Paper A 7

REGIONAL AGRICULTURE SURVEYS USING ERTS-1 DATA

William C. Draeger, James D. Nichols, Andrew S. Benson, David G. Larrabee, William M. Senkus and Claire M. Hay,
Center for Remote Sensing Research, University of California, Berkeley, California 94720

ABSTRACT

During the past year, the Center for Remote Sensing Research (CRSR) at the University of California has conducted studies designed to evaluate the potential application of ERTS data in performing agricultural inventories, and to develop efficient methods of data handling and analysis useful in the operational context for performing large area surveys.

This work has resulted in the development of an integrated system utilizing both human and computer analysis of ground, aerial, and space imagery, which has been shown to be very efficient for regional crop acreage inventories. The technique involves (1) the delineation of ERTS images into relatively homogeneous strata by human interpreters, (2) the point-by-point classification of the area within each strata on the basis of crop type using a human/machine interactive digital image processing system, and (3) a multistage sampling procedure for the collection of supporting aerial and ground data used in the adjustment and verification of the classification results.

INTRODUCTION

In the United States, the Department of Agriculture presently conducts an enumerative program in which virtually all agricultural land is inventoried annually. In addition, numerous other federal, state and local agencies conduct extensive crop inventories, land use surveys, and soils mapping projects of varying magnitude. On a worldwide basis it would seem that the principal obstacles to providing enough food for all persons are not merely ones of production but also problems of allocation and distribution. What is needed is knowledge as to where and how much food is now being produced, and how crop production is changing with time. Considering the present needs for regional, national, and worldwide inventory and evaluation data, coupled with the particular capabilities of the ERTS system, agricultural applications appear to be especially promising as an area in which important benefits might be realized from the use of such technology.

Up to the present time agricultural inventories have required a tremendous effort on the part of on-the-ground enumerators, and have presented a formidable data compilation task. However, a satellite sensing system, with which large areas of land can be surveyed in their entirety on one image, and which can provide uniform worldwide coverage with a relatively small number of images, offers great promise as a data collection tool for alleviating these problems. Furthermore, the dynamic nature of agriculture requires not a single evaluation in most cases, but rather a continual updating of conditions. In fact, it has been shown that desired information about agricultural crops can often be obtained only by capitalizing on a knowledge of the patterns of change exhibited by particular crop types under various growing conditions. Again, this suggests that a satellite sensing system such as ERTS, which makes possible regular, frequent observations of each spot on the earth's surface, can provide a service which is both highly desirable and totally infeasible using conventional techniques.

Based on these facts, plus the encouraging results achieved using both high altitude aircraft and spacecraft imagery for crop inventory experiments over the past several years, the Center for Remote Sensing Research (CRSR) at the University of California undertook the experiment described in this paper. The experiment was designed to evaluate the feasibility of using satellite data regionally to provide needed agricultural information on an operational basis. The experiment was performed in Maricopa County, Arizona and San Joaquin County, California in cooperation with a number of state and federal agencies.

In an effort to accurately determine the amount of detail which can be extracted from ERTS-1 data, and the optimum use of "subsampling" in the form of aerial photography and ground truth data for various agricultural-related tasks, the investigation was carried out in a stepwise fashion beginning with agricultural land use stratification, and progressing to very detailed surveys. These investigations entailed the use of both human image analysts and automatic classification and data handling techniques, and an evaluation of the optimum mix of human and machine techniques for each analysis problem. In each case, an attempt was made to ensure that the types of information compiled (e.g., maps, tabular data, crop acreages, etc.) conformed to actual requirements or desires as expressed by those persons currently involved in resource evaluations and planning in the test site.

In the area of agricultural land stratification, particular attention was paid to quantitative analyses of the stratifications to ascertain the extent to which they did provide meaningful crop type and condition information. In so doing, use was made of ground cell information, point sampling along transects (using observers flying in light aircraft), and comparisons with existing land classifications using the CRSR MAPIT techniques. In addition, a study was made of the variation in delineations made at different times during the growing season.

Crop classification and inventory studies progressed concurrently with the stratification investigations, as it seemed likely that any operational inventory procedures would be heavily dependent on an initial stratification of agriculture on ERTS or other small-scale imagery. Very significant progress was made on the development of automatic data processing techniques for the detailed classification of agricultural lands. In particular, emphasis was placed on the optimum interface between human interpreters and the computer. Thus, initial stratifications of agricultural land were performed manually, and the resultant information used in the classification process. In addition, a sampling procedure was designed which would optimize processing of remote sensing and ground data by reducing the amount of ground information required. Incorporated in the design is provision for the weighting of classification errors based on the relative importance of errors regarding various crops. Finally, studies were conducted to estimate the relative costs of performing crop inventories using various combinations of human and computer data processing inputs.

