Type II Progress Report

December 20, 1972 - June 19, 1973

Crop Identification & Acreage Measurement Utilizing ERTS Imagery 013

Principle Investigator

Donald H. Von Steen AG 328

Introduction

The Statistical Reporting Service (SRS) is interested in the ERTS program as a potential data source for improving current crop acreage estimates. SRS uses ground observations of land area segments across the U.S. in its present agricultural statistics program. These small area segments can be used as ERTS imagery training data for the discriminant analysis. SRS personnel in the Research Division are studying aerial photography as well as satellite imagery to see if it is possible to make a crop inventory of an area by using remotely sensed data.

Personnel and Equipment

With the contract agreement of June 19, 1972, the Research and Development Branch of the Research Division fixed the personnel and equipment for the ERTS project under the direction of principle investigator, Donald H. Von Steen. Paul Hopkins is the EDP specialist. Harold Huddleston, Fred Warren, William Wigton, and Ed Burgess are Mathematical Statisticians who are working on this project. We regret to report that Ron Bosecker and John Ridgely are no longer on the project. The address remains the South Building, Room 4814, but our telephone number has changed to 447-2672. In addition to the equipment of our last report, we have purchased a microdensitometer.

Microdensitometer Purchase

An invitation for bid for a Microdensitometer Data Acquisition System including a Computer Control System was prepared and submitted to GSA for review and delegated procurement authority prior (DPA) to submission to industry. The Solicitation and requested DPA were returned and approved by GSA. Bids were opened on March 15, 1973. Only two bids were received. Both bids were non responsive. One failed to comply to the terms of the solicitation, and the other firm qualified its bid with respect to payment. However, the latter firm met the technical specifications, we asked GSA and received permission to negotiate with them under a new delegation of procurement authority. Negotiations were conducted and a contract signed with Photometric Data Systems Corporation. We expect delivery of the system the 13th of July 1973.

Photography and Digital Tapes

The following tables show the ERTS-1 imagery and digital tapes currently available with useable district coverage.

SOUTH DAKOTA: ERTS PHOTOS & TAPES 1972-73

I.D. Number	Coordinates	Scene Date	70mm rec'd	95" Color Rec'd		CC Tapes Rec'd		Dist. Coverage	Cloud Cover
				Bulk	Precision	Bulk	Precision		
1023-16440		8/15				11/16		SE 40 %	
1024-16491	N44-48 W97-36	8/16				11/9		N 80	20% cc
1025-16545	N44-60 W98-58	8/17	9/22		4/23		2/5	NW 5	10% cc
1025-16551	N43-34 W99-31	8/17	9/22		4/23		2/5	W 10	0% cc
1041-16433	N44-31 W96-21	9/2	10/19	3/29	5/1	11/16	12/19	NE 40	20% cc
1041-16435	N43-05 W96-53	9/2	10/19	3/12	4/25	11/9		SE 30	20% cc
1042-16491	N44 -29 W97-47	9/3	12/28			11/9	N/A	N 80	30% cc
1043-16550	N44-29 W99-21	9/4	10/12	2/26	4/25	11/17	1/10	W 10	10% cc
1060-16491	N44-30 W97-45	9/21	11/2			11/9	6/19	N 80	0% cc
1060-16494	N43-04 N98-17	9/21	11/2			11/15	6/19	SW 25	0% cc
1077-16440	N43-09 W96-46	10/8	11/17	2/20			1/29	SE 30	10% cc
1078-16492	N44-38 W97-40	10/9	11/8			11/9	5/25	N 80	30% cc
1095-16440	N44-29 W96-20	10/26	11/22	2/20	6/1	11/20	2/5	NE 40	10% cc
1095-16442	N43-04 W96-52	10/26	11/22	2/20		11/20	1/8	SE 30	0% cc
1114-16500	N44-32 W97-50	11/14	12/12			1/8		N 80	20% cc
1114-16502	N43-07 W98-22	11/14	12/12			12/19		SW 25	30% cc

