THE AUXILIARY USE OF LANDSAT DATA IN ESTIMATING CROP ACREAGES: RESULTS OF THE 1975 ILLINOIS CROP-ACREAGE EXPERIMENT

Statistical Reporting Service
U.S. Department of Agriculture

SRS-21

The Auxiliary Use of LANDSAT Data in Estimating Crop Acreages: Results of the 1975 Illinois Crop-Acreage Experiment

By

Chapman Gleason, Project Coordinator

and

Robert R. Starbuck, Richard S. Sigman,

George A. Hanuschak, Michael E. Craig,

Paul W. Cook, and Richard D. Allen

Statistical Reporting Service

U.S. Department of Agriculture

Washington, D. C.

October 1977

TABLE OF CONTENTS

LIST OF TABLES

I. Introduction

This report summarizes the results of the Illinois Crop Acreage Experiment, a collaborative investigation by the staffs of the Center for Advanced Computation (CAC) at the University of Illinois and of the Statistical Reporting Service (SRS) of the U.S. Department of Agriculture. The purpose of the experiment was to investigate the usefulness of data collected by the orbiting LANDSAT satellites in improving the precision of crop acreage estimates at several levels--such as counties, groups of counties, such as, Crop Reporting Districts (CRD's), and entire states. The approach of SRS in using LANDSAT data to estimate crop acreages is to use it as an auxiliary variable with existing ground surveys.

This report describes the following phases of the project:

1) Ground data collection procedures.
2) Acquisition and management of project LANDSAT data.
3) Segment location and scene registration.
4) Processing systems developed to interface ground data and LANDSAT data for purposes of estimating crop acreages.
5) Pixel classification procedures and results for LANDSAT imagery collected over Illinois during the 1975 growing season.
6) Crop acreage estimates based on LANDSAT data for each Illinois county.
7) Comparison of crop acreage estimates based on LANDSAT data with June Enumerative Survey (JES) estimates for multi-county regions.

II. Data Sources

A. Ground Data*

1. Enumerator data

In support of this project, all crop and land-use information for the fields in the 300 SRS June Enumerative Survey segments in Illinois were keypunched by the Illinois State Statistical Office (SSO) to create a ground-truth data base. Every month throughout the growing season (July, August, and September) the crop maturity and land-use information for every field in each segment was updated. The Illinois SSO prepared computer programs to print field questionnaires which listed the field and crop acreage, cover type, and intended use from the previous visit as an aid to the field enumerators. The computer-printed questionnaire proved to be an excellent aid since enumerators did not have to copy information from the previous questionnaire to the current one. When the crop or land use changed between visits to the field, enumerators were instructed to accurately draw new color-coded field boundaries on ASCS (Agricultural Stabilization and Conservation Service) aerial photos (8" = 1 mile scale), indicate the date of change, and record the acreage and cover type of each new field on the questionnaire.

In order to interface satellite data processing with the JES and monthly update surveys, it was necessary to redefine fields if land use changed for any part of a JES field any time during the season. The largest part of the field retained the old field number and the next

---

*Questionnaires and methods used in SRS ground surveys are explained in more detail in Appendices A and B.

unused field number was given to the newly created field, (actually a subfield of the original field).

For example, if field 3 of Tract A had 30 acres of winter wheat harvested between the last visit in June and the current visit in July, and then 20 acres of soybeans were planted in one portion of the field with the remaining 10 acres left fallow, this information would be recorded as follows:

Survey	Tract	Field	Sub-Field	Cover Type	Acres	Maturity
June:	A	3		W. Wheat	30	Mature
July:	A	3	3	Soybeans	20	Planted
	A	3	16	Idle	10	Plowed

This, of course, assumes that field 16 was the next unused field number in Tract A.

The usual JES definition of a "field" is not as specific as might be desired for remote sensing purposes. For example, a JES field of 40 acres might contain 35 acres of corn and 5 acres of wasteland in one corner. Enumerators were asked to draw this as two fields, one of 35 acres and one of 5 acres. If fields of this type were not broken out in the JES, enumerators were asked to draw in the proper boundaries and list the correct acreages on a later visit. When this occurred, the ground-truth data for the affected fields were changed for the previous visits rather than defining new subfields.

In September, at the end of four visits to the JES segment, information had been collected on land use, cover types, and crop maturities for each JES field and follow-up survey field. Boundaries on

ASCS photographs were reviewed against the survey acreage data for all segments as a quality control procedure. The data for each field were carefully reviewed and edited. Most editing consisted of only filling in minor inconsistencies of data for non-crop fields. Appendix A contains the ground-data collection forms and the general data collection and editing procedures used by the Illinois SSO for this project.

A magnetic tape of the edited ground data was then delivered to CAC. CAC reformatted the tape and mailed it to Bolt, Beranek, and Newman (BBN) in Cambridge, Massachusetts, in a file format compatible with EDITOR ground-truth files. (EDITOR is an interactive image processing system developed by the Center for Advanced Computation, University of Illinois.)

2. Infrared aerial photography

Another source of ground truth was low altitude color infrared (IR) aerial photography at approximate scale 5" = 1 mile. This imagery was available for a subsample of 202 of the 300 segments. This photography was taken in late July and early August 1975. The tract and field boundaries for the 202 IR segments were transferred from the ASCS photos to the IR imagery. When the field boundaries drawn on the ASCS photos differed from the natural boundaries in the IR imagery, the boundaries in the IR imagery were used. The 202 color IR segment photos and the remaining 98 ASCS segment photos were then sent to CAC for segment digitization.

B. LANDSAT Data*

1. Acquisition

All LANDSAT imagery collected over Illinois during the summer of 1975 was acquired from NASA in the form of 70 mm film transparencies of bands 5 and 7. These were evaluated by both SRS and CAC with regard to project objectives. Ideally it takes only 11 LANDSAT scenes collected in three satellite passes over Illinois, each one day apart, to completely cover the state (see Figure 1). Because of clouds, however, portions of 13 separate LANDSAT scenes from a number of different dates were required for complete coverage of the state.

2. Pre-processing and reformatting

One of the project goals was to provide county, crop reporting district, and state-wide estimates for the entire state of Illinois. Since a county was the smallest geographic unit for which estimates were to be made, all LANDSAT imagery acquired from NASA was reformatted into a set of image-files such that each of the 102 Illinois counties was wholly contained within at least one such image file. To accomplish this, pseudo-frames of LANDSAT digital data were created when a county did not fall wholly within a LANDSAT frame. A pseudo-frame is created by linking data records from the bottom portion of one frame to the data records at the top of an adjoining frame having the same image date. Since different satellite passes have different image dates, pseudo-frames can only be constructed from frames within the same satellite pass. Six such

*Much of the information in this section is taken from the project description in Ray and Huddleston [1].

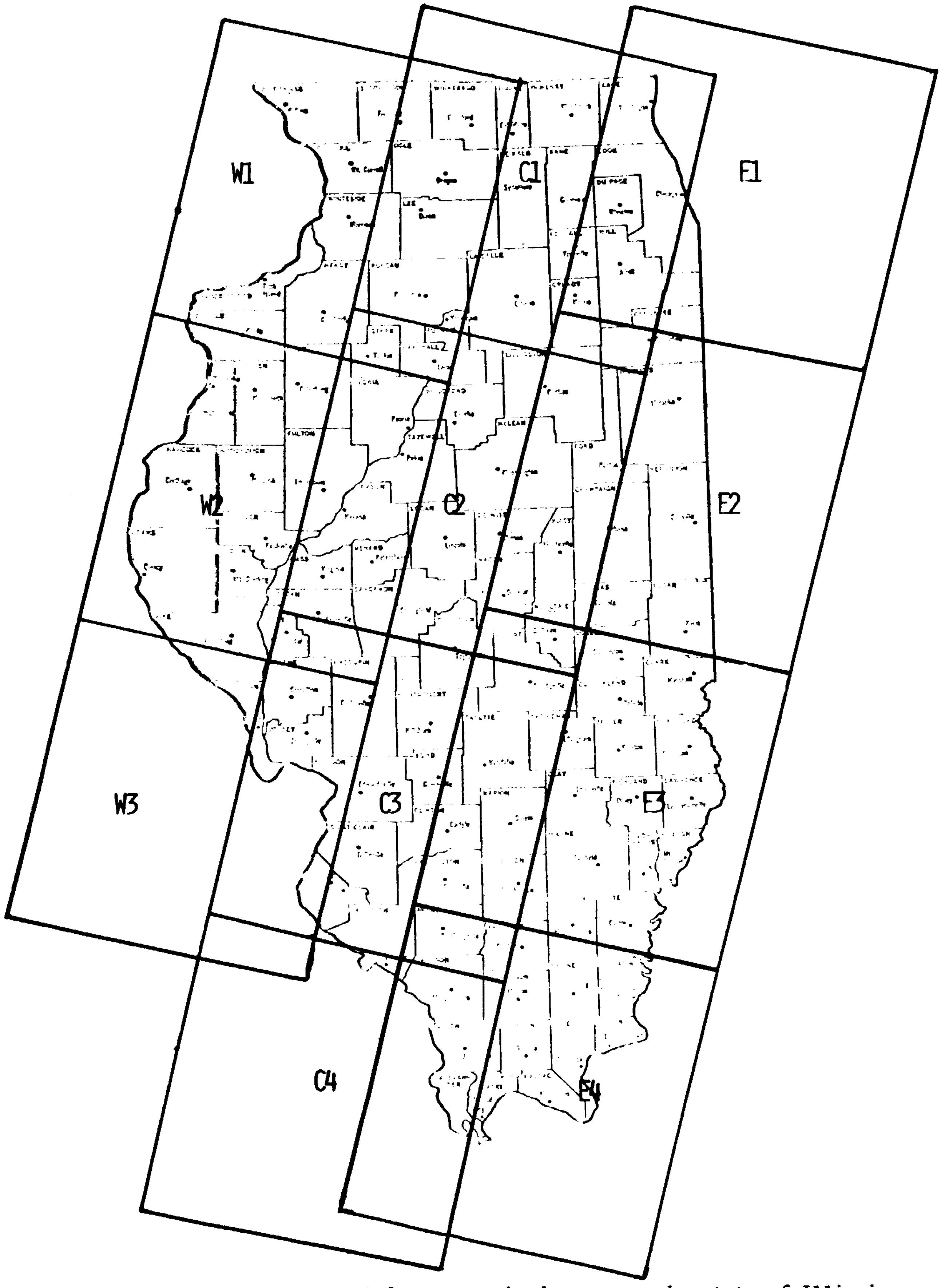


Figure 1. Portions of 11 LANDSAT frames required to cover the state of Illinois. The letter and numbers indicate the frame within a pass.

pseudo-frames were compiled for this project. Table 1 gives the LANDSAT frames and pseudo-frames needed to completely cover each county in Illinois on at least one image file (i.e., frame or pseudo-frame).

Table 1. LANDSAT frames and pseudo-frames for cloud free coverage of Illinois 1975.

Scene or Image File ID	Analysis Designation	Acquisition Date	Frame or Pseudo-frame
2194-16035	W1	August 4	Frame
2194-16042	W2	August 4	Frame
2194-16044*	W3	August 4	Frame
2194+16041	W1+	August 4	Splice of W1, W2
2194+16043	W2+	August 4	Splice of W2, W3
2193-15581	C1A	August 3	Frame
2211-15574	C1	August 21	Frame
2211-15580	C2	August 21	Frame
2211+15576	C1+	August 21	Splice of C1, C2
2175-15592	C3	July 16	Frame
2175-15595	C4	July 16	Frame
2175+15594	C3+	July 16	Splice of C3, C4
2228-15515	E1	September 7	Frame
2228-15522	E2	September 7	Frame
2228-15524	E3	September 7	Frame
2228-15531	E4	September 7	Frame
2228+15523	E2+	September 7	Splice of E2, E3
2228+15529	E3+	September 7	Splice of E3, E4

*This LANDSAT image was never annotated and cataloged by NASA. However, NASA made the first 500 scan lines available to us for this project.

The counties of Sangamon and Christian were not wholly contained in any one LANDSAT frame. Moreover, it was not possible to construct a pseudo-frame to contain these counties because in the selected LANDSAT imagery the candidate frames for building a pseudo-frame had different image dates. Consequently, no analyses of the LANDSAT data for those counties were performed. The geo-numeric numbering scheme used for the LANDSAT image files is shown in Figure 1.

3. Data Management

In addition to the partitioning of the LANDSAT data by image-files (frames and pseudo-frames), the complete set of 102 counties was subdivided into non-overlapping groups of contiguous counties with one county group per image-file. These county groups were called analysis districts and all data management and processing of the LANDSAT data was structured in terms of analysis districts. Fourteen such analysis districts were defined for this project (see Figure 2). These analysis districts became the focal point of a coordinated effort by CAC and SRS to process the data in the 13 LANDSAT image-files.

To process the LANDSAT data the following functions had to be performed:

1. Digitize and calibrate to a map base each of 300 SRS segments.
2. Register each LANDSAT image-file and locate the segments accurately.
3. Digitize the land-use strata maps for each of the 102 counties.
4. Train the classifier for each image-file and classify the entire image file.
5. Estimate the acreages for each image-file.

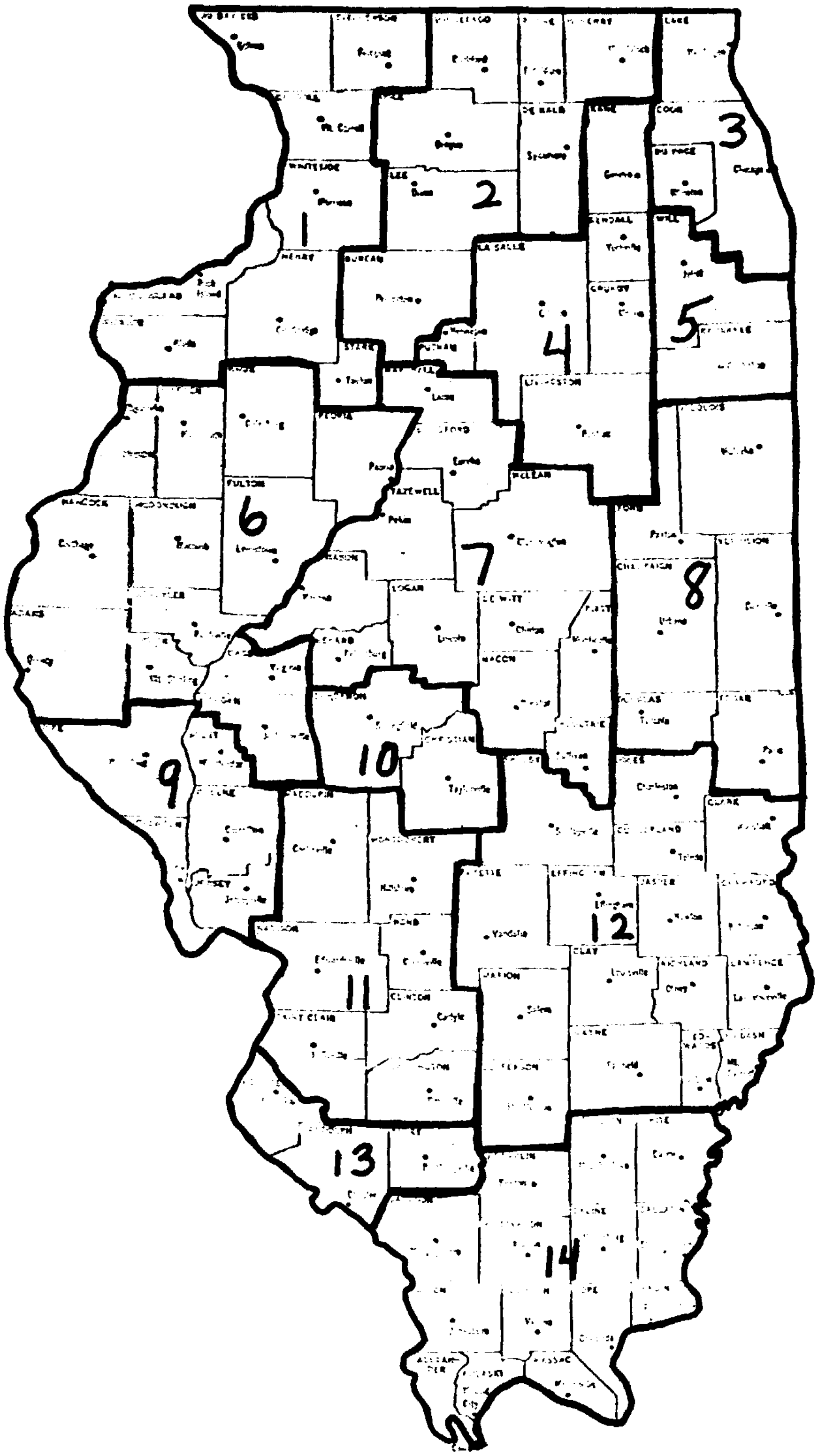


Figure 2. Analysis Districts for 1975 Illinois Acreage Estimation Project.

CAC managed and performed the following functions:

1. Digitization of the 300 SRS area segments.
2. Registration and segment location for W1, W2, W3, C1, C1A and E1.
3. Digitization of all county land-use strata maps for analysis districts 2, 3, 4, 5 and 7.
4. Development of software.

SRS managed and performed the following functions:

1. Ground data collection and editing for four visits to 300 SRS area segments.
2. Registration and segment location of C2, C3, C4, E2, E3 and E4.
3. Digitization of all county land-use strata maps for analysis districts 1, 6, 8, 9, 10, 11, 12, 13 and 14.
4. Software systems design for acreage estimation.
5. Analysis of all data sets.

