Lockheed Electronics Company, Inc.
A SUBSIDIARY OF LOCKHEED CORPORATION
1830 NASA Road 1, Houston, Texas 77058
Tel. 713-333-5411

JSC-14466
NOV
79-10154
CR 151889

Ref: 642-6967
Contract NAS 9-15200
Job Order 73-743-17

TECHNICAL MEMORANDUM

LABEL IDENTIFICATION FROM STATISTICAL TABULATION (LIST) TEST AND EVALUATION

By

M. D. Pore and R. A. Abotteen

(E79-10154) LABEL IDENTIFICATION FROM STATISTICAL TABULATION (LIST) TEST AND EVALUATION (Lockheed Electronics Co.) 35 p HC A03/MF A01 CSCL 05B
N79-18416
Unclas
G3/43 00154

Approved By: J.C. Minter
T. C. Minter, Jr., Supervisor
Techniques Development Section

November 1978

LEC-12381

CONTENTS

TABLES

FIGURE

1. INTRODUCTION

Label Identification from Statistical Tabulation (LIST) is an analyst-interpreter (AI) picture-element (pixel) labeling procedure for making at-harvest percentage small-grain estimates in the Large Area Crop Inventory Experiment (LACIE).[1] In this labeling procedure, the AI is required to answer questions about the segment and pixels which relate to simple properties that discriminate small grains from nonsmall grains. The responses, along with pertinent agricultural and meteorological variables, are statistically weighted to develop a discriminant function which is trained on blind-site ground-truth labels.

Results from an earlier development of LIST were analyzed and reported by Pore in November 1977 (ref. 1). Those results were used to develop a semi-automated, operational LIST reported by Abotteen and Pore in February 1978 (ref. 2). This newly developed operational LIST was tested on both Kansas and North Dakota blind sites, and the results of those tests are reported here.

Section 2 describes the analyses performed, and section 3 gives the results of those analyses. An evaluation and recommendations follow in section 4.

---

[1] The LACIE is a joint undertaking of the U.S. Department of Agriculture (USDA), the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce, and the National Aeronautics and Space Administration (NASA). The procedures which are the subject of this paper were developed at the NASA Lyndon B. Johnson Space Center (JSC) for the Earth Observations Division (EOD), Space and Life Sciences Directorate.

2. ANALYSES

Four AI's were used to test the quality of the questions for discriminating small grains (agricultural crops) from nonsmall grains. Each AI analyzed 16 segments, taking approximately 2-1/2 hours per segment. Each set of 22 932 pixels in a given area is referred to as a segment and covers a rectangular area of approximately 9 by 11 kilometers (5 by 6 nautical miles).

Every pixel at the intersection of a 10-by-10 grid is a grid pixel (or grid dot). Two hundred nine grid dots are in each segment, and all (and only grid dots) were used in this study. An earlier investigation by Register and Hocutt (ref. 3) has indicated that interpixel correlations decrease with distance and that a distance of 10 pixel widths corresponds to negligible correlation. Hence, dot grids are assumed to be independent samples with respect to crop types.

Separate analyses were performed for the 1976 winter and spring wheat sites, there being eight of each. All Kansas blind sites with available ground truth in stratum 11 of the New LACIE Strata were chosen as the winter wheat test sites. The eight spring wheat sites were chosen from the blind sites in stratum 21 (figure 1 shows locations of New LACIE Strata). Since ground truth was required in stratum 21, segments were chosen to be representative of the three-state coverage of the stratum. The data within each stratum were further partitioned into four training and four test segments (table 1).

For each segment, four acquisition dates were chosen arbitrarily without respect to special areal agricultural-meteorological conditions such as cloud cover, etc.; these were chosen to cover generally the 1975-76 growing season for wheat. Table 2 gives these dates and the respective Robertson biostages for winter wheat and spring wheat. Three types of production film converter (PFC) products were generated: type 1, type 2, and the Kraus product (see reference 4, Austin, for a description of these films). The films were made into Research, Test, and Evaluation (RT&E) packets and kept separate from LACIE operational packets. This was done to maintain a

TABLE 1.— LIST DATA SET

Stratum	LACIE segment	Type (a)	Purpose	County	State
11	1019	WW	Training	Norton	Kans.
11	1035	WW	Training	Ford	Kans.
11	1855	WW	Training	Trego	Kans.
11	1865	WW	Training	Stevens	Kans.
11	1020	WW	Test	Rawlins	Kans.
11	1852	WW	Test	Lane	Kans.
11	1860	WW	Test	Hodgeman	Kans.
11	1880	WW	Test	Ellis	Kans.
21	1542	SW	Training	Roosevelt	Mont.
21	1650	SW	Training	Hettinger	N. Dak.
21	1651	SW	Training	Bowman	N. Dak.
21	1667	SW	Training	Harding	S. Dak.
21	1530	SW	Test	Phillips	Mont.
21	1656	SW	Test	Morton	N. Dak.
21	1660	SW	Test	Logan	N. Dak.
21	1668	SW	Test	Perkins	S. Dak.

