, D. C. and its 44 state offices (a single office serves both Maryland and Delaware and one office serves all of the six New England States) issues about 300 national and 9,000 state reports each year (17). Comprehensive current statistics on all aspects of the Nation's agriculture are included in these reports. Content ranges from weekly subjective assessments of weather and its effects on agricultural production to precise quantitative estimates of agricultural inventories, crop and livestock production, prices, marketings, inputs, wage rates, etc.

In addition to its primary mission of collecting and disseminating agricultural statistics, SRS conducts a comprehensive program of methodological research to improve the statistical methods used for its primary mission. The ultimate goal of the research is to enhance the accuracy of the agricultural statistics produced and/or to reduce the costs associated with a given level of accuracy. Improved sampling, data collection and analysis procedures are constantly under study and development with this goal in mind. New technology, such as computers and aerospace remote sensing, as well as applications of new crop production practices constantly require study and research to understand how they affect or can be used to improve agricultural statistics. Since some of the most important statistics produced by the SRS are the estimates of area planted to various crops a significant proportion of the agency's research resources have always been devoted to improving the methods and data used for these estimates. During the past decade this research has been concentrated on the development of methodology to use remote sensing data collected by earth-orbiting satellites for making crop area estimates.

This research will be discussed in detail. First the present operational procedures used for making crop area estimates are briefly described. Then a description of remote sensing technology and the research conducted over the past decade will be discussed. Finally results of using aerospace remote sensing data for a large scale application test in 1978 are presented and prospects for the future examined.

1. Background

1.1. Current Methodology for Estimating Crop Area (17)

SRS estimates of crop area are developed from data collected through sample surveys. Generally these surveys are based on relatively small samples because of costs, and timeliness requirements for data summarization, analysis, and publication (typically the time interval from the start of data collection until estimates are published is about one month). Since 1965 national and state level estimates of crop area have been based primarily on data collected through a probability area sample, commonly referred to as the June Enumerative Survey (JES). This survey uses a national sample of about 16,000 areas of land. These areas of land average about one mile square and are called sampling units or more commonly "segments." They are randomly selected from an area sampling frame. The area sampling frame used for this purpose was developed by using maps, published statistics, aerial mosaics or photographs, and other available materials to stratify the entire land area of the US into major land use strata (e.g., cultivated land, rangeland, woodland, cities and nonagricultural land). This stratification permits very efficient samples to be designed which provide precise and independent estimates at the state and national levels. The sample size for the JES only accounts for about one-half of one percent of the total land area in the country, yet it produces estimates of area planted to various crops which have sampling errors ranging from one to two percent at the national level and three to five percent at the state level.

Data collection for the JES is done through personal interviews with individual farmers. Interviewers who are usually hired on a part-time basis are thoroughly trained before each JES to assure uniformity of data collection methods throughout the nation. In addition to the training program associated with the JES many other quality controls are included in the JES operating procedures. These include detailed instruction manuals and materials, close field supervision of interviewers, built-in questionnaire checks, and re-enumeration of a subsample of segments by supervisors. Aerial photographs are used to locate and identify the segments and to compare area, on a field-by-field basis, reported by farmers with measured area on the photographs.

Interviewers use state and county road maps to locate segments and aerial photographs to define precisely the segment and field boundaries. Each person who operates land within the segment is interviewed to obtain information about area and use for every field within the segment. This information, which includes both net and gross field area for fields used for planted crops, is carefully recorded on the survey questionnaire for later summarization in the state office. Data collection for the JES starts during the last week of May each year and the survey is completed in June. This includes data collection, editing, summarization and analysis, and publication of planted area estimates.

1.2. Aerospace Remote Sensing

"Remote sensing is the acquisition of information about an object without physical contact... However, the use of the term remote sensing usually refers to the gathering and processing of information about the earth's environment, particularly its natural and cultural resources, through the use of photographs, and related data acquired from an aircraft or satellite" (14, p. 1).

The term "aerospace remote sensing" is generally defined to refer to remote sensing from space, especially from earth-orbiting or geo-synchronous satellites. Although some reference to uses of conventional aerial photography will be made in this paper the primary purpose is to describe how aerospace remote sensing data has been and is being used in the US to estimate land area devoted to various crops.

Specifically the paper presents how data from an earth-orbiting series of satellites (originally called the Earth Resources Technology Satellite (ERTS), now referred to as Landsat) is being used by SRS to estimate crop area in selected parts of the US. There have been four Landsat satellites launched since 1972. Each of these has included a four-channel Multispectral Scanner (MSS) sensor as part of its payload. To date, the MSS has provided the most useful data for SRS purposes. Therefore this paper will be limited to describing research for developing techniques and procedures for using MSS data.