III. Illinois Scene Registration and Segment Calibration

To utilize the LANDSAT data, the image files were registered to a map base, usually U.S. Geological Survey (USGS) maps. This process located segment and field data for classifier training and determined the location of land-use strata and county boundaries needed for county crop-acreage estimates [2]. For Illinois a method developed by CAC was used for scene registration [3]. CAC registered the scenes over W1, W2, W3, C1, and E1; whereas, SRS registered the scenes over C2, C3, C4, E2, E3, and E4.

A. Registration Procedure

1. First order registration

Scene registration consisted of two stages. The first stage, called first order registration, developed a linear regression between LANDSAT data (row, column) values and map (latitude, longitude) values. The regression data were the locations of physical features, called control points, which can be located in both the LANDSAT data and on a USGS topographic map; e.g., secondary road intersections, small lakes, groves of trees, clearings in woods, bends in rivers, river-road intersections, etc. The (row, column) values were determined by locating the features on the 1:500,000 scale LANDSAT photos for bands 5 and 7. The corresponding (latitude, longitude) values were determined from 7 1/2 or 15 minute quadrangle maps (i.e., of scales 1:24,000 or 1:62,500).

After selecting fifty such points well scattered throughout the scene, the map-to-LANDSAT linear regression was computed. Row and column residuals were calculated, and points with column residuals in excess of 10 pixels (15 pixels at the extreme edges of the scene) or row residuals in excess of 2 pixels were rejected as outliers. The linear regression was then recomputed from the non-rejected points. The resulting linear regression was then used to "deskew" the image into a more north-south orientation [4].

2. Precision registration

The second stage of registration, called precision registration, increased the degree of the polynomial transformation between LANDSAT data locations and the map coordinate system. To accomplish this the control points were located more accurately than in first order registration by using a light table to overlay 7 1/2 minute quadrangle maps with LANDSAT data greyscales of each control point. For 15 minute quadrangle maps, each greyscale was reduced to approximately 3/7ths of the original size to obtain a useable match of scales.

While the map and greyscale were overlayed, both were marked at the location of the control point. The marks were then digitized on a digitizing tablet to obtain location values needed for the regression calculations. Table 2 gives the precision registration results for quadratic fits in scenes registered by SRS. Comparable registration results were obtained by CAC for the scenes which they registered.

B. Segment Calibration

To determine labeled pixels for classifier training, each segment must be located with an accuracy of 1/2 pixel or better. This was accomplished by the following procedure:

1. At the scale of LANDSAT greyscales (approximately 1/24000), plots showing field boundaries were obtained for each segment.
2. The segment plots were then overlayed on the segment greyscales at the locations predicted by the precision registration polynomial.
3. By examining the greyscale's lightness and darkness patterns corresponding to segment fields, it was determined whether the segment was correctly located. If not, row and column shifts needed to move the segment to its correct location were determined and used as local corrections for locating segment pixels.

Table 2. Residual Mean Square Errors for scenes registered by USDA/SRS.

Scene ID	Root Mean Square Errors*		Maximum Residual*		No. of Points
	Line Error	Column Error	Line Error	Column Error	
2211-15580 (C-2)	0.4911	1.9783	1.3712	4.5184	61
2175-15592 (C-3)	0.6016	2.2984	1.6048	4.5075	61
2175-15595 (C-4)	0.5098	0.8768	1.0739	1.7182	34 (Partial Scene)
2228-15522 (E-2)	0.4019	2.4462	1.3274	4.0270	50
2228-15524 (E-3)	0.5652	2.0156	2.1089	4.5626	64
2228-15531 (E-4)	0.4509	2.2739	1.6470	5.7788	72

*measured in pixels

IV. Data Analysis

A. Processing Systems

To carry out the project objectives, existing in-house computer facilities (Washington Computer Center) could not be used to effectively manage and classify the large volume of data involved. Therefore, SRS contracted software development to the Center for Advanced Computation. CAC working with SRS staff implemented the following EDITOR procedures for this project:

- Registration and digitization systems,
- Segment location and masking systems,
- Data analysis systems, and
- Acreage estimations systems.

These systems are described in detail in [5] and [6].

In the data analysis process, a large number of computer files were created. The development of a self-documenting file-naming convention [6, Appendix] greatly simplified data management.

B. Analysis Results

The statistical methods used in this project have been described in previous reports. The paper by Sigman, Gleason, Hanuschak, and Starbuck [7] (excerpt in Appendix B) gives details on classifier design and acreage estimation with stratified sampling. Two companion papers by Ray and Huddleston [1] and Huddleston and Ray [8] give methodological details of the project for simple random sampling. As explained in the papers by Wigton [9] and Von Steen and Wigton [13], crop acreages were estimated by a regression estimator with enumerator data from the JES as the primary, survey variable and LANDSAT data as the auxiliary variable.

The effectiveness of LANDSAT data to serve as such an auxiliary variable was measured by the relative statistical efficiency of the regression estimator versus the direct expansion estimator based only on enumerator data. In the analysis of the 1975 Illinois LANDSAT data, three major objectives were pursued. These were:

• To investigate the influence of various factors, both methodological and geographical, on classifier performance,

• To compute crop-acreage regression estimates plus the relative sampling errors of these estimates for individual Illinois counties, and

• To compute crop-acreage estimates for various multi-county areas and then compare the precisions of these estimates to the JES direct expansion estimates for these areas.

1. Classifier Performance Study

The classifier performance study was a set of classification trials performed in domains W1, W2, and W123 which investigated the influence of various factors on classifier performance. Traditionally, the performance of a classifier has been measured in terms of a confusion matrix of percents correct and commission error rates. However, if a classifier is being used to estimate crop acreages, then it should be evaluated in terms of how well it does exactly that. Thus, the classifier performance measure used was the variance of resulting regression estimates.

a. Study Variables

The following factors were investigated for their influence on classifier performance:

i. <u>Classifier Domain</u>. This factor investigated the influence of geography, date of imagery, and size of classifier domain on classifier performance. In the August 4 western satellite pass, single-scene classification and multi-scene classification were compared. This was done by analyzing image files W1 and W2 individually and then jointly with W3 as a joined-scene called W123. In the central pass the classifier domains were for three different dates: domain C1A (= image file C1A) on August 3; domain C12 (concatenation of image files C1+ and C2) on August 21; and C33+ (concatenation of image files C3 and C3+) on July 16. In the September 7 eastern pass, the classifier domains were domain E12 (image files E1 and E2) and domain E23+ (image files E2+ and E3+).

Figure 3 is a map of the eight classifier domains. Because the LANDSAT scenes overlap, 16 counties were contained in more than one classifier domain. These counties, called overlap counties, were used to measure the repeatability of the regression estimates. Table 3 shows the distribution of land area by land-use stratum for the eight classifier domains. Items of note in this table are the following:

• In each of the satellite passes there is a north-south gradient in land use. From north to south the proportion of land in stratum 20 increases whereas the proportion in stratum 11 decreases.

• Domain E12, which contains Chicago, is the most heterogeneous of the eight domains.

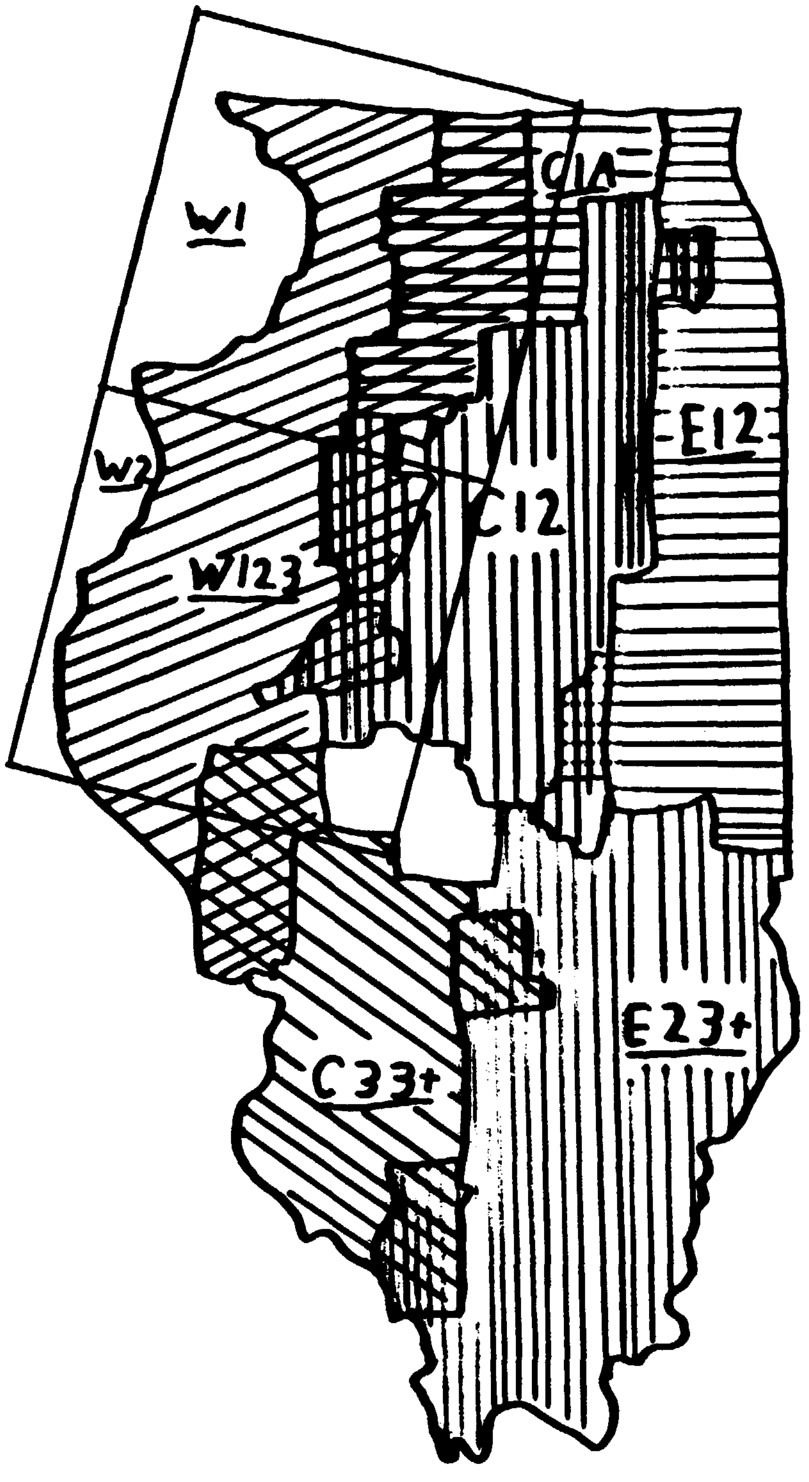


Figure 3. Classifier Domains for 1975 Illinois Acreage Estimation Project.

Table 3. Distribution of land area by land-use stratum within classifier domains

satellite pass	domain	% of domain land area contained in stratum: 11[1]	12[2]	20[3]	30[4]
western	W1	65	16	13	6
	W2	36	19	34	11
	W123	46	18	27	9
central	C1A	73	17	2	8
	C12	75	9	7	9
	C33+	38	24	29	9
eastern	E12	67	7	3	23
	E23+	34	29	24	13

[1] 75% + cultivated
[2] 50% - 75% cultivated
[3] 15% - 49% cultivated
[4] non-cultivated

ii. Number of Classification Categories. This factor investigated various strategies for developing classification categories. The strategies studied were intra-crop clustering to create multiple categories per crop (MCPC), straight supervised training with a single category per crop (SCPC), and pooled crop (PC) categories.

iii. Prior Probabilities. This factor investigated the effect on classifier performance of using "different prior probabilities" for the classification categories. Strictly speaking, there is only one correct set of prior probabilities for a given geographical region. Using "different prior probabilities" actually means using different weighting factors for the likelihood functions in the class discriminant functions.

The two types of prior probabilities studied were unequal priors proportional to expanded reported acres, denoted PER, and equal priors, denoted EP. In a given region the PER prior probability for a particular cover was defined as the ratio of the current year direct expansion estimate to the total land area in the region. Note that the unequal priors are not based on historic crop-acreage estimates.

iv. Training/Test Data Sets. This factor investigated the data sets on which the classifier was trained and tested. The following methods were employed to allocate the LANDSAT data associated with JES segments between the training and test data sets:

• Resubstitution, in which all the segment data, denoted NB for "not background", were used to both train and test the classifier,

• Sample partition, in which the classifier was trained on a 50% sample of segment fields, denoted FLDS, and then tested on all of the segment data, and

• Jackknifing, denoted JK, in which the training set was 3/4 of the data and the test set was the remaining 1/4. This allocation was repeated four times so that the union of the four test sets was the entire collection of segment data.

The jackknifing technique used was that referred to by Toussaint as the Pi-method [10]. Thus, four separate estimates of classifier performance were obtained and averaged to yield the jackknife estimate.

There are two reasons why the training/test factor was of interest. The first reason was the desire to minimize the work involved with evaluating a classifier. The resubstitution and sample partition methods are easy to perform but are known to produce biased evaluations of the classifier in small samples. On the other hand, the jackknife is known to give a less biased evaluation but also involves substantially more effort. Consequently, if the three training/test methods give similar results in the classifier performance study involving domains W1, W2, and W123, this would indicate that resubstitution or sample partition would be sufficient for classifier training and testing in the other Illinois domains.

The second reason for investigating this factor was to study the sensitivity of the classifier to the selection of the training data. This was the purpose of performing sample partition and then comparing the results with those from the other two methods of classifier evaluation.

v. Strata Pooling and/or Deletion. Table 4 shows the distribution of JES segments by stratum for each classifier domain. As can be seen, a number of strata have zero or very few segments in them. Thus, it was necessary to pool and/or delete strata and then compute stratum regression estimates on the pooled, undeleted strata. Some of the strata poolings which were tried are the following:

Pooled Strata #	Original Strata Pooled Together
0	11,12,20,31,32,33,40,61
10	11,12
30	31,32,33,40,61
50	20,31,32,33,40,61

Table 4. Sample sizes by strata for all data sets.

Domain	Total	Number of segments in strata.*							
		11	12	20	31	32	33	40	61
W1	44	30	6	5	2	1	0	0	0
W2	40	16	10	11	1	0	0	1	1
W123	83	44	16	17	3	1	0	1	1
C1A	30	21	4	0	4	1	0	0	0
C12	52	40	2	5	3	1	0	0	1
C33+	43	18	9	9	4	2	0	0	1
E12	56	35	5	1	7	6	2	0	0
E23+	66	26	21	11	2	0	0	5	1

*W1 and W2 entries are on an entire scene basis. All others are for the counties wholly contained in the respective scene.

The strata used in a particular classification trial are identified with a strata-description notation. A "-" is used to separate distinct strata, and parentheses are used to surround pooled strata. For example, 11-(12,20)-30 indicates that stratum 11 is a distinct stratum, strata 12 and 20 are pooled together, and strata 31, 32, 33, 40, and 61 are also pooled together and called 30. Leaving a stratum out of a strata description indicates that the particular stratum was deleted from the classification analysis trial. For example, 11-12-30 indicates the deletion of stratum 20.

Another reason for deleting a particular stratum from the classification analysis was very poor classifier performance in the stratum; i.e., a stratum r-square (see Appendix B) of less than 0.10. When strata were deleted from the classification analysis, "swiss cheese" estimates were computed to estimate crop-acreages. A swiss cheese estimate consists of stratum regression estimates on the strata included in the classification analysis and direct expansion estimates on the strata excluded from the classification analysis.

b. Comparison Measures

In the classification trials the classification objective was to minimize the variance of the resulting regression estimates. As shown in equation (2) of Appendix B, this is accomplished by maximizing the stratum r-squares. Hence, to compare classifier performance on the same stratum, the respective r-squares were compared. For multi-strata regions, classifier performances were compared in terms of the relative efficiencies of the resulting estimates. Two types of relative efficiency were calculated. The first type, denoted RE1, was calculated with respect to the direct expansion estimator which uses the same poolings as the regression estimator. RE1 measures the gain, in terms of lower variance, of the regression estimate over the pooled JES direct expansion estimate. Of course, this doesn't take into account the strata in the direct expansion estimate. However, a second type of relative efficiency, denoted RE2, was calculated with respect to direct expansion over the 11-12-20-30 pooling, or over the best direct-expansion pooling for the region. Thus, RE2 measures the gain, in terms of increased precision, of the regression estimate over the unpooled JES direct expansion estimate.

c. Findings

The classification trials performed in the eight classifier domains are described in Table 5. In Appendix C the corn and soybeans results for the NB and FLDS classification trials are tabulated. In these results the following classification phenomena were common to all eight of the classifier domains:

- PER priors produced higher percents correct* compared to equal priors for both corn and soybeans. However, equal priors yielded higher r-square values compared to PER priors in almost all cases for corn and in several cases for soybeans.

- In the test-data sets (all segment interior pixels) the number of pixels classified as corn or soybeans exceeded the respective number of corn and soybean pixels actually present. For all other covers the opposite was true. The use of equal priors, however, tended to lessen these effects; i.e., there were less commission errors into the major crop categories when equal priors were used.

- Training the classifier on a 50% sample of fields for each cover yielded r-squares very close to those for training on NB (all JES data).

- R-squares in stratum 20 were low for corn, but somewhat better for soybeans.

---

*Percent correct is the percentage of test pixels (all segment-interior pixels, including field boundaries) correctly classified.

Table 5. Summary of Classifier Performance Study.