[a]WW = winter wheat; SW = spring wheat.

TABLE 2.— LIST DATA ACQUISITIONS (1976)

LACIE segment	County	Date	Biostage	
			WW	SW
1019	Norton	Jan. 19	2.4	
		Feb. 6	2.5	
		June 12	4.6	
		June 30	5.4	
1020	Rawlins	Feb. 25	2.5	
		Apr. 10	2.7	
		June 3	3.7	
		July 18	6.0	
1035	Ford	Mar. 13	2.6	
		May 6	3.4	
		June 1	4.1	
		July 8	6.0	
1530	Phillips	June 1	3.5	3.1
		June 18	4.0	3.9
		July 7	5.5	5.0
		Aug. 12	7.0	6.0
1542	Roosevelt	Apr. 25	2.5	1.1
		June 18	4.3	3.4
		July 6	5.7	5.0
		July 24	6.0	6.0
1650	Hettinger	May 9	3.2	2.0
		May 27	3.8	3.0
		Aug. 7	6.0	6.0
		Aug. 25	6.0	6.0
1651	Bowman	May 10	3.3	2.2
		May 29	4.0	3.0
		July 21	6.0	6.0
		Aug. 8	6.0	6.0

TABLE 2.– Continued.

LACIE segment	County	Date	Biostage	
			WW	SW
1656	Morton	May 9	3.0	2.0
		July 2	6.0	4.4
		July 20	7.0	6.0
		Aug. 7	7.0	7.0
1660	Logan	May 7	3.1	2.0
		June 12	4.2	3.7
		Aug. 6	6.0	6.0
		Aug. 23	6.0	6.0
1667	Harding	May 10	3.4	2.3
		May 29	4.3	3.2
		July 21	6.0	5.9
		Aug. 8	6.0	6.0
1668	Perkins	Apr. 22	2.6	1.7
		May 9	3.3	2.3
		May 28	4.0	3.1
		Aug. 7	6.0	6.0
1852	Lane	Mar. 31	2.6	
		May 7	3.2	
		June 20	5.8	
		July 17	6.0	
1855	Trego	Mar. 13	2.6	
		Apr. 18	3.0	
		June 20	5.7	
		July 17	6.0	

TABLE 2.– Concluded.

LACIE segment	County	Date	Biostage	
			WW	SW
1860	Hodgeman	Mar. 13	2.5	
		May 6	3.3	
		June 2	4.1	
		July 8	6.0	
1865	Stevens	Feb. 7	2.4	
		May 15	3.6	
		June 20	5.8	
		July 8	6.0	
1880	Ellis	Mar. 13	2.6	
		May 6	3.2	
		June 10	4.9	
		July 16	6.0	