The MSS views a ground scene of approximately 185 Km by 185 Km in area, with a ground resolution of about 80 meters. For each ground resolution element (referred to as a picture element or "pixel") the MSS measures reflected energy in four spectral bands covering two regions of the electromagnetic spectrum: bands 4 and 5 in the visible range; and bands 6 and 7 in the near-infrared. For each "pixel" the MSS data consists of four values, one measurement in each spectral band. For each scene with 0.8×10^7 "pixels" there are 3.2×10^7 values or measurements which must be processed for making crop area estimates (14). This large data processing task has been the focus of much of the research devoted to using this technology.

2. Remote Sensing Research for Applications to Crop Area Estimation

2.1. Initial Research During 1972-77

In 1972, SRS with joint sponsorship by the National Aeronautics and Space Administration (NASA), began research to develop uses of aerospace remote sensing data collected by the MSS on the first Landsat satellite. The objectives of this research were: to develop new methodology for using MSS data to improve agricultural estimates; to provide knowledge of the information content of the MSS data; and to develop computational capabilities for extracting this information. From the very beginning of this research program emphasis by SRS has been on a "digital" analysis approach rather than on an "image or photo-interpretive" approach (12). Primary reasons for this include:

(a) Lack of trained staff for photo or image interpretation, with little or no prospect of obtaining such expertise, and

(b) The requirement for timely processing and release of agricultural information.

SRS staff assigned to the remote sensing research work in 1972 consisted entirely of statisticians and computer programmers (9). Experience and training in remote sensing techniques would have been extremely valuable; however, it was simply not available within SRS at that time. Since then a few staff members trained in remote sensing have been added to the SRS staff so that a multi-disciplinary approach is possible today. Some "image" analysis approaches particularly for land use/land cover analysis have been developed; however, the predominant research is still aimed at developing more efficient digital analysis techniques. This relates to the need for crop area estimates, whether based on traditional ground surveys such as the JES or on remote sensing data, to be prepared and published within a few days or few weeks after the data are collected.

2.1.1. Methodology

Even though the SRS remote sensing research was directed at producing estimates (of area devoted to specific crops), it never had as a primary goal the complete substitution of remote sensing estimates for those based on ground surveys (e.g., the JES). Indeed it was readily apparent at the beginning of the research in 1972, that a more logical approach would be to develop a methodology that would combine the strengths of the two data sources to produce more precise estimates than either could do independently.

The data collected for the JES provided a natural "ground truth" sample, while Landsat in effect provided a complete census of measurements (if the effect of clouds are ignored). Consequently, research during the 1972-77 period focused on developing the necessary techniques and procedures for integrating these two data sets (JES data with Landsat data) to produce more precise estimates of area devoted to specific crops.

Since the segments enumerated for the JES represent a probability sample of the entire land area of the U.S. a major thrust of the SRS remote sensing research was to use the "inference" capability of this data for processing and analyzing the Landsat data in order to make statistically reliable estimates[1] of crop area. This included the development of registration[2] methods for precise matching of the common data from the two sources, followed by use of a subsample of the matched data to estimate discriminate function parameters for classification[3] of Landsat pixels into desired land use or crop type categories. Once this was accomplished the discriminant functions were tested by classifying another subsample of matched pixels into categories. Results of this classification, while never completely accurate when compared with the ground enumerated crop area data, were sufficiently accurate to bear a close relationship with the ground data. This relationship was used to estimate the parameters of the regression estimator described by Cochran (5). Classification results for entire Landsat scenes or combination of scenes were used in this estimator to make estimates of area devoted to each crop or category of interest. Also, an effectiveness criterion for measuring the success of using Landsat data was relatively easy to derive from this approach. This criterion, the ratio of the variance of the crop area estimate from the ground data only, to the variance of the crop area estimate from the Landsat regression estimator, is generally referred to as the relative efficiency (RE). RE is the multiplier that would be necessary to increase the sample size of the JES in order for the estimates from the ground data alone to be as precise as the regression estimator using both the ground and remote sensing data. When total costs of the two approaches are considered results currently indicate that RE's $\geq$ 2.0 are necessary for the use of Landsat data to be cost-effective.

2.1.2. Development of Computing Capability

While many of the early studies for developing techniques and procedures for using Landsat data for crop area estimates were conducted by using ad-hoc computational resources, it was apparent from the very beginning that a significant investment in data processing, especially software, would be required. The sheer volume of data (3.2×10^7 values for each Landsat scene) plus the statistical computations involved in the multivariate and discriminate analysis, the classification of each pixel, and the final regression estimation of each land use or crop type category, all combined to mandate a significant effort to develop an advanced computational capability.