Trial	Analysis Data Set	No. of Categories/ Type of Pooling Strategy	Priors	Training/ Test	Strata Poolings Tried
W1.1	W1	10-SCPC	PER	NB	0; 10-50; 11-12-20-30
W1.2	W1	10-SCPC	PER	FLDS	0; 10-50; 11-12-20-30
W1.3	W1	10-SCPC	EP	NB	0; 10-50; 11-12-20-30
W1.4	W1	10-SCPC	EP	FLDS	0; 10-50; 11-12-20-30
W2.1	W2	7-SCPC	EP	FLDS	0; 10-50; 11-12-20-30
W2.2	W2	7-SCPC	PER	NB	0; 10-50; 11-12-20-30
W2.3	W2	7-SCPC	EP	NB	0; 10-50; 11-12-20-30
W123.1	W123	10-SCPC	PER	NB	0; 10-50; 11-12-20-30
W123.2	W123	10-SCPC	EP	NB	0; 10-50; 11-12-20-30
W123.3	W123	15-MCPC	EP	FLDS	0; 10-50; 11-12-20-30
W123.4	W123	15-MCPC	EP	NB	0; 10-50; 11-12-20-30
W123.5	W123	10-SCPC	EP	JK	0
C1A.1	C1A	14-MCPC	EP	FLDS	0; 10-50; 11-12-20-30
C1A.2	C1A	14-MCPC	PER	FLDS	0; 10-50; 11-12-20-30
C1A.3	C1A	14-MCPC	EP	NB	0; 10-50; 11-12-20-30
C1A.4	C1A	14-MCPC	PER	NB	0; 10-50; 11-12-20-30
C12.1	C12	26-MCPC & PC	EP	NB	11-12-20-30
C12.2	C12	10-MCPC & PC	EP	FLDS	11-12; 20-Other
C12.3	C12	6-SCPC & PC	EP	NB	11-Other; 11-12; 20-Other
C12.4	C12	5-SCPC & PC	PER	NB	11-12-20
C12.5	C12	4-SCPC & PC	EP	NB	11-12-20; 11-12, 20; 11, 12, 20
C33+.1	C33+	10-SCPC & PC	EP	NB	0; 10-50; 11-12-20-30
C33+.2	C33+	10-SCPC & PC	PER	NB	0; 10-50; 11-12-20-30
C33+.3	C33+	14-SCPC & PC	PER	NB	0; 10-50; 11-12-20-30
C33+.4	C33+	16-MCPC	PER	NB	0; 10-50; 11-12-20-30
C33+.5	C33+	12-MCPC & PC	EP	NB	0; 10-50; 11-12-20-30
C33+.6	C33+	9-SCPC & PC	EP	NB	0; 10-50; 11-12-20-30
C33+.7	C33+	19-MCPC	EP	NB	11-12-20-30
C33+.8	C33+	17-SCPC	EP		11-12-20-30
E12.1	E12	24-MCPC	EP	NB	0; 10-50
E12.2	E12	24-MCPC	PER	NB	0; 10-50
E23+.1	E23+	28-MCPC	EP	NB	0; 10-50; 11-12-20-30; 11-12-50
E23+.2	E23+	28-MCPC	PER	NB	0; 10-50; 11-12-20-30; 11-12-50
E23+.3	E23+	18-MCPC	EP	NB	0; 10-50; 11-12-20-30; 11-12-50
E23+.4	E23+	18-MCPC	PER	NB	0; 10-50; 11-12-20-30; 11-12-50

The optimum strata pooling varied between covers and classifier domains. Within a specific classifier domain, however, the same strata pooling was generally optimum for all classfiers of a given cover.

The low r-squares for corn in stratum 20 are explainable by the very nature of this stratum. Stratum 20 contains 10-49% cropland intermixed with mostly woods and permanent pasture. Thus, because there was considerable overlap in the spectral distribution of woods, permanent pasture, and corn, a large number of woods and permanent pasture pixels were erroneously classified as corn. This caused a very low corn r-square for this stratum.

Figures 4 and 5 plot corn and soybean stratum r-squares against imagery date for the classifier having highest RE2 in each domain for a number of different stratum poolings. The crop development stage and "best" RE2--that is, maximum RE2 over all attempted classifiers and stratum poolings--are also plotted. Table 6 more fully describes the classifiers and stratum poolings having best corn and soybean RE2's in each of the eight domains.

Figure 4 shows that for corn the <u>stratum 11</u> r-squares were largest on August 3 and 4. In <u>stratum 20</u>, however, August 3 and 4 along with August 21 had the smallest corn r-squares.

The high corn r-squares in stratum 11 on August 3 and 4 are possibly explained by the crop condition on these dates. In 1975, corn was nearly 100% silked by the first week in August [11]. The accompanying tassels, which are yellow, possibly distinguished corn from other green crops in

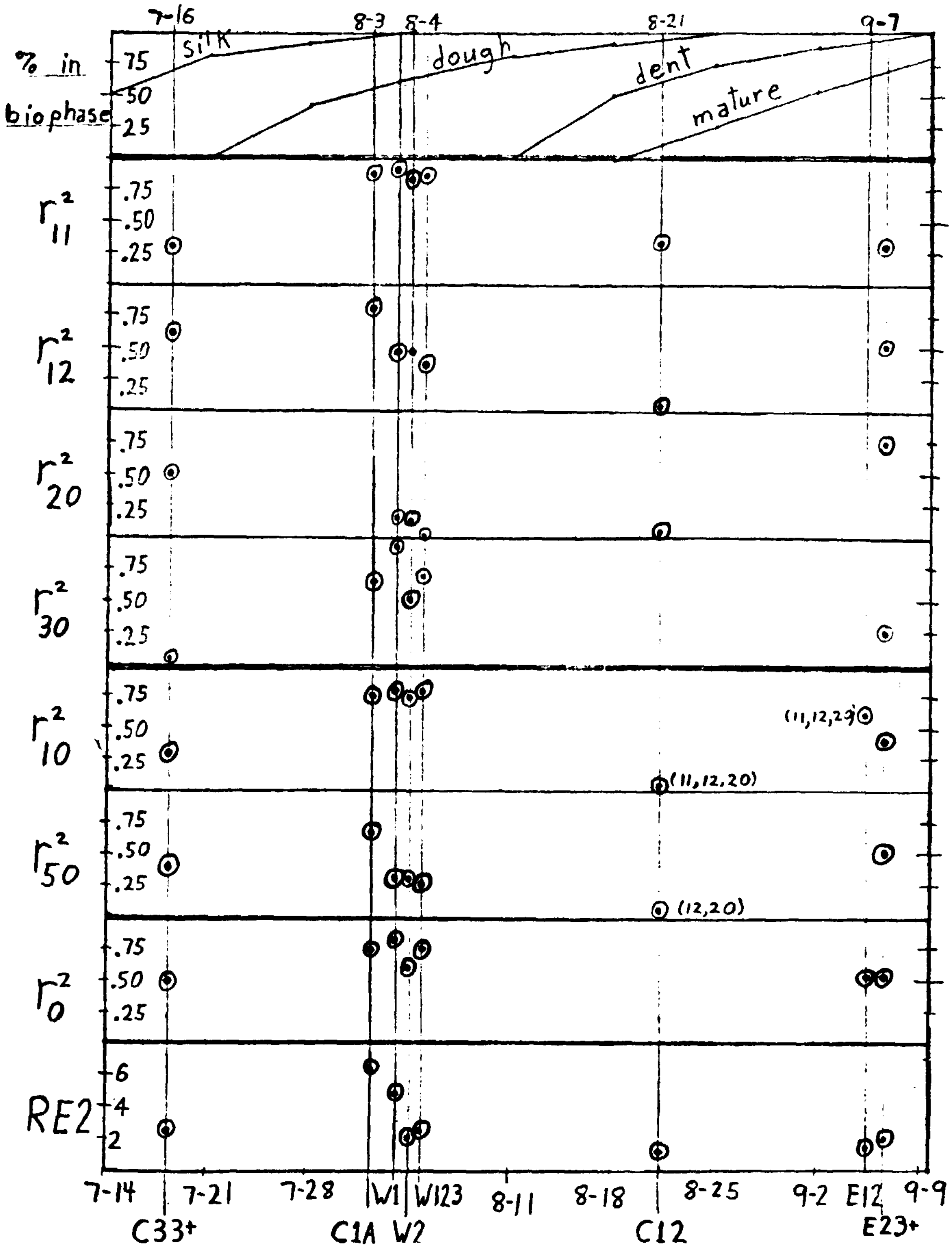


Figure 4. Corn Growth Stage, plus Stratum Coefficients of Determination (r-squares) and Relative Efficiency (RE2) of Best Corn Classifier, as a Function of LANDSAT Imagery Date.

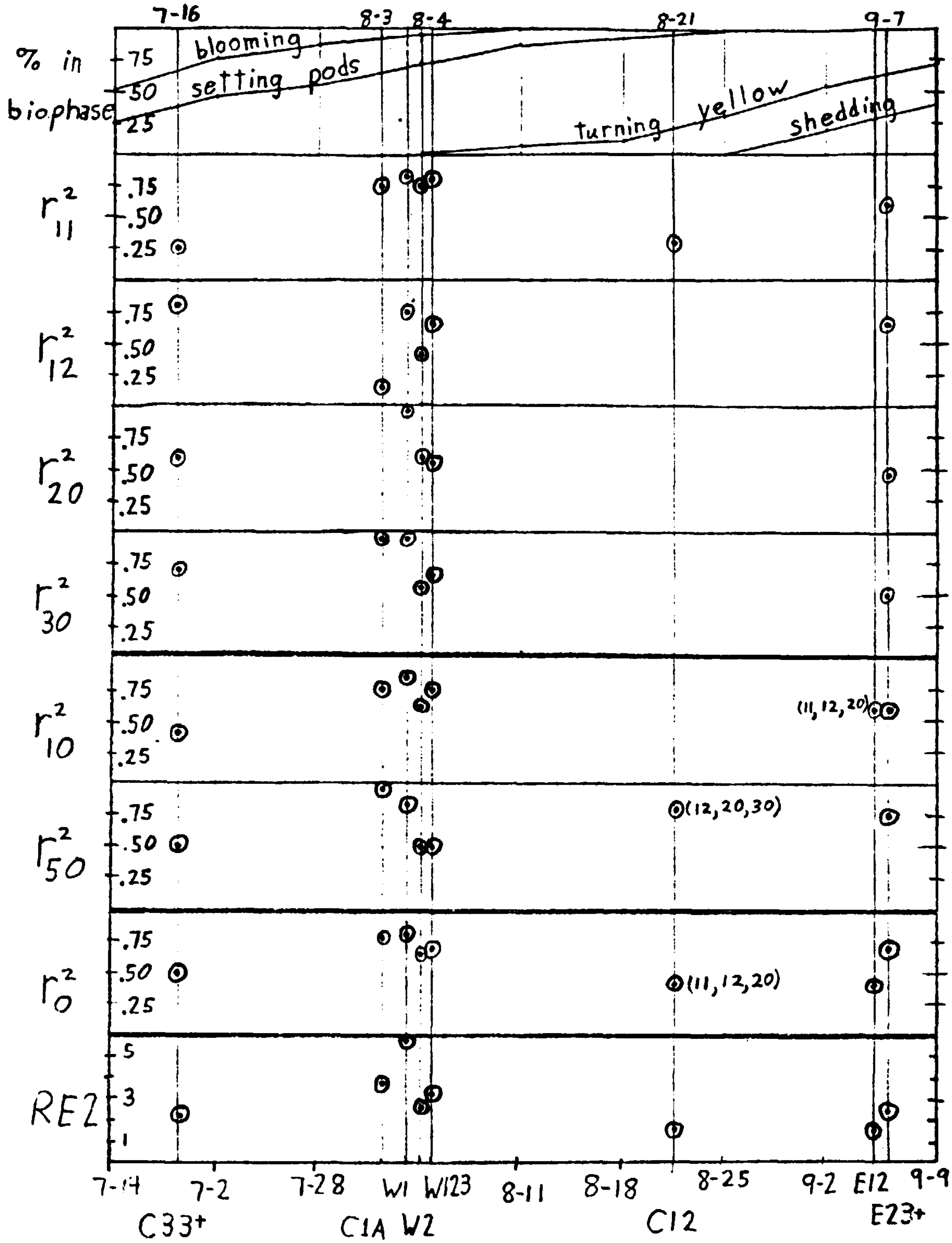


Figure 5. Soybean Growth Stage, plus Stratum Coefficients of Determination (r-squares) and Relative Efficiency (RE2) of Best Soybean Classifier, as a Function of LANDSAT Imagery Date.

Table 6. Corn and Soybean Classifier having maximum RE2 for each data set.

Crop	Data Set	Date	RE2	Categories	Priors	Train/test	Strata Pooling
Corn	W1	Aug. 4	4.58	10/SCPC	EP	FLDS	11-12-20-30
	W2	Aug. 4	2.13	7/SCPC	EP	NB	10-50
	W123	Aug. 4	2.48	15/MCPC	EP	FLDS	11-12-20-30
	C1A	Aug. 3	6.30	14/MCPC	EP	FLDS	11-12-30
	C12*	Aug.21	1.27	4/SCPC&PC	EP	NB	11-(12.20)
	C33+	July16	1.74	10/SCPC&PC	EP	NB	10-50
	E12	Sept.7	1.86	24/MCPC	PER	NB	(11,12,20)
	E23+	Sept.7	1.92	28/MCPC	EP	NB	11-12-20-30
Soybeans	W1	Aug. 4	5.76	10/SCPC	EP	FLDS	11-12-20-30
	W2	Aug. 4	2.34	7/SCPC	PER	NB	0
	W123	Aug. 4	3.22	15/MCPC	PER	FLDS	0
	C1A	Aug. 3	3.83	14/MCPC	PER	FLDS	0
	C12*	Aug.21	1.83	6/SCPC	EP	NB	11-(12,20,30)
	C33+	July16	2.23	10/SCPC&PC	EP	NB	11-12-20-30
	E12	Sept.7	1.06	24/MCPC	PER	NB	0
	E23+	Sept.7	2.38	18/MCPC	EP	NB	11-12-20-30

*Entries are RE1's for this data set.

stratum 11 such as alfalfa and soybeans. In stratum 20, however, the August 3 and 4 crop condition for corn was apparently not a distinguishing feature since very low corn r-squares were obtained in this stratum. In fact, the highest corn r-square in stratum 20 was obtained on September 7, when the majority of corn was in the mature stage.

In the four domains having August 3 or 4 imagery--that is, W1, W2, W123, and C1A--the stratum r-squares for corn were very similar. The best RE2's for these domains were, however, very different. This phenomenon is, in fact, explained by the poor classification results for corn in stratum 20 on August 3 and 4 and by the fact that the four domains have different amounts of land in stratum 20 (see Table 3). Domain C1A had the least amount of stratum 20 land and was thus least affected by poor classifier performance in stratum 20. Consequently, domain C1A had the highest corn RE2. On the other hand, domain W2 had the most stratum 20 land of the four domains and consequently had the lowest corn RE2 of the August 3 and 4 domains.

Figure 5 shows that for soybeans the stratum 11 r-squares were, as for corn, largest on August 3 and 4. Unlike corn, however, poor classification results in stratum 20 were not encountered for soybeans. Also, unlike corn, the superior stratum 11 r-squares on August 3 and 4 were probably not due to soybean growth stage. The reason for this is that the remote sensing appearance of soybeans did not change a great deal over the image dates analyzed. Apparently what happened was that August 3 and 4 produced higher soybean r-squares because it produced higher corn r-squares; i.e., on August 3 and 4 the improved separability for corn decreased the confusion between corn and soybeans and thus the r-squares for both crops increased.

The optimality of August 3 and 4, 1975, for corn and soybean classification confirms 1974 CITARS findings in Illinois [12]. In 1975, crops were approximately 2-3 weeks ahead of the average development stage of the previous three years. Thus, early August 1975 corresponds roughly to late August 1974, which CITARS found to be the optimal 1974 date for corn and soybean discrimination.

Table 6 allows the comparison of best RE2's across the eight classifier domains. Best corn RE2'S ranged from a high of 6.3 in domain C1A (August 3) to a low of 1.3 in domain C12 (August 21). An examination of the C12 imagery, however, revealed the presence of light haze over the entire pseudo-frame, which explains the poor C12 results. Best soybean RE2's ranged from a high of 5.76 in domain W1 (August 4) to a low of 1.06 in domain E12 (September 7).

Table 7 presents the results of trial JK in which jackknife training and testing was used. Table 8 compares the results of this trial to the

Table 7. r-squares for jackknifed classification (W123, SCPC, EP, pooling 0)

cover	pooled-stratum-0 r-square						
	jackknife group						C.V.
	1	2	3	4	Ave	S.E.	(%)
Alfalfa	.002	.001	.195	.078	.069	.09	132.7
Corn	.734	.814	.639	.680	.717	.07	10.5
Dense Woods	.097	.003	.030	.213	.086	.09	109.2
Hay	.017	.245	.042	.271	.144	.13	92.2
Oat Stubble	.000	.016	.119	.004	.035	.06	163.9
Oats	.119	.001	.069	.109	.094	.08	87.8
Permanent Pasture	.339	.304	.552	.269	.366	.13	34.8
Soybeans	.578	.745	.843	.520	.671	.15	22.2
Wasteland	.847	.732	.062	.248	.472	.38	79.9

corresponding resubstitution trial (Trial W123.2). The jackknife and resubstitution r-square values are quite similar, the major dissimilarities being for those cover types which have large coefficients of variation and small r-squares in Table 7. This suggests that for

Table 8. Comparison of jackknifed and resubstitution r-squares (W123, SCPC, EP, Pooling 0)

cover	train/test JK	NB
Alfalfa	.069	.09
Corn	.717	.70
Dense Woods	.086	.01
Hay	.144	.25
Oat Stubble	.035	.06
Oats	.094	.15
Permanent Pasture	.366	.36
Soybeans	.671	.67
Wasteland	.472	.81

sufficiently large sample sizes, the resubstitution method will yield r-square values for major crops whose biases are acceptably small.