A PRACTICAL INVENTORY SYSTEM

The principal result of the ERTS-1 investigation was the development of an integrated data handling and analysis system which utilizes both human and computer analysis of ground, aerial and space imagery and which has been shown to be very efficient for regional crop acreage inventories.

The initial step in the inventory process consists of the delineation of ERTS images into relatively homogeneous strata by human interpreters. The total image is first divided into agricultural and non-agricultural areas, after which the agricultural areas are subdivided on the basis of predominant crop types using only gross image characteristics and a general knowledge of cropping practices.

In an attempt to evaluate the use of satellite imagery for this purpose, all land within San Joaquin County was delineated by image analysts into broad land use and crop category classes based on their appearance on the ERTS-1 July 26 (summer season) color composite image. The stratification of the agricultural land use categories proved to be a relatively simple task, taking each of three interpreters approximately 30 minutes to complete. The three interpretations were quite similar, requiring only minor revisions to produce a "consensus" stratification. A total of thirteen different agricultural strata were recognized, differing both in general field size and relative proportions of crop types and field conditions. Upon comparing these interpretations we concluded that nearly all boundaries were truly representative of differing cropping practices. In a number of cases, the stratifications agreed almost exactly with major soil type boundaries as drawn by earlier soils surveys.

Certainly a much more detailed and up-to-date stratification was produced from the ERTS image than is currently used by the Statistical Reporting Service, USDA. The obvious questions arise, however, as to whether: (1) the strata as drawn on the image are meaningful in terms of actual land use conditions, (2) the strata delineations would change throughout the year, and (3) such a detailed delineation could enable the Statistical Reporting Service to more efficiently and accurately estimate the parameters of interest on a statewide basis.

In addition to their possible use by agencies such as the Statistical Reporting Service, the stratifications performed by the human interpreters proved to be of great value as the preliminary step prior to detailed classification of crop types on a field by field basis, since it was found that by far the most practical and cost-effective method for producing "automated crop inventories" involves the use of manual interpretation at several stages in the process. In particular, it has been found that automatic classification done stratum by stratum, using training data specific to each stratum, results in much greater classification accuracy than would be possible otherwise. Furthermore, it is much more efficient to allow a human to do the preliminary stratification than to attempt this with automated techniques. Thus an interactive man-machine system, in which each is used to perform only those tasks for which it is best suited, results in the greatest overall efficiency in terms of time and money expended for a given level of classification accuracy.

At this point in the inventory process, political and administrative boundaries may also be superimposed on the imagery to define the geographic area of interest. Next, to train the discriminant analysis program, fields identified by ground data or photo interpretation representing the various resource or vegetation types of interest in each stratum are located on small-scale photos for extraction from the digital tapes. The number of training fields required for each crop class depends on the variability of the spectral signature of the various crops present. This variability is caused by such factors as different cropping practices, local soil differences, and genetic variations within a particular crop type. For a crop such as alfalfa where there may be several stages of maturity present at the time of image acquisition, five or more fields per stratum may be required. In the case of less complex crop classes such as corn, one training example may be adequate. These training fields must also be large enough to be identifiable on the imagery acquired by the remote sensing system used in the first stage. On ERTS-1 imagery the minimum area is around 20 acres with a minimum dimension on one side of 800 feet. These fields are identified on and extracted from the spacecraft imagery and supplied as training to the discriminant analysis system. The multispectral data are then run through the discriminant analysis to obtain a point-by-point classification of the entire area by strata (as defined by the human interpreter). This provides an initial estimate of the acreage of the vegetation classes by strata.

The discriminant analysis results must then be sampled in some manner to determine the relationship between the discriminant analysis estimate and the true value or ground estimate of the resource. Sampling units are defined by breaking the entire area into rectangular areas which in the case of the ERTS study were based on the coordinate grid generated by the MSS system. The size and shape of each rectangular area are determined by the information requirements of the manager, the change in variability of the estimates for the SUs as their size is changed, the cost of making further estimates on the SUs, and the difficulty of recognizing the sampling units on conventional larger-scale imagery.

To evaluate the relative utility of the discriminant analysis of ERTS-1 multispectral-multidate imagery in estimating the area of agricultural crops the information obtained from the discriminant analysis, ground data and high flight imagery of the intensive test site in San Joaquin County were used to determine the optimum size of the sampling unit and the number of samples required to obtain acceptable estimates of crop area for the entire county. The optimum size of the primary (first stage) sampling unit was found to be 25 x 35 picture elements (equivalent to 386 hectares on the ground). This was determined from the estimates of the coefficient of variation, and the plot of expected error in transferring the ERTS sampling units to the corresponding photography for precise area measurement.