Starting in 1974, SRS and the Center for Advanced Computation (CAC) at the University of Illinois entered into a joint agreement to develop a software system to process Landsat digital data. An interactive version of Purdue University's LARSYS (one of the first major software systems developed in the US for processing spectral data) (13) developed by CAC and the U. S. Department of the Interior was used as a starting point for the proposed software system. The system (EDITOR) was redesigned to analyze SRS's ground-gathered data from the JES, jointly with the Landsat digital data, to classify all pixels in entire Landsat scenes, and finally to calculate the regression estimates for the crops or land use categories of interest. The EDITOR system was designed to provide a teleprocessing network capability, using computers of different types and sizes located at several locations. Originally the network used included a UNIVAC 1108 on the INFONET network, an IBM 370 at the USDA's Washington Computer Center (WCC), an IBM 370 at CAC in Illinois, a PDP-10 at the Bolt, Beranek, and Newman (BBN) Data Processing Center in Cambridge, Massachusetts, and finally the ILLIAC-IV at the NASA Ames Research Center (ARC) in California. Both the BBN computers and the ILLIAC-IV were on the ARPANET, a teleprocessing network developed by the Department of Defense (DOD). A detailed discussion of the initial capabilities of the EDITOR system is included in a paper written by Ozga, M. et al in 1977 (15). As a system for remote sensing research EDITOR has continued to evolve into a more efficient system for operational processing. This evolution included changes in both the EDITOR software and the computers used for the system. The UNIVAC 1108 has been replaced by an IBM 370 system operated by Martin Marietta Data Systems, which currently has the SRS teleprocessing contract, and the ILLIAC-IV at ARC has been replaced by a CRAY XMP. Almost all of the algorithms in the EDITOR software have been improved or replaced by new algorithms as these have been developed through the SRS remote sensing research program. A partial list of these improvements would include new algorithms for registration (2, 6), classification (3), and statistical estimators (4, 18). Evolution of the EDITOR system is still continuing as new hardware becomes available and the software is being converted into more easily transportable code. Currently a major effort by SRS and ARC is underway to convert the EDITOR code from the original ad-hoc programming languages (SAIL, RATFOR, Assembler, etc.) used for research purposes into PASCAL. Also a new version of EDITOR (called mini-EDITOR) is being developed for use with 16-bit microcomputers. Although data through-put will not be as fast for a microcomputer system it may provide a cost-effective replacement for some functions of EDITOR on the BBN computers and the CRAY. Positive results of remote sensing research by SRS over the past decade are thoroughly integrated by the EDITOR system. Consequently by making the software transportable it will be easier for SRS to share this technology with other users.

2.2. Iowa Project - 1978

Many small and large remote sensing research studies were conducted by SRS in the 1972-77 period to develop methodology and computing capabilities to implement the methodology. These studies were done over agricultural areas in South Dakota, Kansas, Missouri, Idaho, and California (10, 19). Major demonstration tests covering the entire states of Illinois in 1975 and Kansas in 1976 (7, 8) provided significant new evidence of the usefulness of Landsat data for crop area estimation.

Based on results of previous research, SRS decided to demonstrate the EDITOR System by processing Landsat data for the entire state of Iowa in 1978 (11). The goal of this project was to acquire the Landsat data during the optimum time period (late July-early September) for separating corn (maize) and soybeans and to process the data in time for use by the Crop Reporting Board (CRB) in establishing the final 1978 estimates for these crops. Corn and soybeans are the two most important crops in Iowa. Together they accounted for about 87 percent of the area planted to all crops in 1978.

In 1978 the entire Landsat system operated by NASA was still deemed experimental. Because of this, real-time data delivery after satellite acquisition was not possible. Median delivery time after acquisition of Landsat digital data was 49 days with a range from only 32 days to as much as 93 days. Still even with this constraint the goal for the Iowa project was met, and the CRB used the Landsat regression estimates in preparing the Annual Crop Inventory report for 1978.

Because the Landsat satellites are in a polar orbit they pass over the US in a northeast to southwest direction so that data are collected for a given area from east to west.

Figure 1 shows this for the state of Iowa in 1978. With this orbit configuration confounded by "cloud cover" Landsat data for six different dates were required to cover Iowa in 1978. The twelve Landsat scenes are listed in Table 1.

Table 1: Landsat Scenes Used for 1978 Iowa Project

Path	Row	Date	Percent Cloud Cover	Scene ID
30	30	August 19	0	30167-16274
	31	August 19	0	30167-16280
29	30	August 9	0	21295-16013
	31	August 9	40	21295-16020
	32	August 18	0	30166-16224
28	30	September 4	60	30183-16162
	31	September 4	0	30183-16164
	32	September 4	0	30183-16171
27	30	August 7	10	21193-15500
	31	August 7	15	21293-15502
	32	August 7	10	21293-15505
26	31	August 6	0	21292-15444

Even by acquiring data over as many dates as possible some areas of the state were not included in the Landsat analysis due to cloud cover. These areas are shown in Figure 2, which also shows the analysis districts for which regression estimates were computed in 1978. Analysis districts are formed to follow minor civil division boundaries (counties) which are completely covered by a Landsat scene or a Landsat pass (consecutive scenes on the same day). This approach is required because as the atmospheric and vegetation conditions change from day to day reflected energy as measured by the MSS will vary for the same areas. SRS has found that these "signatures" are not extendable over time, therefore the analysis district concept has been included in the SRS remote sensing methodology, with those districts for which no Landsat data are available being estimated by the JES ground data only.