Finally, Table 9 compares classifier performance in domain W123 over all covers and for two different types of prior probabilities. Items to note are the low r-squares and RE1 values for minor crops and the fact that neither type of prior probability, neither EP nor PER, was optimal for every cover. The trends in Table 9 were also demonstrated in the other classifier domains. These results imply that for minor crops, regression acreage estimates are fruitless for the data sets analyzed and for major crops a different classifier should be designed for each major crop type in order to maximize the efficiencies of regression estimates.

Table 9. r-squares and relative efficiencies for all covers (W123, MCPC, FLDS, Pooling 0)

Cover	r-square		RE1	
	EP	PER	EP	PER
Water	.89	.84	8.70	6.23
Waste	.78	.82	4.47	5.45
Soybeans	.62	.71	2.61	3.39
Corn	.75	.57	3.90	2.32
Permanent Pasture	.32	.35	1.44	1.51
Woods	.02	.24	1.01	1.31
Alfalfa	.05	.13	1.04	1.13
Hay	.20	.10	1.24	1.10
Oats	.14	.05	1.15	1.04
Oat Stubble	.01	.03	1.00	1.02

2. Multi-County Crop Acreage Estimates

The relative efficiencies obtained in the majority of classification trials indicated that the auxiliary use of LANDSAT data can reduce the variance of crop acreage estimates for corn and soybeans. Consequently, multi-county regression estimates for corn and soybeans were calculated for the ten-county Western Crop Reporting District (CRD) and for all the classifier domains except domains W1 and W2 since they were subsets of domain W123. The multi-county regression estimates were compared to estimates calculated by direct expansion of enumerator data and to estimates obtained from the summation of final 1975 county estimates published by the Illinois SSO. The final SSO estimates are predominantly based on the Illinois State Farm Census.

In Appendix C the classifiers used for acreage estimation are indicated. Table 10 lists the various multi-county crop acreage

Table 10. Estimated Acres of Corn and Soybeans for wholly contained counties in each analysis area.

Analysis Area *	No. of Counties Wholly Contained On Data Set	Estimator	Corn		Soybeans	
			Acres	C.V.	Acres	C.V.
W123	29	Direct Expansion[1/]	4,110,150	3.6%	1,539,200	7.7%
		Regression[2/]	4,125,400	2.5%	1,681,800	5.2%
		SSO[3/]	3,682,300		1,657,800	
C1A	7	Direct Expansion	1,191,400	7.1%	532,700	13.9%
		Regression	1,180,500	2.9%	523,200	8.2%
		SSO	1,196,900		502,900	
C12	20	Direct Expansion	2,907,700	4.5%	2,217,200	5.5%
		Regression	2,945,100	4.3%	2,127,200	5.1%
		SSO	2,939,700		1,990,400	
C33+	16	Direct Expansion	1,158,000	9.5%	1,675,100	8.6%
		Regression	1,077,000	8.6%	1,540,000	6.8%
		SSO	1,233,000		1,246,000	
E12	12	Direct Expansion	1,781,300	5.6%	1,439,500	6.3%
		Regression	1,577,300	4.1%	1,290,700	6.5%
		SSO	1,792,000		1,383,000	
E23+	32	Direct Expansion	1,669,500	7.5%	2,431,950	5.2%
		Regression	1,615,000	6.9%	2,357,850	3.8%
		SSO	1,767,000		2,045,000	
West CRD	9	Direct Expansion	1,316,000	8.5%	562,000	13.1%
		Regression	1,269,000	4.6%	574,100	10.6%
		SSO	1,125,000		680,000	

* Analysis area = domain (e.g. W123, C12, etc.) or sub-domain (e.g. West CRD).

1/ Planted acres. 2/ Standing acres (at image date). 3/ Harvested acres.

estimates and their coefficients of variation (CV's). For the Western CRD and for domain C1A, substantial decreases in sampling variance were achieved by the regression estimator for both corn and soybeans. Western CRD corn CV's were 8.5% for direct expansion, decreasing to 4.6% for regression; soybean CV's were 13.1% for direct expansion and 10.6% with the regression estimator. Domain C1A corn CV's were 7.1% and 2.9% for direct expansion and regression, respectively; whereas, soybean CV's were 13.9% with direct expansion decreasing to 8.2% with regression. In domain W123 only modest gains in precision were achieved by the regression estimator; while in the other four domains, gains in precision by the regression estimator were marginal. In fact, for soybeans in domain E12 the regression CV was larger than the direct expansion CV; i.e., the regression estimator using both LANDSAT data and enumerator data had a larger variance than the direct expansion estimate using only enumerator data. The reason for this was that because of small sample sizes in a number of E12 strata, it was necessary to pool strata in order to compute a regression estimate. Unfortunately, the loss in estimator precision due to collapsing strata exceeded the gains in precision due to regression.

The gains in precision by the regression estimates for soybean acreages were generally less than the gains for corn. This occurred because in a given domain the same classifier was used for both corn and soybeans. Since the classifier chosen was usually the optimal corn classifier, it was in many cases sub-optimal for soybeans. If optimal soybean classifiers had been used, then the gains in precision by the regression estimator would have been slightly higher for soybeans.

Additional items of note in Table 10 are the following:

• For corn the direct expansion estimate was with two exceptions always between the regression estimate and the SSO estimate. Thus, regression in these cases pulled the direct expansion corn estimates away from the SSO values.

• On the other hand, for soybeans the regression and SSO estimates were in six out of seven cases in the same direction away from the direct expansion value. Thus, for soybeans regression in most cases pulled the direct expansion value toward the SSO estimate.

• For both corn and soybeans, the regression estimate was larger than the direct expansion estimate in five out of seven cases. However, the differences between the regression and direct expansion estimates were less than the standard error of the latter in all but one case for corn and for all except two cases for soybeans. For corn the exception was domain E12 where the difference between the regression and direct expansion estimates was 2.04 standard errors of the direct expansion estimate. For soybeans the exceptions were domains W123 and E12, where the differences between the two types of estimates were between one and two standard errors of the direct expansion estimate.

C. Single-County Crop-Acreage Estimates

Regression estimates were computed for corn and soybeans for each county wholly contained in a LANDSAT frame or pseudo-frame. The actual calculated estimates are tabled in Appendix D. The classifiers used for the single-county estimates were the same classifiers that were used for multi-county estimates.

Figures 6 and 7 plot the final 1975 SSO acreage estimates versus the corresponding regression estimates for soybeans and corn, respectively, in all of the individual counties. In the case of the overlap counties, the estimates for both domains containing the county are plotted.

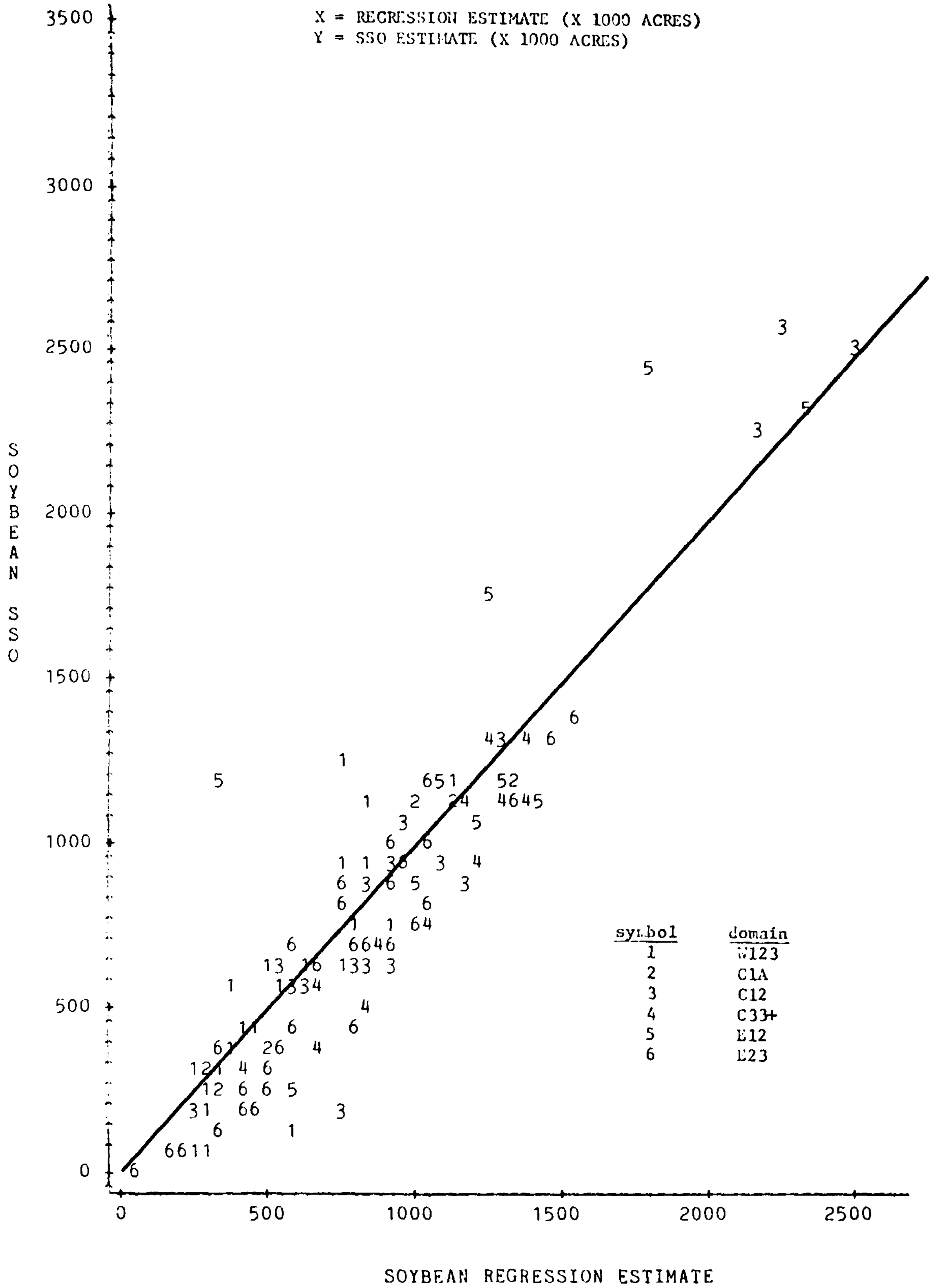


Figure 6. Comparison of regression and SSO county estimates for soybeans.

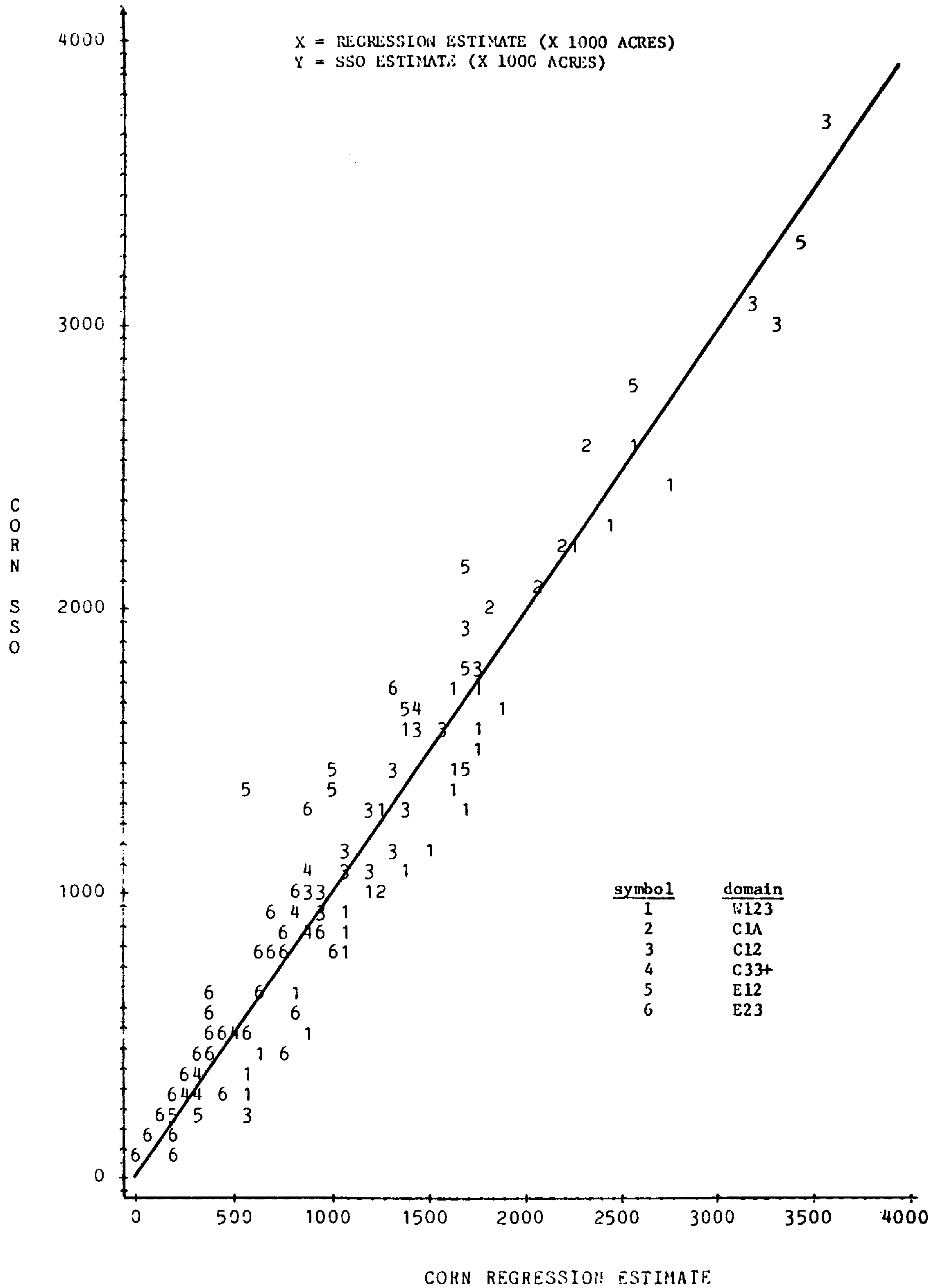


Figure 7. Comparison of regression and SSO county estimates for corn.

For the county soybean estimates in Figure 6, it appears that in a majority of the counties the regression estimate exceeds the SSO value. Moreover, the frequency of the regression over-estimation varies with domain. For example, in domain C33+ nearly all of the regression county estimates for soybean acreages exceed the corresponding SSO county estimates.

In Appendix D it can be seen that if a county is quite dissimilar in land use from its containing domain, then the county regression estimate based on that domain deviates markedly from the county SSO estimate. An example of this is Dupage county which is in domains C12 and E12. Dupage is essentially a suburb of Chicago. Thus, with regards to land use it is more like domain E12 than like domain C12. As can be seen in Figure 6, in Dupage county the soybean regression estimate based on domain E12 is closer to the SSO estimate than is the regression estimate based on domain C12. In fact, in domain E12 it appears that the soybean regression estimates deviate very little from the SSO values in urban counties such as Cook, Dupage, and Champaign, but in highly agricultural counties, such as Ford, Vermillion, and Iroquios, there are quite large differences between the regression and SSO values. This effect is a result of the the highly heterogeneous land-use pattern of domain E12.

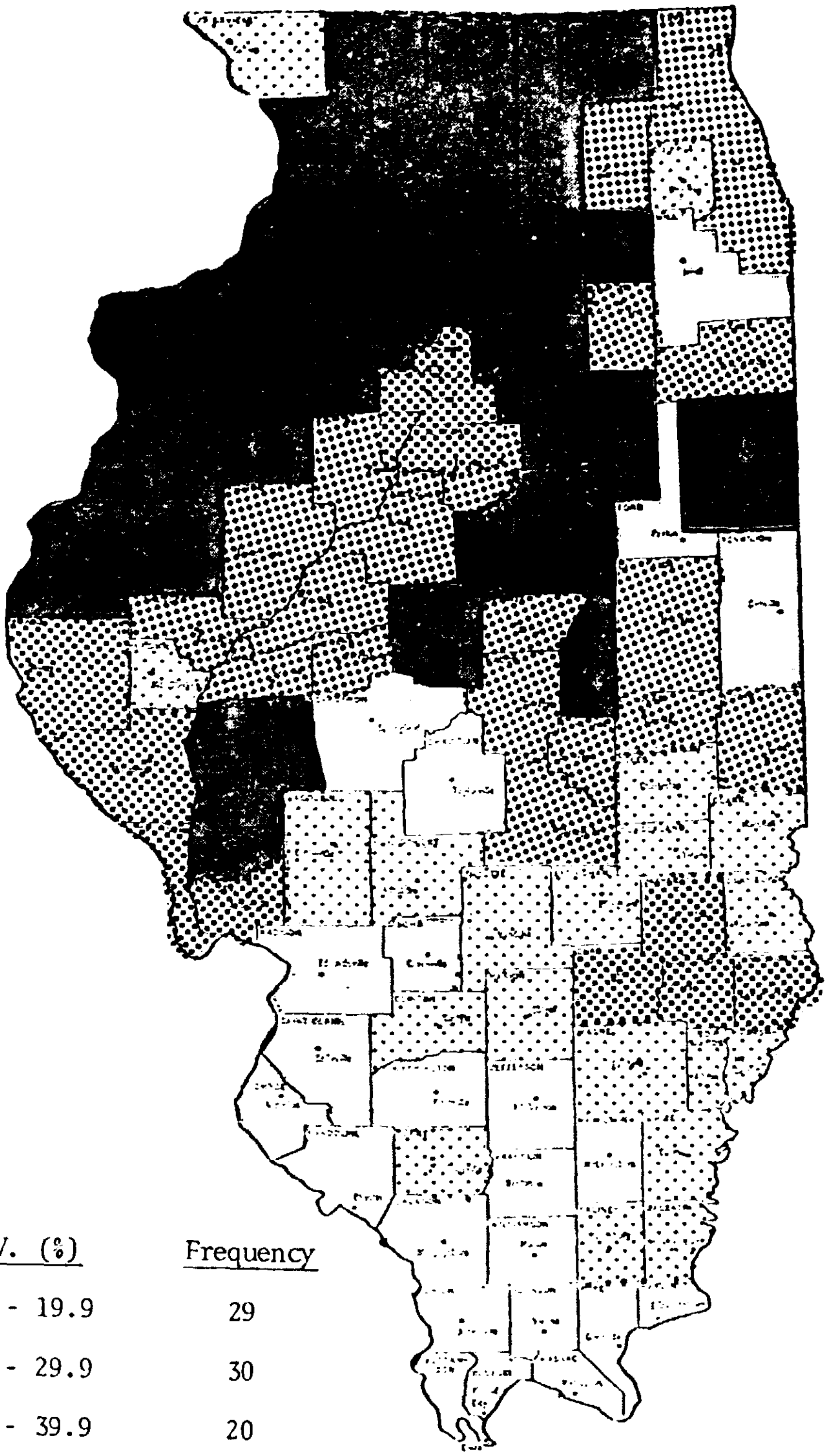
Though Figures 6 and 7 have different scales, it is apparent that there is better agreement between the regression and SSO estimates for corn in Figure 7 than for soybeans in Figure 6. This is further evidenced by the correlations between the two estimates. For the entire state the correlation between regression and SSO estimates is .96 for corn and .91 for soybeans. In domain E12 the correlation between estimates is .95 for corn and .85 for soybeans. In Figure 7 it appears that for corn, unlike soybeans, the number of positive differences between regression and SSO estimates is nearly equal to the number of negative differences. However, several of the domain effects observed for soybean regression estimates persist for the corn regression estimates. For example, the regression estimate for corn acreages are less than the SSO estimates in the agricultural counties of domain E12, as was also the case for soybeans. Moreover, in domain C33+ the differences between regression and SSO estimates for corn are all in the same direction. For corn the regression estimator consistently overestimates in C33+, whereas for soybeans it consistently underestimates there.