6

NASA-S-76-11610A

GREAT PLAINS CROP REPORTING DISTRICTS (CRD'S) AND NEW LACIE STRATA

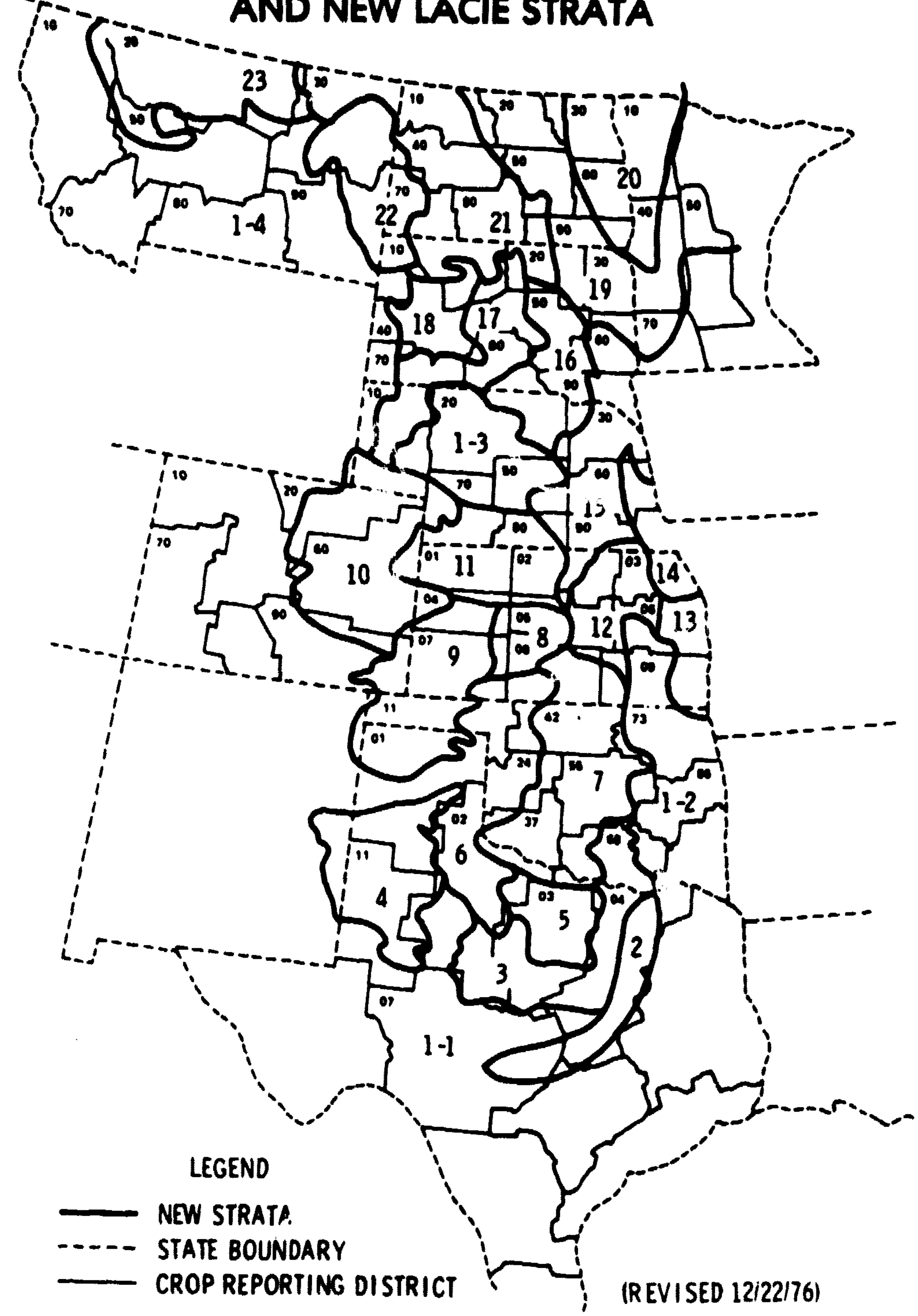

Figure 1.— Map of U.S. Great Plains showing New LACIE Strata.

restricted experimental environment of labeling without a 9- by 9-inch film image (covering a 185- by 185-kilometer track of land) of the broad area of interest and without ancillary agricultural-meteorological information. Hence, accuracies should be below those experienced in an operational labeling system.

The LIST procedure consists of obtaining AI responses to a set of questions directed at describing simple properties of the grid dots. The format used is presented in appendix A. These responses directly yield three categories or labels for the pixels.

a. Column 2 determines a designated other (DO) category.

b. Columns 3, 4, and 5 determine a nonclassifiable category.

c. The balance, those for which columns 6 through 9 are answered, constitutes a category of "pure" or labelable pixels.

The border pixels were omitted from the study, and their disposition will be discussed in section 4. The DO pixels were not part of the analysis or discriminating process but are reported as LIST results. This is because the LIST procedure (as presently defined) accepts the AI designation of DO as a LIST label. Only the pixels which could be labeled were admitted into analysis. This minimizes the effect of outliers and unlabelable pixels, thus producing more precise labeling functions.

The first analysis consisted of a stepwise linear discriminant analysis using the Statistical Package for the Social Sciences (SPSS, ref. 5). The major options were to base prior probabilities of category membership on training sample sizes and to use the minimum residuals method of stepping variables in and out of discrimination. Other analyses were a direct discriminant analysis that automatically uses every variable under consideration and a quadratic discriminant procedure that includes all linear terms and all two-way products (including squared terms). This latter procedure utilized the Patterson-Pitt algorithm as implemented by Thadani (ref. 6) and Ahlers (ref. 7). The discriminants were determined using ground truth on the four training segments, and accuracy was determined using the discriminant function to classify the four test segments. Percentages of pixels correctly labeled were calculated from contingency tables of ground truth by LIST.