Crop area estimates for corn and soybeans were computed at the state, analysis district, and individual county levels. Results are shown in Tables 2 and 3 for the state and analysis district levels. Improvements in precision for the regression estimate (Landsat and ground data) versus the direct expansion estimate (JES ground data only) were substantial, as measured by the relative efficiencies shown in these tables. As shown in Figure 2, Landsat data were not available for 13 of the 99 Iowa counties during the late July to mid-September period in 1978. For these counties estimates from the JES were added to the other analysis district estimates to produce the total estimate for the state.

Coefficients of variation for the individual county level estimates ranged from 7 to 60 percent; however, other small area estimators for counties are being developed which will hopefully provide county level estimates with relative mean square errors of less than 20 percent. The state level Landsat estimates were used by the USDA's Crop Reporting Board in preparing the Annual Crop Summary for Iowa. Although not the only information used for this summary the remote sensing input influenced the Board to change the estimates for both corn and soybeans slightly from what the other data indicated. Subsequent evaluation based on check data (marketings, exports, etc.) indicated that the changes based on Landsat estimates were toward the "true" values for these crops.

Figure 1

LANDSAT Scene Locations

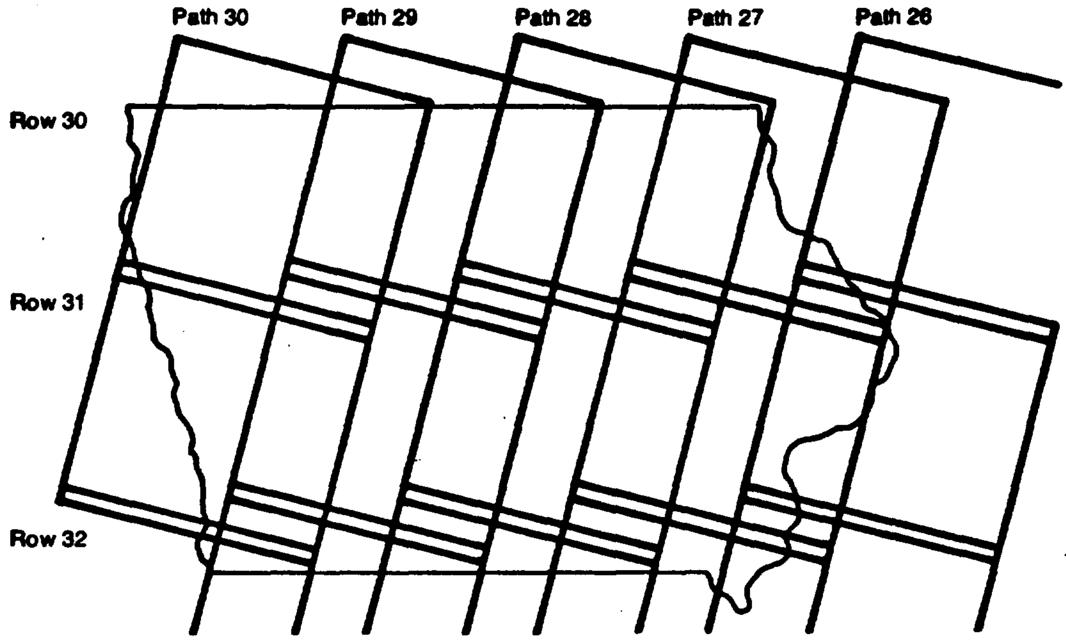

Figure 2

Iowa Analysis Districts

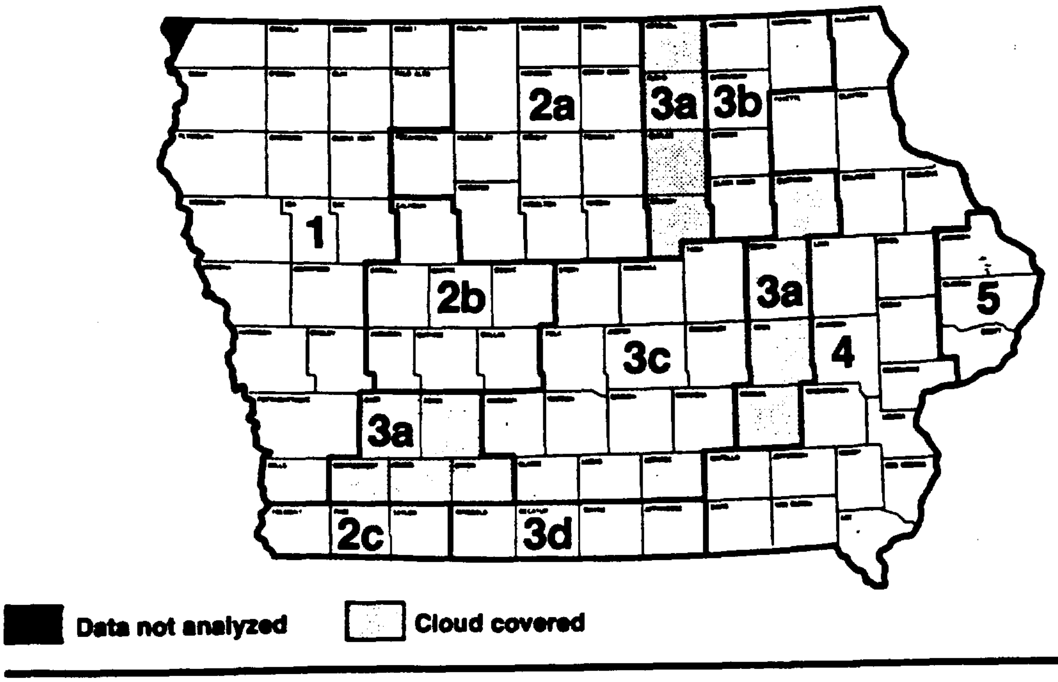

Table 2: Estimates of Corn Area Planted, Coefficients of Variation, and Relative Efficiencies by Analysis District, Iowa, 1978

Analysis District	Classified Pixels 1/	JES Estimates: Area Planted	JES Estimates: Coefficient of Variation	Landsat Estimates: Area Planted	Landsat Estimates: Coefficient of Variation	Relative Efficiency
	(1,000 Hectares)	(1,000 Hectares)	(Percent)	(1,000 Hectares)	(Percent)	
1	1;306	1,462	3.5	1,460	2.2	2.5