The coefficients of variation* for the corn and soybeans regression estimates are mapped in Figures 8 and 9, respectively. In the case of the overlap counties, the lower C.V. is used. The distributions of the C.V.'s are indicated in the figure legends.

As can be seen in Figures 8 and 9, many large C.V.'s for county regression estimates for corn and soybeans occurred--41% of the C.V.'s of county regression estimates for corn acreages exceeded 30%. Similarly, for soybeans 47% of the C.V.'s exceeded 30%. Some moderately small C.V.'s were obtained, however, in domain C1A, for example, all of the county regression C.V.'s for corn were between 10.0 and 12.0%.

---

*As is explained more fully in Appendix B, the variances, and hence coefficients of variation, of the single-county regression estimates given in this report are possibly overstated.



C.V. (%)	Frequency
10.0 - 19.9	29
20.0 - 29.9	30
30.0 - 39.9	20
40.0+	21
TOTAL	100

Figure 8. Distribution of Coefficients of Variation (C.V.'s) of County Regression Estimates for Corn

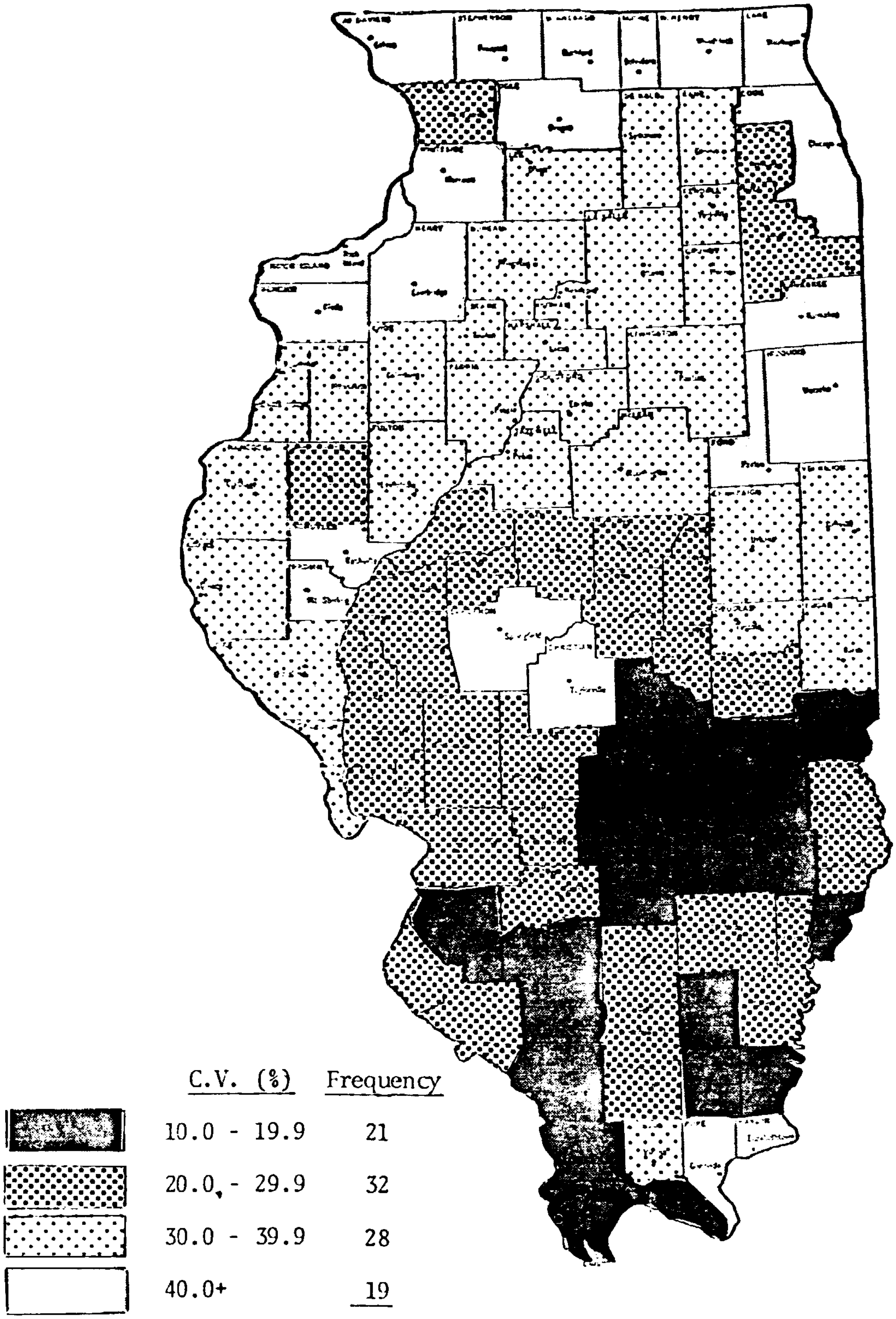


Figure 9. Distribution of Coefficients of Variation (C.V.'s) of County Regression Estimates for Soybeans

In Figure 8 it can be seen that the C.V.'s of county regression estimates for corn are lowest in northern Illinois and highest in the southern part of the state. Figure 9 shows that the opposite is true for soybeans--high C.V.'s in northern Illinois and low C.V.'s in the southern part of the state.

The magnitudes of the regression estimate C.V.'S are partially explained by the very magnitudes of the regression estimates themselves. Figures 10 and 11 show that many of large C.V.'s were for regression estimates which were small in magnitude, and conversely many of the small C.V.'s were for regression estimates which were large in magnitude. Large C.V.'s also occurred in areas where there was considerable spectral confusion. For corn, large C.V.'s occurred in the southern part of Illinois, where considerable spectral confusion between corn and trees occurred. For soybeans, large C.V.'S occurred in the northern part of the state where considerable confusion between soybeans and corn occurred.

Tables 11 and 12 present the regression estimates for the sixteen overlap counties. Because each overlap county is contained in two domains, each tabled county has two regression estimates for each crop. The difference between these two regression estimates, referred to as the overlap difference, was compared in each overlap county to the larger of the standard errors of the two regression estimates, denoted S2.

For corn, six of the sixteen overlap differences exceeded the corresponding S2's. This occurred in the four counties overlapped by C33+ (July 16) and W123 (August 4) and in the two counties overlapped by

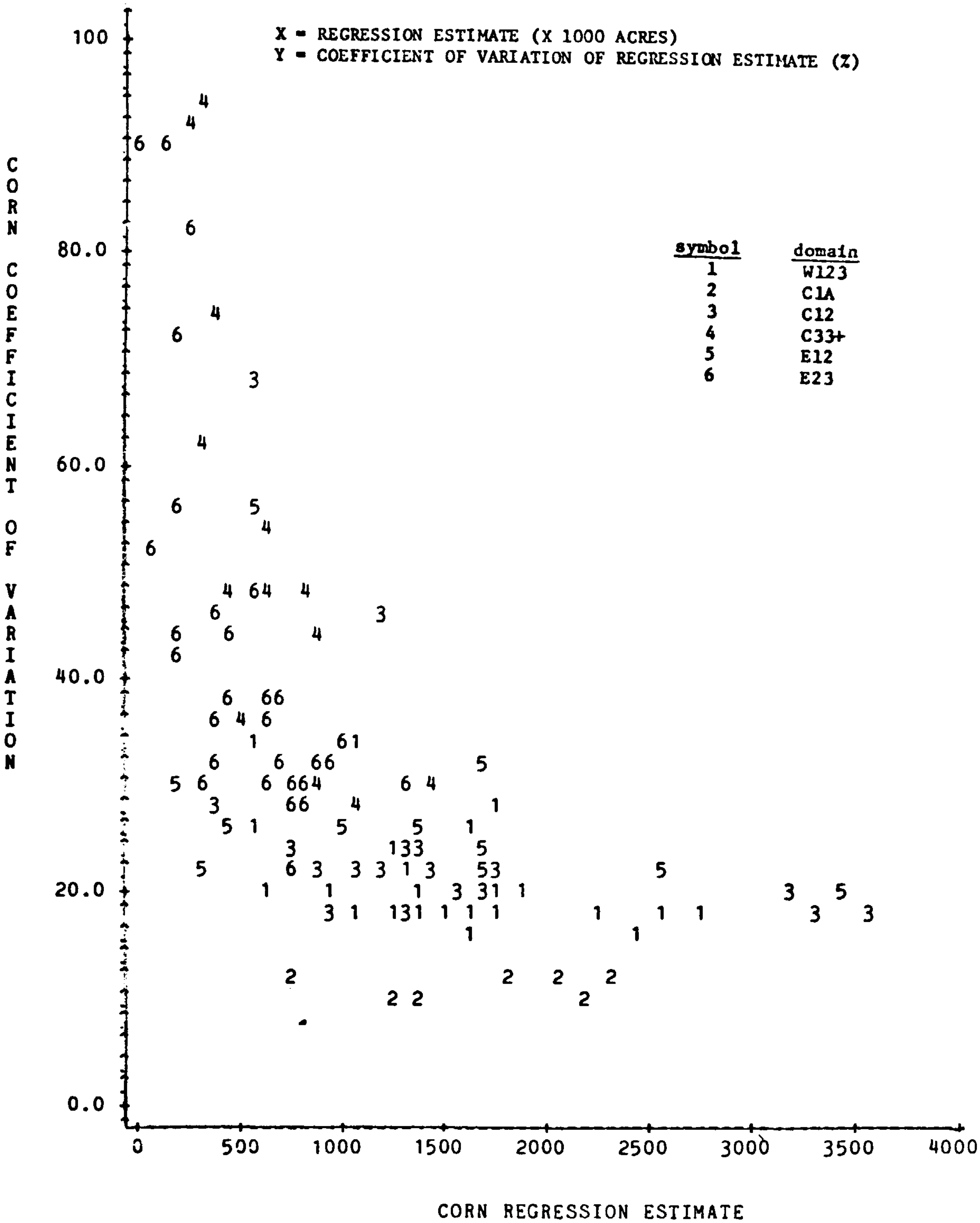


Figure 10. Coefficient of variation of regression estimates as a function of the regression estimates for corn in individual counties.

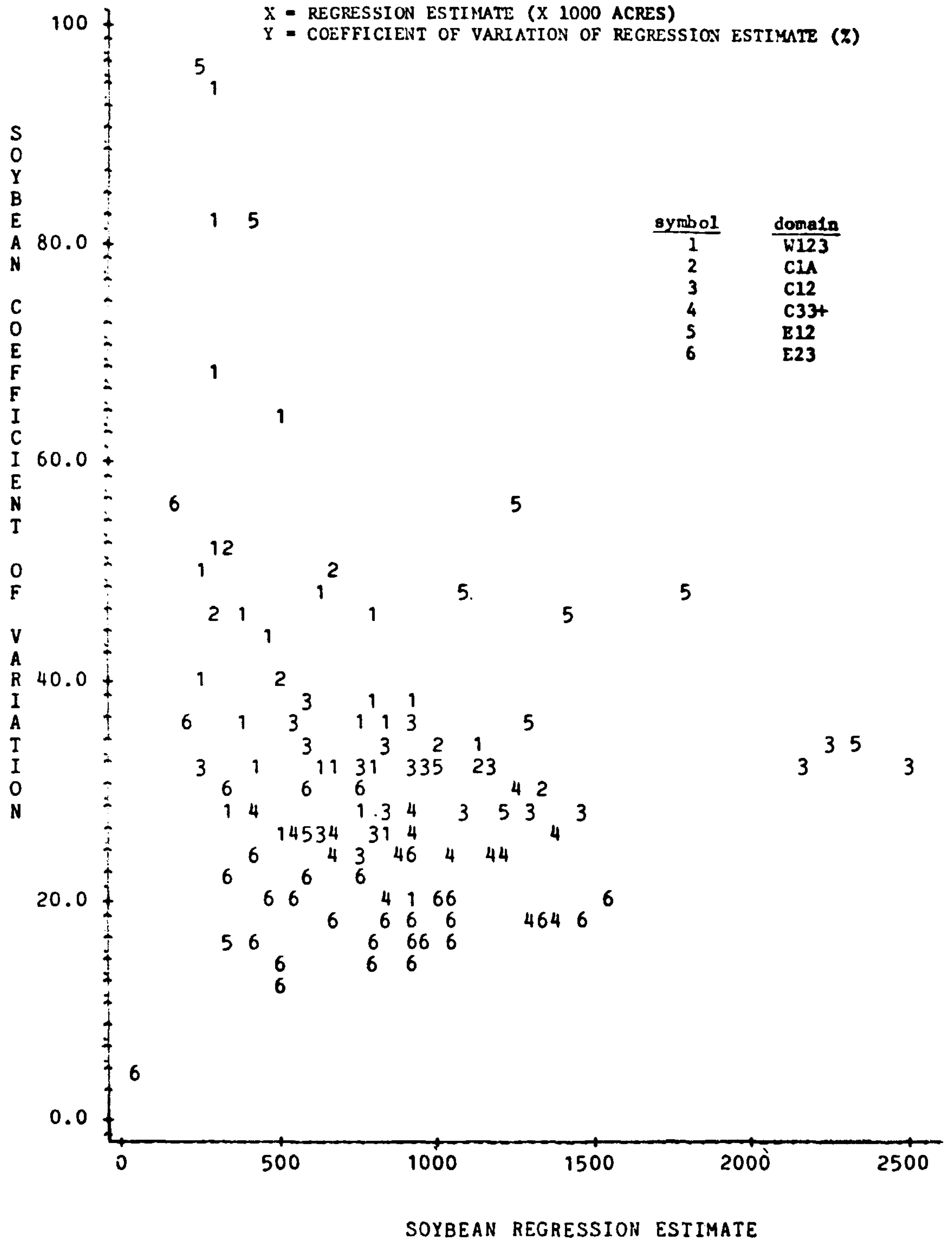


Figure 11. Coefficient of variation of regression estimates as a function of the regression estimates for soybeans in individual counties.

Table 11. Comparison of estimated acres for overlap counties, western and central passes.

		CORN				SOYBEANS			
County	Scene	SSO Estimate	Regression Estimate	S.D.	C.V.	SSO Estimate	Regression Estimate	S.D.	C.V.
Bureau	C1A	260,400	231,950	27,800	12.0	116,600	132,600	40,000	30.2
	W123		254,000	47,500	18.7		110,600	36,900	33.4
Winnebago	C1A	103,000	122,950	13,000	10.6	27,500	33,600	17,400	51.8
	W123		121,500	26,100	21.5		29,600	20,100	68.0
Ogle	C1A	218,000	217,350	23,900	11.0	62,700	68,200	34,200	50.2
	W123		223,000	42,400	19.0		51,500	33,100	64.2
Stark	C12	100,000	96,674	16,752	17.3	42,500	56,644	21,297	37.6
	W123		91,976	16,729	18.2		30,584	25,015	81.8
Peoria	C12	126,000	121,081	56,703	46.8	63,200	84,302	28,825	34.2
	W123		123,965	29,767	24.0		65,320	21,309	32.6
Mason	C12	112,500	132,433	31,782	24.0	92,400	110,136	30,158	27.4
	W123		129,142	27,469	21.3		76,143	21,241	27.9
Morgan	C33+	111,000	103,958	28,671	27.6	73,000	102,606	25,036	24.4
	W123		147,200	25,850	17.6		93,735	19,591	20.9
Scott	C33+	46,000	40,303	13,078	32.4	32,500	40,350	11,101	27.5
	W123		61,073	12,140	19.9		31,501	9,019	28.6
Greene	C33+	104,000	87,194	26,238	30.1	62,200	90,831	23,616	26.0
	W123		136,766	26,275	19.2		76,003	18,840	24.8
Jersey	C33+	53,000	50,535	17,803	35.2	38,100	53,207	13,677	25.7
	W123		85,705	18,528	21.6		48,891	13,179	27.0

Table 12. Comparison of estimated acres for overlap counties, central and eastern passes.