At this point there are two basic models that can be applied to estimate the number of sampling units to comprise the second stage sample. If an estimate of the quantity of a resource present is needed, and if it is found that the variance of the estimate is proportional to the value of the resource, then probability sampling will generally be the most efficient model. If, however, inplace mapping is desired, a sampling scheme using regression estimation to establish the relationship between the discriminant analysis estimate and the ground estimate for the resource is used. Therefore, the sampling units for the second stage are selected using information derived in the first stage initial classification, thus reducing the amount of aerial and ground data needed.

The second stage of the model is based on aerial photography of the selected sampling units on which precise field size measurements can be made. In cases where only surface area cover estimates are needed, the second stage imagery and associated ground data are all that are needed. In other cases where estimates of yield per unit area are required, three or more stages may be required to obtain adequate information. When the "correct" area and classification for each field in the sampling units has been determined, this information is used to adjust the estimates obtained in the initial classification.

CONCLUSIONS

The techniques and results discussed in this paper have their real significance in that they indicate the very real possibility of eventually performing operational agricultural surveys on a regular basis using satellite data as a basic input. The particular procedure which has been explained is presented, not as an answer to the entire problem of agricultural surveys using spacecraft data, but rather as an example of how the problem can be approached and as an indication of some of the possible techniques that might be used.

Obviously there are a number of questions that must be answered before the design and feasibility of an operational system can be defined. Among the unanswered questions are those relating to the area effectiveness of training and calibration data (i.e., over how large an area is a given set of training data useful), to the applicability of such techniques to all crops in all parts of the world (and what kinds of adaptations of the techniques might prove necessary), and to the accuracy with which crop yields as well as acreages might be estimated. Furthermore, a limited study of the sort described here working with data from an experimental satellite such as ERTS-1 can only indicate the potential usefulness of such a system. Certainly a commitment by agencies actually involved in the collection of agricultural information, and an operational data processing system providing for the rapid availability of data which is so crucial in the agricultural situation are necessary before the final questions can be answered.

Town or City	1969 Urbanized Area (Acres)	1972 Urbanized Area (Acres)	Amount of Agricultural Land (Acres) Lost
Chandler	1,920	2,880	960
Glendale	4,000	6,240	2,240
Mesa	8,320	13,440	5,120
Phoenix*	56,680	64,680	8,000
Scottsdale*	6,400	9,600	3,200
Tempe	7,200	15,840	8,640

*1972 urbanized area totals do not include sparse residential developments in wildland areas.

Figure 1. Measurements of urban areas in Maricopa County, Arizona made using stratifications of Apollo 9 photos taken in 1969 and ERTS-1 imagery obtained in 1972 indicate the extent to which prime agricultural areas are being converted to urban use.

FIELD BY FIELD RESULTS OF CLASSIFICATION OF TEST AREAS FROM ERTS-I TAPE DATA

STRATUM 16 SAN JOAQUIN COUNTY, CALIFORNIA

		GROUND DATA										TOTAL	COMMISSION ERROR
		ASPA	CORN	HARV	BARE SOIL	POTA	SAFF	SUG. BEET	FL. IRR.	ALFA	TOMA		
CLASSIFICATION	ASPA	126			4	3						133	5
	CORN	7	59									66	11
	HARV			19								19	0
	BARE SOIL	5			5							10	50
	POTA	2				4						6	35
	SAFF	1	1				6					8	25
	SUG. BEET		1			2		5				8	37
	FL. IRR.								3			3	0
	ALFA											0	–
	TOMA		3				1					4	100
TOTAL		141	64	19	9	9	7	5	3	0	0	257	
% CORRECT		89	92	100	56	44	86	100	100	–	–		227/257

TOTAL PERCENT CORRECT - 88.3

Figure 2. Preliminary classification tests on sample areas within a field-crop stratum in San Joaquin County resulted in an overall percent correct identification based on number of fields of 88 percent.