2A	924	829	4.5	819	2.5	3.3
2B	464	332	11.5	454	3.4	6.0
2C	125	106	11.0	110	9.5	1.2
3A	2/	657	4.4	2/	2/	2/
3B	345	276	10.0	268	8.5	1.5
3C	590	551	7.5	542	6.0	1.6
3D	59	84	17.8	83	18.6	.9
4	1,059	1,030	6.7	896	4.5	3.0
5	132	148	11.1	150	6.0	3.3
State Total	5,661 3/	5,475	2.3	5,439 3/	1.5	2.4

1/ Converted to hectares.
2/ Landsat data not available.
3/ Includes JES estimate for analysis district 3A.

Table 3: Estimates of Soybean Area Planted, Coefficients of Variation, and Relative Efficiencies by Analysis District, Iowa, 1978

Analysis District	Classified Pixels 1/	JES Estimates: Area Planted	JES Estimates: Coefficient of Variation	Landsat Estimates: Area Planted	Landsat Estimates: Coefficient of Variation	Relative Efficiency
	(1,000 Hectares)	(1,000 Hectares)	(Percent)	(1,000 Hectares)	(Percent)	
1	760	748	8.1	782	4.0	3.7
2A	650	655	6.8	675	3.4	3.7
2B	244	257	12.9	256	6.1	4.6
2C	94	95	25.0	97	11.7	4.4
3A	2/	402	9.2	2/	2/	2/
3B	84	87	28.0	125	9.4	4.3
3C	370	329	14.5	338	7.1	4.0
3D	79	83	32.6	96	10.2	7.6
4	343	441	12.7	425	8.0	2.7
5	35	47	29.2	49	12.5	5.1
State Total	3,060 3/	3,144	3.9	3,245 3/	2.5	2.4

1/ Converted to hectares.
2/ Landsat data not available.
3/ Includes JES estimate for analysis district 3A.

3. AgRISTARS

Considerable detail has been provided for the 1978 Iowa project, because it produced for the first time quasi-operational estimates based on use of Landsat data. It provided a means of integration of much of the research completed at that time and it provided a partial base from which the USDA could start a major new remote sensing program in 1980. This program was called "Agriculture and Resource Inventory Surveys Through Aerospace Remote Sensing" (AgRISTARS) (1).

3.1. Departmental Program

The AgRISTARS program was initiated in fiscal year 1980 in response to an initiative issued by the USDA. Led by USDA, the program was developed as a cooperative effort with NASA, the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce (USDC), the U. S. Department of the Interior (USDI), and the U. S. Agency for International Development (AID).