KANSAS: ERTS PHOTOS & TAPES 1972-73

I.D. Number	Coordinates	Scene Date	70mm Rec'd	95" Color Rec'd		CC Tapes Rec'd		Dist. Coverage %	Cloud Cover
				Bulk	Precision	Bulk	Precision		
1007-16563	N38-50 W101-12	7/30	8/31		4/13		10/20	N 20	10% cc
1023-16454		8/15	-			11/15		SE 10	
1025-16565	N37-60 W101-29	8/17	10/5	4/4			12/4	W 60	20% cc
1025-16571	N36-34 W101-57	8/17	10/5	4/4	4/27	11/15	2/28	SW 15	30% cc
1043-16570	N37-22 W101-45	9/4	10/6	4/6		11/15		W 60	50% cc
1060-16505	N38-48 W99-47	9/21	11/3	6/13		11/22	5/14	NE 15	10% cc
1060-16512	N37-23 W100-15	9/21	11/3			11/15	5/22	E 80	20% cc
1061-16564	N38-48 W101-12	9/22	10/25			2/21	6/19	N 20	0% cc
1061-16570	N37-23 W101-41	9/22	10/25			2/5		W 60	0% cc
1095-16460	N37-22 W98-52	10/26	11/22	2/20	5/17	11/22	1/10	E 20	30% cc

MISSOURI: ERTS PHOTOS & TAPES, 1972-73

I.D. Number	Coordinates	Scene Date	70mm Received	95" Color Rec'd		CC Tapes Received		Dist. Coverage	Cloud Cover
				Bulk	Precision	Bulk	Precision		
1034-16052	N37-22 W88-44	8/26	10/6	3/9	4/23	11/17	12/22	% NE 50	40%cc
1034-16055	N35-56 W89-12	8/26	10/6	2/26	4/23	11/22	12/19	S 40	30%cc
1035-16112	N36-34-W90-30	8/27	11/9				5/8	95	20%
1052-16052	N37-22 W88-41	9/13	12/26			11/15	5/10	NE 50	0%
1052-16055	N35-56 W89-09	9/13	12/26			11/18	5/10	S 40	10%
1070-16052	N37-31 W88-42	10/1	11/1	2/28	4/25	1/29	3/8	NE 50	0%
1070-16055	N36-06 W89-09	10/1	11/1	2/28		12/27	N/A	S 40	0%
1071-16111	N37-34 W90-05	10/2	11/1	2/28	6/1	12/27	2/28	NE 50	0%
1071-16113	N36-08 W90-33	10/2	11/1	2/28	4/25	12/22	2/28	S 40	0%
1089-16113	N37-25 W90-14	10/20	12/6			1/24	5/10	S 75	30%
1089-16120	N36-01 W90-42	10/20	12/6			1/24	N/A	SW 60	30%
1106-16060	N37-17 W88-54	11/6				12/8		NE 60	10%
1106-16063	N35-51 W89-22	11/6				12/8		S 50	20%

IDAHO: ERTS PHOTOS & TAPES, 1972-73

I.D. Number	Coordinates	Scene Date	70mm Rec'd	95" Color Rec'd		CC Tapes Rec'd		Dist. Coverage	Cloud Cover
				Bulk	Precision	Bulk	Precision		
1034-17470	N43-03 W112-41	8/26	10/20	4/6	5/23	11/22	1/24	% NE 25	10%
1034-17473	N41-38 W113-12	8/26	10/20	3/20	6/1	11/9	12/19	SE 40	20%
1035-17525	N43-02 W114-07	8/27	10/20	4/6	6/13	11/15	12/22	N 80	0%
1035-17531	N41-36 W114-39	8/27	10/20	3/12	4/25	11/8	12/19	SW 40	30%
1036-17583	N43-04 W115-33	8/28	10/6	3/20	4/25	11/8		NW 15	20%
1052-17470	N43-06 W112-38	9/13	11/3			11/8		NE 25	0%
1052-17472	N41-41 W113-10	9/13	11/3			11/8	6/19	SE 40	0%
1053-17524	N43-06 W114-03	9/14	10/19	2/26	4/25	11/8	12/19	N 80	0%
1053-17531	N41-42 W114-35	9/14	10/19	2/26	5/1	11/8	12/19	SW 40	0%
1054-17583	N43-04 W115-28	9/15	10/18	3/16	4/25	11/9	12/19	NW 15	0%
1054-17585	N41-38 W116-00	9/15	10/18	3/16	5/10	11/9	12/19	SW 5	0%
1071-17524	N43-16 W113-57	10/2	11/6			11/1		N 80	0%
1071-17531	N41-51 W114-29	10/2	11/6			12/8		SW 40	0%
1072-17583	N43-11 W115-24	10/3	11/9				6/19	NW 15	0%
1107-17532	N43-01 W114-09	11/7	12/6			12/8	5/15	N 80	10%

Ground Observations

The cost data for the ground observations have been summarized. There are two comprehensive parts to this summarization. The first part includes the enumerators time and mileage costs for the field enumeration of the survey data. Table 1 shows a summary of the ground survey expenses.