		CORN				SOYBEANS			
County	Scene	SSO Estimate	Regression Estimate	S.D.	C.V.	SSO Estimate	Regression Estimate	S.D.	C.V.
Dupage	C12	19,500	55,961	38,267	68.4	18,900	76,285	17,668	23.2
	E12		17,743	5,352	30.2		26,568	25,289	95.2
Platt	C12	145,000	133,491	24,633	18.5	105,000	97,801	30,948	31.6
	E12		100,748	25,970	25.8		119,443	34,161	28.6
Fayette	E2+*	83,500	93,491	29,977	32.1	112,600	134,861	23,303	17.3
	C33+**		86,836	37,505	43.2		120,476	29,667	24.6
Moultrie	E2+	98,500	80,531	22,546	28.0	69,700	57,148	17,086	29.8
	C33+**		85,490	18,541	21.7		77,733	20,487	26.4
Perry	E3+	29,500	42,010	15,727	37.4	44,300	79,496	12,359	15.5
	C33+		23,290	21,572	92.6		83,149	15,994	19.2
Jackson	E3+	27,500	42,019	18,169	43.2	69,300	93,749	13,682	14.6
	C33+		29,421	27,767	94.4		86,103	21,189	24.6

*11-12-20-30

**10-50

C12 (August 21) and E12 (September 7). In the latter two counties the overlap differences were between 1.0 and 1.5 times S2, and in the four C33+/W123 counties the overlap differences were from 1.5 to 2.0 times S2.

For soybeans, four overlap differences exceeded the corresponding S2 values. Of these four, three were between 1.0 and 1.5 times S2--Stark, Mason, and Moultrie, where the corn overlap differences were all less than corn S2's--and one was between 1.5 and 2.0 times S2, namely Dupage, where the corn overlap difference was also greater than its corresponding S2 value.

Even though many of the overlap differences were less than or only slightly larger than S2, a number of the overlap differences were nevertheless fairly large because S2's were large. For example, for corn, in Dupage county the regression estimate based on domain C12 was more than 300% above the regression estimate based on domain E12. This was caused by the different land-use distributions in the two domains and by the different strata poolings used for county estimates in E12 and C12. (The same strata pooling is used for all county estimates in the same domain, however.) The E12 estimates were made using a "swiss cheese" estimator for pooled stratum 30; i.e., in domain E12, regression estimates were computed for strata 11, 12, and 20 and a direct expansion estimate was computed for stratum 30. This eliminated a commission-error bias in the regression estimate which would have occurred had stratum 30 been used for regression. In domain C12, however, stratum 30 was pooled with strata 12 and 20. For the corn regression estimator based on domain C12, the stratum estimates for corn in Dupage county were the following:

Strata	Estimate	Standard Deviation	C.V.(%)
11	18894	4556	24.1
(12,20,30)	37067	37995	102.5
TOTAL	55961	38267	68.4

Note that the contribution of pooled stratum (12,20,30) was 67% of the total estimate. If instead a "swiss cheese" estimate had been used, the contribution of pooled strata (12,20,30) would have been considerably less.

V. Conclusions

It was found that classifier performance was influenced by a number of temporal, methodological, and geographical factors*. Best results were obtained when corn was tasselled and near dough stage of development. Dates earlier or later in the growing season produced poor results. However, the effects of atmosphere on the results obtained cannot be independently measured or completely separated from the effects due to the maturity stage of the crops. Also, poor classifier performance was observed in areas where considerable spectral confusion was present. This suggests that multi-temporal LANDSAT data should be investigated as a means to decrease spectral confusion between crops.

*Another factor affecting classifier performance is average field size. The magnitude of this effect is being assessed by comparing the results of the Illinois Crop Acreage Experiment to results from similar studies in other states. These comparisons will be presented in future reports. Average field sizes in acres in Illinois by crop type were woods - 21.1, corn - 29.1, oats - 14.2, winter wheat - 17.9, sorghum - 14.6, soybeans - 28.9, alfalfa - 14.4, clover - 12.0, and permanent pasture - 17.0.

Resubstitution was found to be an acceptable method of classifier training and testing for a classification domain which contained 84 segments. Equal priors proved to be the best type of "prior probabilities" to use for estimating corn acreages. However, for soybeans, the best type of priors varied by domain. Minor crops could not be distinquished with any degree of consistency or accuracy and it is felt that the project methodolgy will not improve minor crop acreage estimates.

For major crops, however, increases in precision of acreage estimates for counties and groups of counties can be achieved using LANDSAT data with the methodology developed in this project. However, the large coefficients of variation make the majority of the county estimates unsuitable for operational use with the present area-sample size. Nevertheless, estimates for groups of counties appear quite encouraging when sufficient spectral separability is present in the LANDSAT data. The reported variances of the single-county regression estimates may be overstated but are, nevertheless, a function of spectral separability and regression-domain homogeneity.

In order to perform the developed methodology, LANDSAT frames had to be joined together in several cases to provide sufficient data for designing the classifier and for estimating strata regression parameters. It is felt that when an adequate number of segments for classifier training and testing is available that only 8 to 14 counties should define a regression domain. These counties should be spatially contiguous and the resulting domain should be as homogeneous as possible with regards to intensity of cultivation.

Finally, even though the ability of LANDSAT data to improve acreage estimates varied widely across the data sets analyzed, it is felt that when improved sensor technology is realized or possibly in geographical areas with larger field sizes that the developed methodology may provide county acreage estimates for major crops with precisions suitable for operational use.

VI. References

1. Ray, Robert M., III and Huddleson, Harold F., "Illinois Crop-Acreage Estimation Experiment", Proceedings of the 1976 Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana.

2. Cook, Paul, ""Procedures for Boundary Location in LANDSAT MSS Digital Data", Statistical Reporting Service, United States Department of Agriculture, Washington, D.C., in progress.

3. Donovan, Walter E.; Ozga, Martin; and Ray, Robert M.; "Compilation and Geometric Registration of ERTS Multitemporal Imagery," CAC Technical Memo No. 52, Center for Advanced Computation, University of Illinois at Urbana-Champaign, Urbana, Illinois, May 1975.

4. Donovan, Walter E., "Oblique Transformation of ERTS Images to Approximate North-South Orientation," CAC Technical Memo No. 38, Center for Advanced Computation, University of Illinois at Urbana-Champaign, Illinois, November 1974.

5. Ozga, Martin; Donovan, Walter E.; and Gleason, Chapman P., "An Interactive System for Agricultural Acreage Estimates Using LANDSAT Data," Proceedings of the 1977 Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana.

6. Starbuck, Robert R., "Overview and Examples of the EDITOR System for Processing LANDSAT Data", Statistical Reporting Service, United States Department of Agriculture, Washington, D.C., March 1977.

7. Sigman, Richard R.; Gleason, Chapman P.; Hanuschak, George A.; and Starbuck, Robert A.; "Stratified Acreage Experiments in the Illinois Crop-Acreage Experiment", Proceedings of the 1977 Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana

8. Huddleston, Harold F., and Ray, Robert M. III, "A New Approach to Small Area Crop-Acreage Estimation", Annual Meeting of the American Agricultural Economics Association, Pennsylvania State University, State College, Pennsylvania, August 1976.

9. Wigton, William H., "Use of LANDSAT Technology by Statistical Reporting Service", Proceedings of the 1976 Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana.

10. Toussaint, Godfried T., "Bibliography on Estimation of Missclassification", IEEE Transactions on Information Theory. Vol. IT-20, No. 4, (July, 1974), pages 472-79.

11. "Illinois Weekly Weather and Crop Bulletin", Illinois Cooperative Crop Reporting Service, Springfield, Illinois, 1975.

12. "Crop Identification Technology Assessment for Remote Sensing (CITARS), Volume X, Interpretation of Results", NASA, Lyndon B. Johnson Space Center, Houston, Texas, December, 1975.

13. Von Steen, Donald H. and Wigton, William H., "Crop Identification and Acreage Measurement Utilizing LANDSAT Imagery", Statistical Reporting Service, United States Department of Agriculture, Washington, D.C., March 1976.

VII. CONTRIBUTIONS BY AUTHORS

Chapman Gleason: Project Coordinator; Analysis of W123, E2+, E3+; systems development; statistical methodology; ground data collection coordinator; project write-up.

Robert Starbuck: Analysis of W123 jackknifing, E1, E2; systems development; statistical methodology.

Richard Sigman: Analysis of W1, C1+, C2, C3, C3+; systems development; statistical methodology; project write-up.

George Hanuschak: Analysis of W2, C1+, C2; ground data collection; statistical methodology.

Michael E. Craig: Analysis of C1A.

Paul W. Cook: Registration of C2, C3, C4, E2, E3, E4.

Richard D. Allen: Coordination and management of the ground data collection in Illinois.

Appendix A

Supplementary Survey Questionnaires
for 1975 Illinois Crop
Acreage Experiment

Questionnaires:

UNITED STATES DEPARTMENT OF AGRICULTURE
STATISTICAL REPORTING SERVICE
Form Approved
O. M. B. Number 40-S75033
Approval Expires 12-31-75

Segment ________ Page ___ of ___
State ________ County ________
Enumerator ________

1975 SATELLITE CROP INFORMATION SUPPLEMENT

Complete one line for each JES Section A field except Item 5 (Farmstead, roads, ditches, woods, etc.)

Month of Visit (Check)

- ☐ July = 07 (Green pencil)
- ☐ August = 08 (Blue pencil)
- ☐ September = 09 (Red pencil)

Time Spent In Segment

Month	Starting	Ending	Office Use
July........			
August......			
September....			

Code	Field Appearance Definition	Code	Field Appearance Definition
10	Green cover (not in planted crop)	50	Mature (turning or ready for harvest)
20	Prepared land (worked land including planted but not emerged)	60	Harvested crop (but not worked or prepared)
30	Emerged (less than 50% of field covered with green foliage)	70	Dried or cut vegetation (brown pasture, cut hay, etc.)
40	Green (50% or more of field covered with green foliage, but not mature)	80	None of above, (explain in notes)

Date of Visit (Mo./Day)	PREVIOUS VISIT: Tract Code Letter	Field Number: JES	Field Number: Followup	Office Use	Crop or Land Use (Specify)	Acres In: Field (Acres)	Acres In: Crop (Acres)	Has crop or land use changed for any part of field since previous visit? (Check)	If yes, enter date of harvest or land use change (Mo./Day)	Has field boundary changed? (Check) If yes, correct and assign new field number(s). Complete 1–4 and skip to 12.	CURRENT VISIT: Acres In: Field (Acres)	Acres In: Crop (Acres)	Office Use	Crop or Land Use (Specify)	Office Use	Intended Use (Specify)	Field appearance Code	Notes On Crop Or Field Conditions
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						
/						.	.	YES ☐→ NO ☐→18	/	YES ☐ NO ☐→12	.	.						

UNITED STATES DEPARTMENT OF AGRICULTURE
STATISTICAL REPORTING SERVICE
Form Approved
O. M. B. Number 40-R2766
Approval Expires 4-30-77

State ____________ Page ___ of ___

Enumerator ____________

1975 ILLINOIS AND KANSAS SATELLITE CROP INFORMATION SUPPLEMENT

COMPLETE FOR ALL FIELDS WITH PASTURE, HAY CROPS, AND FALL SEEDED SMALL GRAINS.

Code	Definition
10	Green cover (not in intended crop).
20	Prepared land (including planted but not emerged).
30	Emerged (less than 50% of field covered with green foliage.
40	Green (50% or more of field covered with green foliage but not mature).
50	Mature (turning or ready for harvest).
60	Crop has been harvested.

Date of Visit	Segment Number	Tract Code Letter	Field Number	Acres In: Field (Acres)	Acres In: Crop (Acres)	Office Use	Current crop or land use	Office Use	Intended crop use	Field Code (See code above)	Date of harvest, if field has been harvested	Notes on unusual planting patterns or crop condition
1	2	3	4	5	6	7	8	9	10	11	12	13

UNITED STATES DEPARTMENT OF AGRICULTURE
STATISTICAL REPORTING SERVICE
ILLINOIS COOPERATIVE CROP REPORTING SERVICE
ROOM 218 P.O. BUILDING P.O. BOX 429
SPRINGFIELD ILLINOIS 62705

1975 SATELLITE CROP INFORMATION SUPPLEMENT

FORM APROVED
O. M. B. NUMBER 40- S75033
APPROVAL EXPIRES 12/31/75

XXXX SEGMENT
1-A TRACT
ILLINOIS

XXX XXX-XXXX TELEPHONE

DOE, JOHN
R.R.1
STILLMAN HOLLOW LOG

XXXX SEGMENT
1-A TRACT

DATE OF VISIT (MO/DAY) ____/____

ENUMERATOR ______________________________

PREVIOUS VISIT						HAS CROP OR LAND USE CHANGED FOR ANY PART OF THE FIELD SINCE PREVIOUS VISIT?		HAS FIELD BOUNDARY CHANGED?	CURRENT VISIT						
FIELD NUMBER		CROP OR LAND USE		ACREAGE IN					ACREAGE IN:						
		CODE	NAME	FIELD	CROP	(CIRCLE YES OR NO)	ENTER DATE OF HARVEST OR CHANGE (MO./DAY)	(CIRCLE YES OR NO)	FIELD XXX.X	CROP XXX.X	OFFICE	CROP OR LAND USE (SPECIFY NAME)	OFFICE	INTENDED USE (SPECIFY USE)	FIELD CARD CODE
-3-	-4-	-5-	-6-	-7-	-8-	-9-	-10-	-11-	-12-	-13-	-14-	-15-	-16-	-17-	-18-
						YES >>>> NO>-18-	____/____	YES >>>>>> NO >>	____.__	____.__					

FIELD # 1 NOTES: ______________________________

						YES >>>> NO>-18-	____/____	YES >>>>>> NO >>	____.__	____.__					

FIELD # 2 NOTES: ______________________________

						YES >>>> NO>-18-	____/____	YES >>>>>> NO >>	____.__	____.__					

FIELD # NOTES: ______________________________

						YES >>>> NO>-18-	____/____	YES >>>>>> NO >>	____.__	____.__					

FIELD # NOTES: ______________________________

SEGMENT 5001
TRACT 1-A
PAGE 1

ILLINOIS SATELLITE DISCREPANCY CORRECTION FORM

The following discrepancies were noticed during review of the aerial photos and the field information obtained from the JES. In most cases boundaries and field acreages are in question. Please resolve the following inconsistencies and add any comments that may further explain the situation. RETURN THIS CORRECTION FORM WITH THE SEGMENT KIT AFTER THE SATELLITE VISIT.

Enumerator ______________________________

Segment ________________

Tract	Field	Data Problems to be Resolved	Resolved - Check and Explain		
			Yes	No	Notes

INSTRUCTIONS FOR COMPLETING 1975 SATELLITE CROP INFORMATION SUPPLEMENT

I. PURPOSE:

Research is being conducted this crop year in Kansas, Illinois, and Texas to investigate the potential operational use of satellite data to improve crop acreage estimates at the State and County level. Crop or land use information collected in the June Enumerative Survey (JES) along with followup visits to the segments will be used to aid in computer identification of different crops.

You will be either conducting an interview with the tract operator or observing each agricultural field in agricultural tracts and recording its crop or land use. If the crop or land use has changed since the last time the field was visited, the current crop or land use is to be recorded, and the date of harvest or land use change is to be acquired from the tract operator.

II. DEFINITIONS:

A. All JES definitions hold including:

Field - a continuous area of land inside a tract devoted to one crop or land use.

B. For this survey, some additional clarification of crop or land use is as follows:

Crop - record the crop name for an field seeded to one agricultural product, such as winter wheat.

Land use - record a specific use for a field not in any planted crop.

Examples are permanent pasture (note type of grass grown), summer fallow and idle crppland. NOTE: Alfalfa hay is a crop use and not a land use.

Change in Crop or Land Use from Previous Visit - a crop change refers to any change from the previously reported crop planted (winter wheat to soybeans, etc.) or crop appearance (winter wheat now harvested to idle cropland or alfalfa just cut for hay). A land use change refers to any change in land utilization such as cropland pasture now plowed up or summer fallow now planted to winter wheat.

III. PROCEDURE FOR INTERVIEW OR OBSERVATION

A. PRIOR TO VISITING SEGMENT *(At home before enumeration)*

1. Columns 2, 3, 6, 7, and 8: For the July visit, complete these columns by copying the crop/land use from Line 2 of the JES tract questionnaire and acreage field data from the JES Section A - Acreage of Fields and Crops in Tract. Copy information for all tracts with agricultural field data reported in Section A of JES Part A questionnaire. You do not have to record or observe any field which farmstead, roads, ditches, woods, etc. (Any JES Line 5 field).

2. Column 4 (Followup field number): This column must be used when a JES field is subdivided and different crop or land use is made of any part of a field since the time of the previous visit.

3. Identify tracts where the operator will have to be contacted. These tracts can be identified since they were selected for a July Update or Objective Yield interview or because there is a likelihood of a crop or land use change for a field since the last time the segment/tract was visited. Examples of fields likely to have changes are: winter wheat, any hay crop, intentions to plant a spring sown crop or harvest of a spring sown crop such as soybeans. Contact the operators of these tracts and obtain the field information for the satellite supplement without observing fields.

   Try to observe the fields in tracts not to be contacted. If necessary, contact the tract operator to obtain the satellite supplement information.

B. VISITING THE SEGMENT

1. Tract operators requiring an interview - For all operators requiring a visit, obtain satellite supplement information for each agricultural field in the tract. Interview the tract operator if this is possible. If operator is not available, obtain survey data from a reliable source, such as wife, hired man, etc. Follow the instructions as given on the supplement for the interview.