The program goal was to determine the usefulness, cost, and extent to which aerospace remote sensing could be integrated into existing or future USDA systems to improve the objectivity, reliability, timeliness, and adequacy of information required to carry out USDA missions. The program was designed to provide research for addressing a broad range of information requirements for both international and domestic programs concerned with monitoring and inventory of renewable resources. Included were studies of using remote sensing data for global crop production forecasts and estimates, for rangeland and forest inventories, for land productivity analysis, for conservation practices inventory, and for pollution detection and monitoring. The program which is currently planned to continue through 1986 has produced many research results that are now in daily operational use by USDA agencies.

3.2 SRS Role

Because of previous remote sensing research, SRS was selected by USDA as a major participant in AgRISTARS, in both coordination and research roles. Under the research role, SRS has been primarily concerned with applications research for domestic (US) crop estimates. The major contribution of SRS to this program has been the continued refinement of concepts developed during the 1972-77 period, and the proof of concept demonstrated in the 1978 Iowa project. These refinements have been integrated into the SRS methodology which has been used for quasi-operational Large Scale Application Tests (LSAT) in 1980, 1981, 1982, and 1983. Under the AgRISTARS program, SRS plans were to conduct LSAT's over two states (Iowa and Kansas) in 1980, with an additional two states added each year until by the end of the program at least ten states would be included. This goal was met for 1981 and 1982; however, due to budget restraints, only one new state was added in 1983 and no new states are planned for 1984. In 1984, this quasi-operational LSAT will be completed for the following seven states: Iowa, Kansas, Oklahoma, Missouri, Colorado, Illinois, and Arkansas. Primarily corn and soybeans estimates will be prepared for Iowa, Illinois, and Missouri; winter wheat estimates for Kansas, Oklahoma, and Colorado; and cotton, soybeans, and rice estimates for Arkansas. While results from these LSAT's have been quite varied due to many problems with the Landsat program, with data delivery, with cloud cover, etc., the resulting estimates have provided a useful input to the CRB in preparing more accurate crop area estimates for these states. Operational procedures have been improved and costs reduced to the point where the methodology is cost-effective when compared to comparable results by simply expanding the JES ground survey.

4. Future Use of Aerospace Remote Sensing

The future for aerospace remote sensing for domestic crop estimates is not clear at this time. Primarily, because of uncertainties associated with the US land remote sensing program. These include questions about whether the program (if there is one) will be operated by the government or by private firms, about whether sensor and data delivery improvements can be made, and whether or not total costs of a remote sensing system for agricultural estimates will continue the downward trend established during the past decade (SRS costs on a per state basis for applying this methodology has been reduced by a factor of seven since 1975). Those involved with this research in SRS feel that the potential of aerospace remote sensing for US crop estimates has been demonstrated and plans are currently being made to refine the present methodology so that by the end of AgRISTARS in 1986, SRS managment will be able to make a decision whether or not to implement the technology on an ongoing operational basis. This decision will undoubtedly be influenced not only by the US land remote sensing program (private or public) but also by similar programs in other nations, such as France with its "Systeme Probatoire d' Observation de la Terre" (SPOT) satellite, and Japan with its Earth Resources Satellite (ERS 1) (14)--satellites for land remote sensing which are planned for launch in 1985 and 1986, respectively.

Summary

Agriculture has long been identified as the field where applications of aerospace remote sensing technology has the potential of making the largest impact in terms of producing new information or improving existing information systems (16). The SRS of USDA over the past decade has viewed aerospace remote sensing in the latter context viz., that it represents a new source of data particularly on agricultural crops, which should be exploited for improving the current agricultural statistics program. Consequently, beginning soon after the launch of Landsat 1 in 1972, SRS has pursued an active research program to develop methodology and data processing systems for using Landsat digital MSS data for improving the existing US program of crop area estimates. After five years of developmental studies, a major pilot test of the resulting methodology and data processing system was conducted in 1978 for the entire State of Iowa. Based on results of this test, which are described in this paper, SRS, under the AgRISTARS program,

developed a plan for a series of LSAT's beginning in 1980, designed to further demonstrate the usefulness of Landsat data for crop area estimates. This plan included additional research to refine and enhance the methodology for possible operational use after the end of AgRISTARS in 1986. The 1983 LSAT covered seven (Kansas, Colorado, Iowa, Oklahoma, Missouri, Illinois, and Arkansas) major crop states. Although results have been variable, due to the quasi-research nature of the test, quality of Landsat data, and problems with Landsat data delivery, the estimates have been significantly more precise than those based on ground surveys and have proven useful in establishing final end-of-year estimates for crops grown in the seven states.

Although the LSAT's will be continued through 1986, the decision to proceed with an operational remote sensing system for US crop area estimates will also depend on the status of land remote sensing systems in operation, or anticipated, at that time. The usefulness of aerospace remote sensing for crop area estimates has clearly been demonstrated for Landsat MSS type data. It is expected that data from future sensors, such as the one planned for the French SPOT satellite, will prove even more useful. However, costs of data from these future systems are still to be determined and their ultimate use by SRS is still questionable. Current prospects would seem to indicate at least a continued use by SRS of aerospace remote sensing data for information on US crops. Expansion to other states and/or crops will be limited by scarce resources but will probably continue to occur as research provides further gains in productivity for using aerospace remote sensing.