FIELDS CLASSIFIED AS:	GROUND DATA: ASPA	SAFF	SUG BEET	ALFA	TOMA	WAL	CORN	HARV	BARE SOIL	PLOW	PAST	BURN	ROAD	WAT	GRAP	TREE	LIMA	RICE	PEP	SORG	SQUA	POTA	OTHER	TOTAL	COMMISSION ERROR
ASPARAGUS	132	1		4	1				4													3		145	7
SAFFLOWER	1	12					1																	14	10
SUGAR BEETS	.		30	2	4		1			1							5					2		45	31
ALFALFA	1	2	2	69	2						2							2						80	11
TOMATO	18	4	1	3	33		3	2									1				1			66	43
WALNUT						11	2										2		1					16	28
CORN	7						63																	70	9
HARVESTED					1			23			2							1						27	12
BARE SOIL	5		.						5															10	36
PLOWED	2			2				1		18	2				3				1					29	37
PASTURE				3		2					27						2		1					35	20
BURNED												1												1	0
ROAD				3									0											3	100
WATER														10										10	0
GRAPES								2							8			5						15	39
TREES														1		6								7	14
LIMA BEANS				1	1												10							12	9
RICE																		12						12	0
PEPPER																			1					1	0
SORGHUM				1																1				2	20
SQUASH																					1			1	0
POTATO	2																					4		6	18
OTHER				1	1															3			3	8	62
TOTAL	168	19	33	89	43	13	70	28	9	19	33	1	0	11	11	6	20	20	4	4	2	9	*	612	—
% CORRECT	79	63	91	78	77	85	90	82	55	95	82	100	—	91	73	100	50	60	25	25	50	44	(*)	—	[illegible]

TOTAL % CORRECT = 77.9

CROP TYPE	STUDY AREA: CLASSIFIED POINTS	RECLASSIFIED POINTS	RECLASSIFIED W/WEIGHTS	ACRES	% OF TOTAL	TEST EXTRACTIONS: CLASSIFIED POINTS	ACRES	% OF TOTAL
ASPARAGUS	56,371	56,833	57,857	62,070	28.23	17,446	19,190	33.21
HARVESTED	32,103	31,282	30,416	35,377	16.09	5,159	5,675	9.83
CORN	26,662	31,320	31,290	29,375	13.36	8,720	9,592	16.60
SUGARBEET	15,646	14,689	15,046	17,238	7.84	4,309	4,740	8.20
SAFFLOWER	11,089	9,127	9,513	12,225	5.56	2,945	3,240	5.61
WATER	10,189	12,188	10,189	11,235	5.11	2,404	2,644	4.58
POTATO	5,749	3,022	3,137	6,332	2.88	1,866	2,053	3.55
BARE SOIL	1,450	973	1,450	1,583	.72	647	712	1.23
TOMATO	30,409	29,584	30,059	33,487	15.23	7,019	7,721	13.37
ALFALFA	9,932	10,582	10,643	10,950	4.98	1,994	2,193	3.80
TOTALS	199,600	199,600	199,600	219,872	100.00	52,509	57,760	99.98

Figure 3. The table on the left illustrates the results of preliminary classification of over 600 test fields in the field-crop strata in San Joaquin County. On the right are the results of a classification of all of one field crop strata comprising an area of over 200,000 acres (80,000 ha). This essentially comprises a description of the entire stratum in terms of the area classified as each of the various crop types.

	STRAT-IFIED NO CALSCAN	MULTI DATE CALSCAN			CALSCAN NO STRAT-IFICATION
		ONE DATE	TWO DATES	THREE DATES	
STRATIFICATION	120	120	120	120	0
TRAINING EXTRACTION	0	660	660	660	1370
CALSCAN ANALYSIS	0	210	570	1176	8260
SAMPLING GROUND ENUMERATION and CALCULATION	n=82 8835	n=22 2430	n=11 1275	n=9 1060	8835
TOTAL COST	8955	3420	2625	3021	18,465
RELATIVE COST	1.0	.38	.29	.34	2.06

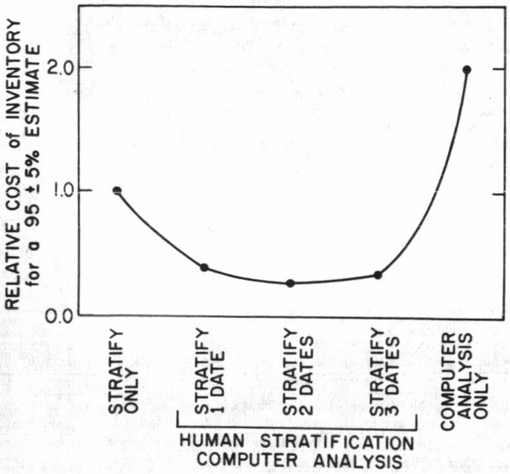

Figure 4. The table and accompanying plot illustrate the estimated relative costs in dollars of performing a crop inventory in San Joaquin County. The estimates assume an accuracy of classification of ±5 percent at the 95 percent confidence level. Note that it would appear that the least expensive survey would involve both initial stratification by human interpreters and computer classification of crops within strata.