	Number of * segments per State	Average hours per segment	Average Miles Driven per segment
MISSOURI	449	5.42	82.22
IDAHO	54	5.75	82.80
SOUTH DAKOTA	350	6.27	95.91
KANSAS	456	6.44	107.10
Four State weighed Average	1309	6.02	94.57

* This study uses all segments in Missouri, South Dakota, and Kansas rather than just the segments in our study area.

The ground work update phase includes the costs for the field enumeration of the survey data. Update costs per segment are about 42% of the original costs. Update costs are reduced because the ERTS fieldwork only included locating the segments and recording crops present and their conditions.

Penn State Classifier

Summary

Version 1 of this classification system was received from Dr. Borden in December '72. A support subroutine REREAD would not function for us and had to be replaced, and all main programs were modified to accept the new program @INCORE. When the software was being brought up, some of the control card formats did not match the documentation sent us. All these discrepencies have been corrected.

The classification system assigned non-unique names to functionally similar, but operationally different subroutines. This slowed the implementation of 4 main programs PACLAS, PACLUS, PSTATS, and later PSUBSET. All known identically named, but operationally different decks have been isolated and renamed.

The cluster analysis routine, PACLUS, has been temporarily shelved. It has a few bugs in it that are to be corrected by the next version of the system, and we decided that this routine was not that essential to the immediate needs of the project.

Progress April 20 - June 19

Except for PACLUS the Penn State system is up and running. The system suffers from design problems and occasional bugs.

Some of the successful conclusion messages can be reached even though there are serious errors. Recently PSUBSET produced a file with a missing line, and PNMSP looped endlessly looking for that line. The endless loop has been removed from both mapping programs. The mapping programs sometimes indicate successful conclusion, when no maps are produced.

Some error messages give the user little clues as to which routine issued it, and therefore insufficient trace information. These are being corrected as they occur, or as a routine is updated for some other purpose.

The original system issued a variety of completion codes at the conclusion of the run. In some cases the successful completion codes were greater than the error codes and nearly all were greater than 8. This made it impossible for us to run programs in series within a job. Successful completion codes were reduced to 0 or 1, and error codes increased to 99 or larger. Subsequent steps within a job can be terminated as a result of failure of a prior step. This works well except that occasionally errors occur and a successful completion code is issued anyway.

We have learned that we cannot access any block of data specified in the table of contents record without the user processing all prior blocks. The Penn State Computer Center installed a system modification that allows the user to bypass errors associated with wrong length records. We are attempting to add special bypass coding into our version.

We are experimenting with using interactive terminals to run the Penn State System. Because of the above mentioned problem, we are having problems with the mapping programs. The PSTATS program I/O unit definitions are being changed to conform with the other system components. This is required because the TSO implementation uses a portioned data set, and IBM support software provides limited support with FORTRAN coded programs.

Statistical Analysis System

Summary

The Statistical Analysis System (SAS) has been successfully installed at the USDA Washington Computer Center. The SAS discriminant function was used successfully to classify some aircraft test data.

Segment Location Program

Our original approach to the problem of identifying and extracting data for individual segments and fields from the ERTS digital tapes was to develop a series of computer programs which would use geographic locations and other parameters to identify individual segments and fields with a minimum of human intervention. This system is not yet operational.

Reasons for the incompleteness of this system include:

1. Technical difficulties in obtaining nearly exact measurements of the other parameters required by the programs as indicated below.

2. Acquisition of the Penn State Classification System. (This provided an alternate approach and reduced the emphasis and completing the geographic location set of programs), and

3. Conflicting job priorities.

The following parameters must be known with some precision if the geographical coordinate system is to work efficiently for locating segments and to work at all in locating fields within segments.

1. The direction of the satellite ground path with respect to north

2. The angle of the scan line with respect to the ground path

3. The perpendicular distance between scan lines

4. The distance between data points (picture elements or pixels) on the scan line

5. The geographic coordinates of the ERTS "picture" format center and of the individual segments.

The direction of the ground path for a particular picture possibly could be obtained as a function either of the spacecraft heading or of the latitude of the picture center. The direction has been obtained empirically from printouts of picture 1061-16570. For this picture, it was 195°32', 5°32' more than the indicated spacecraft heading. Again for this name picture the direction of a perpendicualr to the scan line was 192°21', 2°21' more than the indicated spacecraft heading. Because the spacecraft heading is reported only to the nearest degree, these differences are not as precise as they would seem.