Notes

[1] Direct expansion estimator ($\hat{Y}_{DE}$) and the regression estimator ($\hat{Y}_R$) used by SRS

Y = Population total (for example, total area planted to corn)

$\hat{Y}_{DE}$ = Estimated total (using JES ground data only)

$$= \sum_{i=1}^{L} N_i \left(\sum_{j=1}^{n_i} y_{ij} \right) / n_i$$

where,

L = the number of land use strata in the state.

N_i = population number of sampling units (su) in stratum i.

n_i = sample number of su in stratum i.

y_{ij} = area planted to corn in the j^{th} su of the i^{th} stratum.

Then the variance of $\hat{Y}_{DE}$ is:

$$V(\hat{Y}_{DE}) = \sum_{i=1}^{L} \frac{N_i^2}{n_i(n_i-1)} \cdot \frac{N_i-n_i}{N_i} \cdot \sum_{i=1}^{L} (y_{ij}-\bar{y}_i)^2$$

where,

$$\bar{y}_i = \sum_{j=1}^{n_i} y_{ij}/n_i$$

$\hat{Y}_R$ = Estimated total (using Landsat and JES data)

$$= \sum_{i=1}^{L} N_i \cdot \bar{y}_i(reg)$$

where,

$$\bar{y}_i(reg) = \bar{y}_i + b_i \; (\bar{X}_i - \bar{x}_i)$$

and

b_i = the estimated regression coefficient for the i^{th} stratum when regressing ground reported corn area on classified pixels for the n_i su.

$\bar{X}_i$ = the population average number of pixels of corn per su in the i^{th} stratum.

$\bar{x}_i$ = the sample number of pixels of corn per su in the i^{th} stratum.

Then, the estimated (large sample) variance of $\hat{Y}_R$ is:

$$V(\hat{Y}_R) = \sum_{i=1}^{L} \frac{N_i^2}{n_i} \cdot \frac{N_i - n_i}{N_i} \cdot \frac{1 - r_i^2}{n_i - 2} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

where,

r_i^2 = sample coefficient of determination between reported corn area and classified corn pixels for the n_i su in the i^{th} stratum.

$$r_i^2 = \frac{\left(\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(x_{ij} - \bar{x}_i)\right)^2}{\left(\sum_{i=1}^{n_i} (y_{ij} - \bar{y}_i)^2\right) \left(\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2\right)}$$

Note that

$$V(\hat{Y}_R) = \sum_{i=1}^{L} \frac{n_i - 1}{n_i - 2} (1 - r_i^2) \; V(\hat{Y}_{(DE)i})$$

So that lim $V(\hat{Y}_{Ri}) = 0$ as $r_i^2 \rightarrow 1$ for fixed n_i.

The relative efficiency of $\hat{Y}_R$ compared to $\hat{Y}_{DE}$ is defined as follows:

$$R.E. = \frac{V(\hat{Y}_{DE})}{V(\ddot{Y}_R)}$$

[2]Registration

As used in this paper, registration refers to the procedures for determining the geographic location of each pixel in a given Landsat scene. It is the process of matching the Landsat row-column coordinates with map latitude-longitude coordinates. The process involves finding common points, such as road and/or rail intersections, on both Landsat images and index maps. By digitizing map coordinates of these points, a mathematical transformation can be calculated to predict the latitude-longitude coordinates of all pixels in the Landsat scene. Further precision adjustments are made to accurately locate the JES segments field-by-field in the Landsat scene. Over time this precision adjustment has evolved from a largely manual approach to one that uses computer algorithms to detect field edges within the Landsat data. The ultimate goal of registration is to locate JES segments in the Landsat data to an accuracy of $\pm$ 0.5 pixel.

[3]Discriminant Analysis - Classification

The basic use of discriminant analysis is to assign an observation to one of two or more groups on the basis of its value. In this paper the observation is the "pixel" and its value is the vector of measurements recorded by the MSS on Landsat. The multivariate measurements represented by this vector were used to assign (classify) each pixel into a crop or land use type (corn, soybeans, cities, water, etc.).

A sample of fields from the JES for each crop or land use type was selected and the MSS data were summarized to obtain the distributions for each type. With four measurements for each pixel the measurement space (MS) is four-dimensional. This largely precludes a visual partitioning of the MS by simple visual means which may be appropriate for separating observations into groups for one or two dimensions. With four-dimensional data, boundaries between groups become contour "surfaces" in the MS. These dividing surfaces are constructed so that observations falling on the "surface" have equal probabilities of being in either group. Those observations not on the dividing surface always have a greater probability of being classified into the group for which the observation is interior to the contour surface.