2. Tract operators not requiring an interview - Observing crop/land use and field appearance instructions are as follows:

Task 1: Locate the tract and record the starting time (Military) when you started to observe fields. Record ID information in upper right hand corner.
Task 2: Enter date of visit in Column 1, example (July 24 = 07/24).
Task 3: Verify the pre-entered tract and field data for the tract in Columns 2 through 8.
Task 4: Complete the field observation and verification. Observe each field in the tract by driving past the field and identifying the field's current crop or land use. If no portion of the field has changed land use from the previous visit, check a "no" (Column 9) and enter the field appearance code (Column 18). Then complete any notes on this particular field in Column 19. When the crop or land use has changed, follow the Flow Diagram for Task 4 to record the changes.
Task 5: Verify the pre-entered tract and field data (Columns 2, 3, 6, 7, & 8) for another tract in the segment and continue until all tracts are covered.
Task 6: Contact tract operator(s) for fields that have crop or land use changes since the previous visit. (Yes, checked in Column 9), and complete two or more lines for each field with a crop or land use change. *(See Task 6 in the flow diagram)*.
Task 7: Record ending time when you leave segment.

C. <u>AFTER VISITING THE SEGMENT</u>

1. For the August and September visits: Copy the previous visits field data into Columns 2 through 8. Pre-enter data in Column 4 <u>only</u> when a JES field has been subdivided into two or more fields on a previous visit.

2. <u>Mail</u> the completed Satellite Supplement for the visit just completed to the State Office in the envelope provided.

Appendix B

Estimation Methods and Classifier Design Procedures Used in the Illinois Crop Acreage Experiment*

I. STATISTICAL THEORY AND METHODOLOGY

A. DIRECT EXPANSION ESTIMATION (GROUND DATA ONLY)

Aerial photography obtained from the Agricultural Stabilization and Conservation Service is photo-interpreted using the percent of cultivated land to define broad land-use strata. (See Table B1.) Within each stratum, the total area is divided into N_h area frame units. This collection of area frame units for all strata is called an area sampling frame. A simple random sample of n_h units is drawn within each stratum. The Statistical Reporting Service then conducts a survey in late May, known as the June Enumerative Survey (JES). In this general purpose survey, acres devoted to each crop or land use are recorded for each field in the sampled area frame units. Intensive training of field statisticians and inteviewers is conducted providing rigid controls to minimize non-sampling errors.

The scope of information collected on this survey is much broader than crop acreage alone. Items estimated from this survey include crop acres by intended utilization, grain storage on farms, livestock inventory by various weight categories, and agricultural labor and farm economic data.

Let h = 1, 2,..., L be the L land-use strata. For a specific crop (corn, for example) the estimate of total crop acreage for all purposes and the estimated variance of the total are as follows:

Let Y = Total corn acres for a state (Illinois, for example).

$\hat{Y}$ = Estimated total of corn acres for a state.

y_{hj} = Total corn acres in j^{th} sample unit in the h^{th} stratum.

Then

$$\hat{Y} = \sum_{h=1}^{L} N_h \left(\sum_{j=1}^{n_h} y_{hj} \right) / n_h \qquad (1)$$

---

*Excerpted from Sigman, Richard R.; Gleason, Chapman P.; Hanuschak, George A.; and Starbuck, Robert A.; "Stratified Acreage Experiments in the Illinois Crop-Acreage Experiment", Proceedings of the 1977 Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana.

The estimated variance of the total is:

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

Note that we have not yet made use of an auxiliary variable such as classified LANDSAT pixels. The estimator in (1) is commonly called a direct expansion estimate, and we will denote this by $\hat{Y}_{DE}$.

As an example, for the state of Illinois in 1975, the direct expansion estimates were:

Corn $\hat{Y}_{DE}$ = 11,408,070 Acres
Relative Sampling Error = 2.4% = $\sqrt{v(\hat{Y})} / \hat{Y}$

Soybeans $\hat{Y}_{DE}$ = 8,569,209
Relative Sampling Error = 2.9% = $\sqrt{v(\hat{Y})} / \hat{Y}$

B. REGRESSION ESTIMATION (GROUND DATA AND CLASSIFIED LANDSAT DATA)

The regression estimator utilizes both ground data and classified LANDSAT pixels. The estimate of the total Y using this estimator is:

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

where

$$\bar{y}_{h(reg)} = \bar{y}_h + \hat{b}_h (\bar{X}_h - \bar{x}_h)$$

and $\bar{y}_h$ = the average corn acres per sample unit from the ground survey for the h^{th} land-use stratum

$$= \sum_{j=1}^{n_h} y_{hj} / n_h$$

$\hat{b}_h$ = the estimated regression coefficient for the h^{th} land-use stratum when regressing ground-reported acres on classified pixels for the n_h sample units.

$$= \frac{\sum_{j=1}^{n_h} (x_{hj} - \bar{x}_h)(y_{hj} - \bar{y}_h)}{\sum_{j=1}^{n_h} (x_{hj} - \bar{x}_h)^2}$$

$\bar{X}_h$ = the average number of pixels of corn per frame unit for <u>all</u> frame units in the h^{th} land-use stratum. Thus <u>whole</u> LANDSAT frames must be classified to calculate $\bar{X}_h$. Note that this is the mean for the population and not the sample.

$$= \sum_{i=1}^{N_h} X_{hi}/N_h$$

X_{hi} = number of pixels classified as corn in the i^{th} area frame unit of the h^{th} strata.

$\bar{x}_h$ = the average number of pixels of corn per sample unit in the h^{th} land-use stratum

$$= \sum_{j=1}^{n_h} x_{hj}/n_h \ .$$

x_{hj} = number of pixels classified as corn in the j^{th} sample unit in the h^{th} strata.

The estimated (large sample) variance for the regression estimator is

$$v(\hat{Y}_R) = \sum_{h=1}^{L} \frac{N_h^2}{n_h} \frac{N_h - n_h}{N_h} \cdot \sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)^2 \cdot \frac{1 - r_h^2}{n_h - 2}$$

where

r_h^2 = sample coefficient of determination between reported corn acres and classified corn pixels in the h^{th} land-use stratum.

$$= \frac{\left[\sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)(x_{hj} - \bar{x}_h)\right]^2}{\left[\sum_{j=1}^{n_h} (y_{hj} - \bar{y}_h)^2\right]\left[\sum_{j=1}^{n_h} (x_{hj} - \bar{x}_h)^2\right]}$$

Note that,

$$v(\hat{Y}_R) = \sum_{h=1}^{L} \frac{n_h - 1}{n_h - 2} (1 - r_h^2)\, v(\hat{Y}) \qquad (2)$$

and so lim $v(\hat{Y}_R) = 0$ as $r_h^2 \to 1$ for fixed n_h. Thus a gain in lower variance properties is substantial if the coefficient of determination is large for most strata.

The relative efficiency of the regression estimator compared to the direct expansion estimator will be defined as the ratio of the respective variances:

$$R.E. = v(\hat{Y}_{DE}) \;/\; v(\hat{Y}_R) \qquad (3)$$

When LANDSAT passes do not cover the entire state on one date, it is necessary to work with analysis districts (domains) which are wholly contained within a LANDSAT scene or pass. In this study the analysis districts were collections of counties wholly contained in a LANDSAT pass. The regression estimate for the i^{th} analysis district is

$$\hat{Y}_{Ri} = \sum_{h=1}^{L_i} N_{hi}\, \bar{y}_{hi(reg)}$$

where

$$\bar{y}_{hi(reg)} = \bar{y}_{hi} + \hat{b}_{hi}\,(\bar{X}_{hi} - \bar{x}_{hi}).$$

When analysis districts are used, degrees of freedom for least squares regression by strata can become small. Under these circumstances it is necessary to pool strata, and the regression estimate for the i^{th} analysis district becomes:

$$\hat{Y}_{Ri} = \sum_{k=1}^{L_i^*} N_{ki}^*\, \bar{y}_{ki(reg)}^*,$$

where L^* = total number of pooled strata for the i^{th} analysis domain,

$$\bar{y}_{ki(reg)}^* = \bar{y}_{ki}^* + \hat{b}_{ki}^*\,(\bar{X}_{ki}^* - \bar{x}_{ki}^*)$$

for $k = 1, 2, \ldots, L_i^*$, and N_{ki}^*, X_{ki}^*, x_{ki}^*, y_{ki}^* are adjusted for varying sizes of the sample units in each stratum. (Thus, h indexes individual stratum; whereas, k indexes pooled stratum. Consequently, the * notation is redundant and will not be used in the next section.)

C. COUNTY ESTIMATES USING A REGRESSION ESTIMATOR

Let $N_{k,c}$ = total number of area frame units in the k^{th} pooled strata for a set of C counties.

$\bar{X}_{k,c}$ = total number of pixels in the set of C counties classified as corn for the k^{th} pooled stratum divided by $N_{k,C}$.

Then an estimate based on the regression estimator of the total corn acreage for the C counties is:

$$\hat{Y}_{REG,c} = \sum_{k=1}^{L} N_{k,c}(\bar{y}_k + \hat{b}_k\,(\bar{X}_{k,c} - \bar{x}_k)) \qquad (4)$$

$$v(\hat{Y}_{REG,c}) = \sum_{k=1}^{L} N_{k,c}^2 \; \frac{N_k - n_k}{N_k} \; S_{k,y}^2 \; \frac{n_k - 1}{n_k - 2} \; .$$

$$(1 - r_k^2) \; (I(C) + \frac{1}{n_k} + \frac{(\bar{X}_{k,c} - \bar{x}_k)^2}{\sum_{i=1}^{n_k} (x_{ki} - \bar{x}_k)^2} \;)$$

where

$I(C) = 1$ if $O(C) <$ total number of counties wholly contained in the analysis district
$= 0$ otherwise

$O(C)$ is the cardinality of the set C.

$S_{k,y}^2$ = variance for the corn reported acreage for the k^{th} pooled stratum

$$= \sum_{j=1}^{n_k} (y_{kj} - \bar{y}_k)^2 \; / \; (n_k - 1)$$

Note that when $I(C) = 1$, the variance formula contains a term which is not present when $I(C) = 0$. This extra term occurs because the statistical treatments of these two cases are quite different. When C is the entire regression domain [$I(C) = 0$], the problem is one of estimating the population total for the regression domain. On the other hand, when C is a subset of the regression domain [$I(C) = 1$], the problem is one of predicting a sub-population total using the stratum regression equations developed for a sample from the entire population.

In this latter case, the variance formula given above is derived by treating the part of C contained in stratum k as a single (fictitous) segment in which the number of pixels classified as the crop of interest is $\bar{X}_{k,c}$. This is equivalent to assuming that there is no variation at all among the "errors"--i.e., reported acres minus regression-predicted acres for the crop of interest--for the (actual) segments in C. If there is such variation, and preliminary investigation suggest that there is, then the stated variance formula is conservative and overstates the variability of the county regression estimates. Attempts to more accurately model the structure of the regression-error are currently being pursued and if successful will be described in a future report.

II. DESIGNING A CLASSIFIER

The pixel classifier is a set of discriminant functions corresponding one-to-one with a set of classification categories. Each discriminant function consists of the category's likelihood probability multiplied by the category's prior probability. If the prior probabilities used are correct for the population of pixels being classified, then the resulting Bayes classifier minimizes the posterior probability of misclassifying a pixel for a 0-1 loss function.

In crop-acreage estimation, however, the objective is to minimize the variance of resulting acreage estimates. Since minimizing the posterior probability of misclassification does not necessarily achieve this objective, optimum acreage estimation may require the use of prior probabilities different than the optimum Bayes set.

For the case of multivariate normal signatures, the category likelihood functions are completely specified by the population means and covariances of the category signatures. Thus, the calculation of category discriminant functions involves the estimation of signature means and covariances and category prior probabilities.

Designing the classifier for this experiment consisted of the fllowing steps:

1. Identification of classification categories.

2. Calculation of signature means and covariances and category prior probabilities from a training set of labeled pixels (called "training the classifier").

3. Measurement of classifier performance on a test set of labeled pixels (called "testing the classifier").

4. Heuristic optimization of the classifier by repeating steps 1 through 3 for different numbers of categories and/or different prior probabilities, and then proceeding to step 5 for the "optimized" classifier.

5. Estimation of classifier performance in classifying the entire pixel population.

Because of the availability of ground data, which supplied the location and cover type of agricultural fields, supervised identification of classification categories was possible. A classification category was created for each cover type in which the number of training pixels exceeded a specified threshold, usually 100 pixels. In addition, a classification category for surface water was created using pixels from rivers, lakes, and ponds.

A classifier was heuristically optimized through a series of classification trials using field-interior pixels to train and all segment-interior pixels to test. The various trials used different combinations of the number of categories and the method of computing prior probabilities.

Table B1. Stratum numbers and definitions

stratum	description	sub-stratum	description	
10	intensive agriculture	11	75%+ cultivated	
		12	50% - 75% cultivated	
50	non-intensive agriculture	20	15% - 49% cultivated	
		31\		\
		32	:urban	:non-
		33/		:agricultural
		40	range land	: (30)
		61	proposed water	:
		62	water	/

Appendix C: Results of Individual Classification Trials

Table C1. Summary of Corn Classification trials for data set W1.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	54	0	.83	5.69	3.03
			10-50	.80,.36	3.95	3.78
			11-12-20-30	.86,.62,.09,1.0		4.25
	PER	88	0	.64	2.74	1.46
			10-50	.56,.50	2.15	2.06
			11-12-20-30	.65,.60,.06,.95		2.46
FLDS	EP	57	0	.84	5.97	3.18
			10-50	.82,.31	4.20	4.02
			11-12-20-30	.89,.57,.15,1.0		4.58
	PER	84	0	.70	3.26	1.74
			10-50	.62,.51	2.44	2.33
			11-12-20-30	.72,.56,.07,.97		2.77

Table C2. Summary of Soybean Classification trials for data set W1.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	72	0	.81	5.25	4.73
			10-50	.82,.83	5.26	4.81
			11-12-20-30	.82,.70,.98,.98		5.56
	PER	74	0	.82	5.42	4.89
			10-50	.83,.83	5.43	4.97
			11-12-20-30	.83,.72,.98,.98		5.76
FLDS	EP	71	0	.81	5.20	4.69
			10-50	.82,.84	5.25	4.81
			11-12-20-30	.82,.75,.99,.98		5.62
	PER	74	0	.82	5.41	4.87
			10-50	.82,.84	5.42	4.96
			11-12-20-30	.82,.72,.97,.98		5.74

Table C3. Summary of Corn for Classification trials for data set W2.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	51	0	.63	2.66	1.61
			10-50	.66,.19	1.68	1.76
			11-12-20-30	.66,.71,.06,.28		1.27
	PER	85	0	.41	1.65	1.00
			10-50	.55,.15	1.47	1.54
			11-12-20-30	.72,.48,.25,.00		1.15
FLDS	EP	54	0	.69	3.16	1.91
			10-50	.74,.30	2.03	2.13
			11-12-20-30	.82,.58,.12,.53		1.67

Table C4. Summary of Soybean Classification trials for data set W2.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	65	0	.62	2.53	2.26
			10-50	.60,.49	2.10	2.18
			11-12-20-30	.73,.31,.63,.55		1.97
	PER	63	0	.63	2.63	2.34
			10-50	.62,.49	2.15	2.23
			11-12-20-30	.73,.38,.58,.55		1.97
FLDS	EP	65	0	.63	2.60	1.67
			10-50	.61,.51	2.16	2.13
			11-12-20-30	.73,.34,.63,.02		1.91

Table C5. Summary of Corn Classification trials for data set W123.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	52	0	.70	3.34	1.73
			10-50	.72,.21	2.23	2.00*
			11-12-20-30	.78,.54,.00,.58		2.23
	PER	86	0	.52	2.08	1.07
			10-50	.56,.18	1.74	1.56
			11-12-20-30	.67,.57,.00,.20		1.81
FLDS	EP	48	0	.75	3.90	2.02
			10-50	.77,.27	2.54	2.28
			11-12-20-30	.86,.47,.01,.70		2.48
	PER	84	0	.57	2.32	1.20
			10-50	.59,.21	1.86	1.67
			11-12-20-30	.71,.54,.01,.23		1.91

*Classifier used for crop-acreage estimates.

Table C6. Summary of Soybean Classification trials for data set W123.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	63	0	.67	2.99	2.84
			10-50	.69,.49	2.56	2.60*
			11-12-20-30	.77,.44,.57,.56		2.52
	PER	67	0	.74	3.32	3.15
			10-50	.74,.50	2.78	2.82
			11-12-20-30	.78,.62,.55,.66		2.91
FLDS	EP	47	0	.62	2.61	2.48
			10-50	.64,.47	2.29	2.33
			11-12-20-30	.68,.50,.56,.55		2.31
	PER	66	0	.71	3.39	3.22
			10-50	.74,.52	2.84	2.89
			11-12-20-30	.78,.64,.56,.66		2.97

*Classifier used for crop-acreage estimates.

Table C7. Summary of Corn Classification trials for data set C1A.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	41	0	.71	3.30	1.71
			10-50	.70,.59	3.15	3.08
			11-12-30	.84,.77,.59		5.39
	PER	87	0	.71	3.30	1.14
			10-50	.37,.78	1.53	1.49
			11-12-30	.53,.61,.78		2.01
FLDS	EP	44	0	.77	4.24	2.19
			10-50	.75,.66	3.81	3.72
			11-12-30	.86,.79,.66		6.30 *
	PER	87	0	.59	2.34	1.21
			10-50	.41,.75	1.64	1.60
			11-12-30	.58,.60,.75		2.20

*Classifier used for crop-acreage estimates.

Table C8. Summary of Soybean Classification trials for data set C1A.

Train/ Test	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	61	0	.66	2.88	2.62
			10-50	.62,.96	2.59	2.39
			11-12-30	.61,.24,.96		2.38
	PER	68	0	.66	2.88	3.53
			10-50	.71,.96	3.46	3.19
			11-12-30	.72,.12,.96		3.11
FLDS	EP	62	0	.71	3.34	3.05
			10-50	.67,.98	3.03	2.79
			11-12-30	.66,.30,.98		2.76*
	PER	68	0	.77	4.20	3.83
			10-50	.74,.98	3.78	3.48
			11-12-30	.74,.15,.98		3.39

*Classifier used for crop-acreage estimates.