For the same tape, the perpendicular distance between scan lines was computed as 79.546 meters, as 79.0 meters from the Users Handbook. Also, the distance between picture elements was computed as 57.094 meters.

The geographic coordinates of individual segments from measurements on country road maps can be computed to within one data point. However, the reported format center of the digital data tapes is rounded to the nearest minute, i.e. to within 13 pixels (longitude) and 12 scan lines (latitude).

A great deal of additional work would be needed to make this system fully operational. Since the segment and field location procedure based on the Penn State system does work, only those parts of the geographic coordinate system (particularly the field location progress) which supplement the Penn State system are likely to be implemented in the near future.

Data Analysis

Segment and Field Location

We are now using the subsetting and mapping routines from the Penn State system in the segment location procedure. The flow of work in this procedure follows.

1. Identify, on 1/250,000 false color imagery, the approximate location of each segment.

2. Use the Penn State system to subset large areas of data from the ERTS system corrected digital data tapes and to print an 8 or 9 level grey-scale representation of band 5 for that area.

3. Locate and plot the approximate segment location on the area printouts.

4. Use the Penn State system to subset segment areas from the larger areas subsetted in Step 2 and to printout new grey-scale representations of each of these segments.

5. On these (Step 4) printouts, precisely define the segment and field boundaries.

6. Punch field identification cards as input to the analysis subroutines.

Particular disadvantages of this procedure are:

1. A large amount of manual work is involved, and

2. None of the output from this system is directly useable as input into other analytical programs.

At this time segment locations have been marked on the 1/250,000 false color imagery for at least one pass over each of the four test sites. We plan to locate segments and fields first in Idaho, then Kansas, South Dakota, and Missouri, in that order.

At present we have identified 37 of 44 June Enumerative Survey (JES) segments and 13 of 14 special training segments in Idaho. Individual fields have been delineated in 7 JES segments and in 2 training segments. Progress at this particular moment has been slowed both by reassignment of personnel to and from this group and be extremely slow turnaround time at the Department of Agriculture Washington Computer Center.

Target dates for completing the first round of segment and field locations (one set of pictures per state) are:

Idaho - July 6
Kansas - July 27
South Dakota - August 17
Missouri - September 7

Classification Programs

The Penn State University ACLASS program is now operating. It was written primarily for aircraft data where sun angles have a drastic effect on classification errors. The original data in n-space must be normalized before one classifies the data. Geometrically, the data points are "reduced" to the surface of a unit hypersphere. The object is to calibrate all points to the same average brightness so that the effect of sun angles will be minimized. This program is implemented in two steps; first, one takes data from known crops and calibrates or trains the discriminant procedure. This is done with the STATS program that estimates population mean vectors. Mean vectors are then located on the unit sphere and all unknown data that falls within a certain distance or within a specified angle of the normalized mean data will be classified into that population. We are able to run this classification procedure and preliminary results indicate that normalizing ERTS data is rather wasteful. However, some classification will be done to more thoroughly evaluate ACLASS.

We do have another classification procedure that is presented in a book by C. R. Rao, Linear Statistical Inference and Its Applications, page 488. Briefly this discriminant procedure assumes that the measurements of the different crops come from normal distributions with mean vectors μ_i and covariance matrics Σ_i. The normal distribution assumed here is not to be confused with the normalized data ACLASS uses.

An unknown sample point is then checked to see which multivariate normal distribution the point could have come from. It would be classified into the population that has the nearest mean vector. This procedure is programmed in SAS. We do not have software to get ERTS data into SAS efficiently, but we are preparing programs to alleviate this problem.

UNITED STATES DEPARTMENT OF AGRICULTURE
STATISTICAL REPORTING SERVICE
WASHINGTON, D.C. 20250

11

July 6, 1973

SUBJECT: ERTS-1 Progress Report

TO: ERTS Program Manager
NASA Headquarters
Code ER
Washington, D. C.

Enclosed is a copy of our type II progress report covering December 20, 1972 - June 19, 1973.

Harold F. Huddleston

HAROLD F. HUDDLESTON
Principle Research Statistician
Research & Development Branch
Research Division

Enclosure