An SPSS program listing for the spring wheat site LIST is given in appendix B as representative documentation of the automation process.

3. RESULTS

The particular variables admitted by a stepwise discriminant procedure are the number of training samples, the variability of the particular area sampled, acquisition dates, etc. Certainly, it is not recommended that a training sample of the size used here be implemented in LACIE; hence, discriminant vectors and tests for category mean differences will not be presented here. Instead, tables for test accuracy (on segments not used in training) are presented. Table 3 is a key to these contingency tables.

Four analyses were performed on the winter wheat segments: two using the quadratic discriminator (Q), one using the stepwise discriminant, and one using the AI labels. Table 4 gives these results for all four AI's, each responding to the four winter wheat test segments. Table 5 lists the variables used in the respective parts of table 4. Appendix C gives variable definitions for all analyses. As presently programmed, the quadratic discriminator was determined to accrue numerical analysis errors of computation at an unacceptable rate and was not used in the spring wheat site analyses.

All spring wheat sites were treated as mixed-wheat sites, even where winter wheat analysis was patently unnecessary. The mixed-wheat philosophy was to give positive responses automatically where indicated for either spring or winter wheat. For example, if the canopy trajectory for a pixel is similar to a winter wheat trajectory (SUM is high for winter wheat biostage numbers) while it is dissimilar for spring wheat (SUM is low for spring wheat biostage numbers), then KEYS and SUM are based on winter wheat biostages for that pixel. Tables 6 and 7 give the results for the spring wheat sites.

The AI percentage of small grains and the LIST percentage of small grains were consistently below the ground-truth percentage of small grains ($m < \ell$ in table 3), regardless of the type of discriminant used. This is partially attributed to the fact that (1) omission rates apparently are always less

TABLE 3.— CONTINGENCY TABLE KEY

Type of labeler

		SG	Non	ℓ
GT	SG	a	b + e	g
	Non	c	d + f	h
	m	i	j	k

PCL

Variable	Definition
a, b, c, d	Raw pixel counts for the four test segments
e, f	Raw pixel counts for the DO pixels
g, h, i, j	Marginal probabilities (expressed as percentages) of correct labeling (PCL's):
g	$\frac{a}{a + b + e} \times 100 = [1 - Pr(omission)] \times 100$
GT	Ground truth
h	$\frac{d + f}{c + d + f} \times 100 = [1 - Pr(commission)] \times 100$
i	$\frac{a}{a + c} \times 100$
j	$\frac{d + f}{b + d + e + f} \times 100$
k	$a + b + c + d + e + f$
ℓ	$\frac{a + b + e}{k} \times 100$ = ground-truth percentage of small grains
m	$\frac{a + c}{k} \times 100$ = LIST labeled percentage of small grains
Non	Nonsmall grains
PCL	$\frac{a + d + f}{k} \times 100$ = the probability (expressed as a percentage) of correct labeling
SG	Small grains

TABLE 4.— LIST TEST ACCURACY ON WINTER WHEAT SITES

AI labels

		SG	Non	21.1%
	SG	482	44 + 65	82%
GT	Non	73	586 + 1553	97%
	19.8%	87%	95%	2803

PCL = 93.5%

Q with B&G only

		SG	Non	21.1%
	SG	465	61 + 65	79%
GT	Non	81	578 + 1553	96%
	19.5%	85%	94%	2803

PCL = 92.6%

Linear discriminant

		SG	Non	21.1%
	SG	491	35 + 65	83%
GT	Non	86	573 + 1553	96%
	22.7%	85%	96%	2803

PCL = 93.4%

Q17

		SG	Non	21.1%
	SG	476	50 + 65	81%
GT	Non	85	574 + 1553	96%
	20.0%	85%	95%	2803

PCL = 92.9%

TABLE 5.– LIST TEST VARIABLES FOR WINTER WHEAT SITES

Title	Variables
AI labels	Analyst label.
Linear discriminant	G1, canopy trajectory, B4, GREEN3, B2, G4, KEY4, B1, G2, PCGW, G3, KEY3, GREEN2, KEY2, BIO4, GREEN4, BIO2.
Q with B&G only	B1, B2, B3, B4, G1, G2, G3, G4, and all possible interactions.
Q17	BIO2, BIO4, B1, B2, B4, G1, G2, G3, G4, GREEN2, GREEN3, GREEN4, PCGW, KEY2, KEY3, KEY4, canopy trajectory, and all possible interactions.