Table C9. Summary of Corn Classification trials for data set C12.

Train/ Test	Priors	Categories	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	26/MCPC & PC	51	11-12-20-30	.17,.42,.15,.00	1.09	
FLDS		10/MCPC & PC	64	11-(12,20)-30	.20,.20,.00	1.12	
NB	PER	6/SCPC & PC	89	(11,12,20)	.07	1.06	
			89	11-(12,20)-30	.33,.07,.00	1.20	
	EP		50	11-(12,20,30)	.02,.02	.98	
	PER	5/SCPC & PC	90	11-12-20	.29,.09,.01	1.16	
	EP	4/SCPC & PC	88	(11,12,20)	.05	1.04	
				11-(12,20)	.33,.05	1.27	*
				11-12-20	.33,.09,.02	1.21	

*Classifier used for crop-acreage estimates.

Table C10. Summary of Soybean Classification trials for data set C12.

Train/ Test	Priors	Categories	% Correct	Strata Pooling	R_h^2	RE1	RE2
NB	EP	26/MCPC & PC	76	11-12-20	.35,.61,.79	1.68	
FLDS		10/MCPC & PC	56	11-(12,20,30)	.25,.79,.56	1.59	
NB	PER	6/SCPC & PC	70	(11,12,20)	.44	1.77	
			70	11-(12,20)-30	.33,.82,.66	1.78	
	EP		67	(11,12,20)	.40	1.64	
			67	11-(12,20)	.29,.81	1.61	
			67	11-(12,20,30)	.29,.79	1.83	
	PER	5/SCPC & PC	72	11-12-20	.34,.84,.83	1.72	
	EP	4/SCPC & PC	76	11-12-20	.36, .79,.80	1.74	*

*Classifier used for crop-acreage estimates.

Table C11. Summary of Corn Classifiction trials for data set C33+, train/test on NB.

Priors	Categories	% Correct	Strata Pooling	R_n^2	RE1	RE2
EP	9/SCPC & PC	62	11-12-20-30	.26,.47,.38,.12	1.44	1.44
	10/SCPC & PC	48	0	.58	2.36	1.53
			10-50	.46,.52	1.86	1.74*
			11-12-20-30	.30,.52,.47,.22	1.60	1.60
	12/MCPC & PC	21	0	.47	1.87	1.21
			10-50	.39,.40	1.60	1.49
			11-12-20-30	.28,.61,.51,.01	1.67	1.67
	14/SCPC & PC	09	11-12-20-30	.01,.02,.34,.02	1.08	1.08
	16/MCPC & PC	07	11-12-20-30	.00,.05,.52,.11	1.17	1.17
	17/SCPC	08	11-12-20-30	.01,.02,.33,.02	1.07	1.07
	19/MCPC	07	11-12-20-30	.00,.06,.47,.11	1.15	1.15
PER	9/SCPC & PC	87	11-12-20-30	.49,.14,.00,.00	1.21	1.21
	10/SCPC & PC	86	11-12-20-30	.52,.15,.00,.00	1.22	1.22
	14/SCPC & PC	86	11-12-20-30	.52,.15,.00,.00	1.22	1.22
	16/SCPC & PC	58	0	.03	1.01	.66
			10-50	.33,.04	1.29	1.21
			11-12-20-30	.70,.03,.17,.02	1.42	1.42
	17/SCPC	87	11-12-20-30	.56,.11,.00,.00	1.24	1.24
	19/MCPC	58	11-12-20-30	.70,.06,.07,.02	1.36	1.36

*Classifer used for crop-acreage estimates.

Table C12. Summary of Soybean Classification trials for data set C33+, train/test on NB.

Priors	Categories	% Correct	Strata Pooling	R_h^2	RE1	RE2
EP	9/SCPC & PC	42	0	.37	1.56	1.16
			10-50	.27,.54	1.55	1.55
			11-12-20-30	.10,.75,.57,.50	1.97	1.97
	10/SCPC & PC	29	0	.48	1.90	1.41
			10-50	.40,.52	1.76	1.76
			11-12-20-30	.22,.79,.58,.71	2.23	2.23*
	12/MCPC & PC	70	11-12-20-30	.09,.70,.08,.23	1.38	1.38
	14/SCPC & PC	19	11-12-20-30	.13,.68,.55,.62	1.87	1.87
	16/MCPC & PC	38	11-12-20-30	.20,.63,.46,.81	1.79	1.79
	17/SCPC	19	11-12-20-30	.14,.67,.55,.73	1.89	1.89
	19/MCPC	38	11-12-20-30	.21,.62,.46,.83	1.80	1.80
PER	9/SCPC & PC	57	11-12-20-30	.14,.67,.52,.50	1.81	1.81
	10/SCPC & PC	48	0	.38	1.58	1.17
			10-50	.37,.50	1.68	1.68
			11-12-20-30	.19,.75,.55,.39	1.98	1.98
	14/SCPC & PC	48	0	.38	1.58	1.17
			10-50	.37,.50	1.68	1.68
			11-12-20-30	.19,.75,.55,.39	1.98	1.98
	16/MCPC & PC	80	11-12-20-30	.16,.78,.12,.17	1.50	1.50
	17/SCPC	63	11-12-20-30	.15,.67,.49,.55	1.79	1.79
	19/MCPC	80	11-12-20-30	.15,.78,.12,.16	1.49	1.49

*Classifier used for crop-acreage estimates.

Table C13. Summary of Corn and Soybean Classifications trials for data set E12, train/test on NB.

Cover	Priors	% Correct	Strata Pooling	R_h^2	RE1	RE2
Corn	EP	49	0	.35	1.50	.55
			(11,12,20)	.57	2.28	1.60
	PER	79	0	.34	1.48	.54
			(11,12,20)	.63	2.65	1.86*
Soybeans	EP	46	0	.44	1.75	.79
			(11,12,20)	.39	1.60	.97
	PER	64	0	.58	2.38	1.06*
			(11,12,20)	.42	1.68	1.01

*Classifer used for crop-acreage estimates.

Table Ci4. Summary of Corn Classification for data set E23+, train/test on NB.

Priors	Categories	% Correct	Strata Pooling	R_h^2	RE1	RE2
EP	18/MCPC	44	0	.38	1.59	.96
			10-50	.31,.25	1.36	1.15
			11-12-50	.19,.48,.25	1.38	1.27
			11-12-20-30	.19,.48,.37,.09		1.44 *
PER		68	0	.46	1.84	1.11
			10-50	.32,.44	1.51	1.27
			11-12-50	.27,.27,.44	1.42	1.31
			11-12-20-30	.27,.27,.82,.29		1.65
EP	28/MCPC	43	0	.53	2.08	1.26
			10-50	.43,.53	1.79	1.51
			11-12-50	.29,.52,.53	1.74	1.61
			11-12-20-30	.30,.52,.74,.23		1.92
PER		63	0	.50	1.97	1.19
			10-50	.37,.51	1.65	1.38
			11-12-50	.29,.35,.51	1.55	1.43
			11-12-20-30	.29,.35,.78,.19		1.72

*Classifier used for crop-acreage estimates.

Table C15. Summary of Soybean Classification trials for data set E23+, train/test on NB.

Priors	Categories	% Correct	Strata Pooling	R_h^2	RE1	RE2
EP	18/MCPC	44	0	.68	3.08	1.33
			10-50	.61,.75	2.78	1.68
			11-12-50	.62,.65,.75	2.94	2.21
			11-12-20-30	.62,.65,.44,.48		2.38*
PER		70	0	.66	2.86	1.23
			10-50	.60,.64	2.48	1.50
			11-12-50	.63,.60,.64	2.52	1.90
			11-12-20-30	.63,.60,.23,.11		2.11
EP	28/MCPC	27	0	.53	2.09	.90
			10-50	.44,.65	1.95	1.18
			11-12-50	.45,.36,.65	1.89	1.43
			11-12-20-30	.45,.36,.23,.77		1.56
PER		71	0	.64	2.73	1.18
			10-50	.57,.64	2.38	1.44
			11-12-50	.60,.58,.64	2.43	1.83
			11-12-20-30	.60,.58,.22,.18		2.02

*Classifier used for crop-acreage estimates.

Appendix D

Regression Estimates for Corn and Soybean Acreages in Individual Illinois Counties

Abbreviation	Meaning
CREGES (SBREGE)	Regression estimate for corn (soybeans) acreage [hundreds of acres]
CORNCV (SBCV)	Coefficient of variation of corn (soybeans) regression estimate [%]
CORNSS (SBSSO)	SSO estimate for corn (soybean) acreage [hundreds of acres]

COUNTY ESTIMATES FOR DOMAIN W123

COUNTY	DOMAIN	PLTSYM	CREGES	CORNCV	CORNSS	SBREGE	SBCV	SBSSO
ADAMS	1	1	1666	24.0	1300	836	35.3	1127
BROWN	1	2	537	33.4	355	243	50.7	311
BUREAU	1	3	2540	18.7	2604	1106	33.4	1166
CALHOUN	1	4	567	25.1	280	233	39.9	74
CARROLL	1	5	1265	17.5	1300	572	29.6	105
CASS	1	6	917	20.3	840	541	25.5	563
FULTON	1	7	1721	29.0	1510	914	37.8	866
GREENE	1	8	1368	19.2	1040	760	24.8	622
HANCOCK	1	9	1905	19.3	1630	748	36.2	1240
HENDERSON	1	A	1040	17.3	925	371	36.4	370
HENRY	1	B	2768	17.2	2450	794	46.6	723
JERSEY	1	C	857	21.6	530	489	27.0	381
JODAVIESS	1	D	1083	34.1	750	271	94.2	68
KNOX	1	E	1741	19.5	1740	796	31.6	763
MASON	1	F	1291	21.3	1125	761	27.9	924
MCDONOUGH	1	G	1625	17.4	1450	825	26.3	913
MERCER	1	H	1398	18.7	1552	439	43.4	412
MORGAN	1	I	1472	17.6	1110	937	20.9	730
OGLE	1	J	2230	19.0	2180	515	64.2	627
PEORIA	1	K	1240	24.0	1260	653	32.6	632
PIKE	1	L	1601	25.7	1380	783	37.3	701
ROCK ISLAND	1	M	1070	18.7	830	275	52.7	207
SCHUYLER	1	N	840	29.0	620	367	46.2	560
SCOTT	1	O	611	19.9	460	315	28.6	325
STARK	1	P	920	18.2	1000	406	32.1	425
STEPHENSON	1	Q	1721	18.6	1602	306	81.8	168
WARREN	1	R	1618	16.5	1720	641	32.2	650
WHITESIDE	1	S	2428	16.2	2250	624	49.0	650
WINNEBAGO	1	T	1215	21.5	1030	296	68.0	275

N=29

COUNTY ESTIMATES FOR DOMAIN C1A

COUNTY	DOMAIN	PLTSYM	CREGES	CORNCV	CORNSS	SBREGE	SBCV	SBSSO
BOONE	2	1	769	12.4	775	294	46.8	307
BUREAU	2	2	2320	12.0	2604	1326	30.2	1166
DEKALB	2	3	1828	12.8	2010	990	34.1	1112
LEE	2	4	2090	12.1	2100	1108	32.9	1143
MCHENRY	2	5	1396	10.9	1270	498	40.2	399
OGLE	2	6	2174	11.0	2180	682	50.2	627
WINNEBAGO	2	7	1230	10.6	1030	336	51.8	275

N=7

COUNTY ESTIMATES FOR DOMAIN C12

COUNTY	DOMAIN	PLTSYM	CREGES	CORNCV	CORNSS	SBREGE	SBCV	SBSSO
DEWITT	3	1	1050	22.2	1160	839	27.7	893
DUPAGE	3	2	560	68.4	195	763	23.2	189
GRUNDY	3	3	1165	21.7	1100	744	32.0	799
KANE	3	4	1384	23.8	1275	904	36.1	609
KENDALL	3	5	958	18.4	920	554	35.7	606
LASALLE	3	6	3184	19.9	3070	2148	31.7	2244
LIVINGSTON	3	7	3321	18.0	3000	2252	33.6	2593
LOGAN	3	8	1679	19.9	1930	1471	27.7	1343
MACON	3	9	1544	20.4	1580	1306	27.8	1324
MARSHALL	3	A	1057	22.5	1090	604	34.3	578
MASON	3	B	1324	24.0	1125	1101	27.4	924
MENARD	3	C	769	24.8	755	642	26.6	569
MCLEAN	3	D	3556	18.0	3680	2484	31.8	2498
MOULTRIE	3	E	855	21.7	985	777	26.4	608
PEORIA	3	F	1211	46.8	1260	843	34.2	632
PIATT	3	G	1335	18.5	1450	978	31.6	1050
PUTNAM	3	H	387	27.7	452	235	32.6	206
STARK	3	I	967	17.3	1000	566	37.6	425
TAZEWELL	3	J	1734	22.3	1770	1147	31.4	374
WOODFORD	3	K	1412	21.2	1600	914	31.1	340

N=20

COUNTY ESTIMATES FOR DOMAIN C33+

COUNTY	DOMAIN	PLTSYM	CREGES	CORNCV	CORNSS	SBREGE	SBCV	SBSSO
BOND	4	1	440	48.4	520	683	25.7	591
CLINTON	4	2	694	37.6	815	932	25.1	682
FAYETTE	4	3	868	43.2	835	1205	24.6	940
GREENE	4	4	872	30.1	1040	908	26.0	622
JACKSON	4	5	294	94.4	310	861	24.6	690
JERSEY	4	6	505	35.2	530	532	25.7	381
MACOUPIN	4	7	1409	31.0	1660	1358	26.6	1333
MADISON	4	8	789	47.4	900	1156	24.7	1130
MONROE	4	9	306	61.4	355	656	23.1	391
MONTGOMERY	4	A	1317	31.0	1420	1241	29.6	1314
MORGAN	4	B	1040	27.6	1110	1026	24.4	730
PERRY	4	C	233	92.6	295	831	19.2	470
RANDOLPH	4	D	373	74.3	540	920	27.7	607
SCOTT	4	E	403	32.4	460	403	27.5	325
ST. CLAIR	4	F	608	53.3	780	1390	17.4	1110
WASHINGTON	4	G	620	47.8	760	1300	18.8	1144

N=16

COUNTY ESTIMATES FOR DOMAIN E12

COUNTY	DOMAIN	PLTSYM	CREGES	CORNCV	CORNSS	SBREGE	SBCV	SBSSO
CHAMPAIGN	5	1	2561	22.4	2810	2324	33.2	2343
COOK	5	2	285	21.9	185	588	25.3	236
DOUGLAS	5	3	989	25.1	1390	990	32.7	860
DUPAGE	5	4	177	30.2	195	266	95.2	189
EDGAR	5	5	1358	27.0	1650	1278	37.0	1189
FORD	5	6	545	55.7	1370	325	16.2	1215
IROQUOIS	5	7	3440	19.5	3260	1807	47.6	2454
KANKAKEE	5	8	1658	21.2	1800	1075	48.5	1205
LAKE	5	9	419	26.6	250	430	81.1	195
PIATT	5	A	1007	25.8	1450	1194	28.6	1050
VERMILLION	5	B	1674	31.5	2110	1230	56.2	1740
WILL	5	C	1659	24.0	1450	1400	45.8	1154

N=12

COUNTY ESTIMATES FOR DOMAIN E23+

COUNTY	DOMAIN	PLTSYM	CREGES	CORNCV	CORNSS	SBREGE	SBCV	SBSSO
ALEXANDER	6	1	158	41.6	105	493	12.2	293
CLARK	6	2	623	36.8	755	970	16.9	933
CLAY	6	3	743	27.9	440	898	18.2	990
COLES	6	4	870	32.0	1300	897	24.4	1024
CRAWFORD	6	5	629	38.3	670	744	29.2	881
CUMBERLAND	6	6	408	38.7	530	808	15.0	695
EDWARDS	6	7	286	30.8	405	323	21.2	365
EFFINGHAM	6	8	658	31.7	790	908	16.8	860
FAYETTE	6	9	935	32.1	835	1349	17.3	1126
FRANKLIN	6	A	248	81.9	335	757	22.8	786
GALLATIN	6	B	394	35.4	640	664	17.3	622
HAMILTON	6	C	401	46.6	455	930	16.5	1006
HARDIN	6	D	30	89.5	65	38	4.1	8
JACKSON	6	E	420	43.2	275	937	14.6	693
JASPER	6	F	769	29.1	850	1038	16.3	1159
JEFFERSON	6	G	542	48.6	500	1017	21.0	774
JOHNSON	6	H	98	89.9	205	215	35.3	75
LAWRENCE	6	I	732	21.7	795	573	21.3	459
MARION	6	J	782	30.4	570	1053	17.3	986
MASSAC	6	K	201	43.1	255	427	15.5	180
MOULTRIE	6	L	805	28.0	985	571	29.8	697
PULASKI	6	M	201	43.0	100	481	13.6	233
PERRY	6	N	420	37.4	295	795	15.5	443
POPE	6	O	82	52.4	130	149	55.7	90
RICHLAND	6	P	628	29.9	615	818	17.9	667
SALINE	6	Q	355	35.7	520	529	19.7	375
SHELBY	6	R	1290	29.4	1700	1523	19.3	1360
UNION	6	S	208	56.8	140	453	19.9	198
WABASH	6	T	371	32.0	560	407	23.6	226
WAYNE	6	U	974	34.8	785	1451	19.0	1285
WHITE	6	V	713	37.9	950	1035	20.4	833
WILLIAMSON	6	W	176	72.3	115	328	29.6	128

N=32