For SRS applications, discriminate functions have been computed by assuming the MSS measurements are from a multivariate normal density distribution. Departures from normality has not been a serious problem as long as precautions have been taken to insure unimodal data. This has meant that in certain cases the data analyst has to recognize, usually by using cluster analysis, that the same crop may fall into more than one group, e.g., MSS data for late planted corn may follow an entirely different distribution than corn planted early. After assuring that the normality assumption is approximately satisfied, a multivariate normal distibution was estimated for each group. This was based on a sample of fields from the JES for each crop or land use group. This sample was used to calculate mean vectors and variance-covariance matrices for each crop group. Using these sample statistics, the quadratic discriminate function was calculated. This function was used to classify all unknown pixels into one of the crop groups for which the crop mean vector is closest to the point based on the Mahalanobis distance. That is, the crop for which the probability is highest.

References

1. "AgRISTARS Technical Program Plan," January 15, 1980, AP-JO-00628, JSC-17153, Lyndon B. Johnson Space Center, Houston Texas 77058.
2. Allen, R. D., "USDA Registration and Rectification Requirements." Proceedings of the NASA Workshop on Registration and Rectification, Nevin A. Bryant, Editor, JPL

Publication 82-83, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, June 1, 1982.

3. Amis, M. L., and others, "Evaluation of Large Area Crop Estimation Techniques Using Landsat and Ground-Derived Data." AgRISTARS Report DC-L1-04051, JSC-17116, LEMSCO-15763, March 1981.

4. Cardenas, M., M. Blanchard, M. Craig, "On the Development of Small Area Estimators Using Landsat Data as Auxiliary Information," Research report, ESCS, USDA, Washington, D. C., 1978.

5. Cochran, W. G., 1977. Sampling Techniques - Third Edition, John Wiley and Sons, Inc., New York.

6. Cook, P. W., "Landsat Registration Methodology, USDA, SRS, 1972-1982." Research report, Remote Sensing Branch, Statistical Research Division, SRS, May 1982, Washington, D. C.

7. Craig, M., R. Sigman, and M. Cardenas, "Area Estimates by Landsat: Kansas 1976 Winter Wheat," Research report, ESCS, USDA, Washington, D. C., 1978.

8. Gleason, C., and others, "The Auxilliary Use of Landsat Data in Estimating Crop Acreages: Results of the 1975 Illinois Crop Acreage Experiment." Research report SRS-21, USDA, Washington, D. C., October 1977.

9. Hanuschak, G. A., R. D. Allen, and W. H. Wigton, "Integration of Landsat Data into the Crop Estimation Program of USDA's Statistical Reporting Service, 1972-1982," Proceedings Eighth International Symposium, Machine Processing of Remotely Sensed Data with Special Emphasis on Crop Inventory and Monitoring, Purdue University, Laboratory for Applications of Remote Sensing, July 7-9, 1982.

10. Hanuschak, G., and K. Morrissey, "Pilot Study of the Potential Contribution of Landsat in the Construction of Area Sampling Frames," Research report, SRS, USDA, Washington, D. C. 1977.

11. Hanuschak, G., and others, "Obtaining Timely Crop Area Estimates Using Ground-Gathered and Landsat Data." Technical Bulletin 1609, ESCS, USDA, Washington, D. C., August 1979.

12. Huddleston, H. F., and W. H. Wigton, "Use of Remote Sensing in Sampling for Agricultural Data," ISI/IASS Invited Paper, Warsaw, 1975.

13. Lindenlaub, J. C., "The LARSYS Software System--An Overview," 136 p. Purdue University, Laboratory for Applications of Remote Sensing, West Lafayette, Indiana.

14. Manual of Remote Sensing, Second Edition, Volume 1, Published by American Society of Photogrammetry, 1983.

15. Ozga, M., W. Donovan, and C. Gleason, "An Interactive System for Agricultural Acreage Estimates Using Landsat Data." Proceedings 1977 Symposium on Machine Processing of Remotely Sensed Data, Purdue University, Laboratory for Applications of Remote Sensing.

16. Remote Sensing, with Special Reference to Agriculture and Forestry, Committee on Remote Sensing for Agricultural Purposes, Agricultural Board, National Research Council, National Academy of Sciences, Washington, D. C., 1970.

17. Scope and Methods of the Statistical Reporting Service, U.S. Department of Agriculture, Miscellaneous Publication No. 1308, Revised September 1983.

18. Sigman, R., and G. Walker, "Comparison of Auxiliary Variables for Satellite-Data Regression Estimators in Crop Surveys." Contributed paper, 1983 American Statistical Association Annual Meeting, Toronto, Ontario, Canada, August 1983.

19. Von Steen, D., and W. Wigton, "Crop Identification and Acreage Measurement Utilizing Landsat Imagery," Research report, SRS, USDA, Washington, D. C., 1976.