TABLE 6.– LIST TEST ACCURACY ON SPRING WHEAT SITES

AI label

		SG	Non	8.8%
GT	SG	113	32 + 79	50%
	Non	47	166 + 2106	98%
	6.3%	71%	95%	2543

PCL = 93.3%

Linear with B-G-BIO step

		SG	Non	8.8%
GT	SG	105	40 + 79	47%
	Non	47	166 + 2106	98%
	6.0%	69%	95%	2543

PCL = 93.5%

Linear discriminant

		SG	Non	8.8%
GT	SG	105	40 + 79	47%
	Non	67	146 + 2106	97%
	6.8%	61%	95%	2543

PCL = 92.7%

Linear with B-G-BIO direct

		SG	Non	8.8%
GT	SG	100	45 + 79	45%
	Non	41	172 + 2106	98%
	5.5%	71%	95%	2543

PCL = 93.5%

TABLE 7.– LIST TEST VARIABLES FOR SPRING WHEAT SITES

Title	Variables (in order of inclusion)
AI labels	AI label.
Linear discriminant	Canopy trajectory, G1, G3, B4, B1, GREEN1, G2, G4, GREEN4, PCGW, B3, KEY3, B2.
Linear with B-G-BIO step	Canopy trajectory, GS1, GW3, BS4, BW1, GREEN1, GW1, GS4, GS2, GREEN4, BS3, KEY3, PCGW, BW3, BS2.
Linear with B-G-BIO direct	GW1, GW2, GW3, GW4, GS1, GS2, GS3, GS4, BW1, BW2, BW3, BW4, BS1, BS2, BS3, BS4, PCGW, PCGS, canopy trajectory, KEY1, KEY2, KEY3, KEY4, GREEN1, GREEN2, GREEN3, GREEN4.

than commission rates ($b < c$ in table 3) and (2) a fairly consistent tendency exists for nearly 4 percent of the DO pixels to be small grains $\left(\frac{e}{e + f} \approx 0.038\right)$.

Mid-season estimation cannot be analyzed effectively because (1) acquisition date selection for end-of-season estimation is usually inappropriate for mid-season estimation and (2) specialized mid-season questions (e.g., automated prototype green number trajectories) have not been developed. Nevertheless, such an analysis is presented here, recognizing that lower than realistic accuracy is expected. Such an analysis indicates the efficacy of present keys and may be of heuristic value in pointing to new developments. A rather high accuracy (PCL in the terminology of table 3) and a moderate decrease in the percentage of small grains reported ($m < \ell$ in the terminology of table 3) are demonstrated in tables 8 and 9.

TABLE 8.— MID-SEASON TEST ACCURACY

Winter sites

		SG	Non	21.1%
	SG	409	117 + 65	69%
GT	Non	113	546 + 1553	95%
	18.6%	78%	92%	2803

PCL = 89.5%

Spring sites

	SG	Non	8.8%
SG	85	60 + 79	38%
Non	38	175 + 2106	98%
4.8%	69%	94%	2543

PCL = 93.0%

TABLE 9.— MID-SEASON TEST VARIABLES

Title	Variables
WW sites	BIO1, BIO2, G1, B1, G2, B2, KEY2, GREEN1, GREEN2, GW1, GW2, BW2.
SW sites	SBIO1, SBIO2, WBIO1, WBIO2, GW1, GW2, BW1, BW2, GS1, GS2, BS1, KEY1, KEY2, GREEN1, GREEN2.

4. EVALUATION AND RECOMMENDATIONS

The inclusion of all possible interactions, as is accomplished routinely using the Patterson-Pitt quadratic discriminator, does not appear to increase classification accuracy because of the inclusion of too many spurious variables. However, the selective construction of greenness/biostage and brightness/biostage interactions does appear to raise the PCL and could be incorporated beneficially in succeeding LIST developments.

The phenomenon of nearly 4 percent DO being small grains constitutes a source of bias that is apparently consistent over diverse geographic regions and is readily measurable. A study to measure this bias and develop a bias-correction procedure would be beneficial in the development of an operational LIST system.

The unexpectedly high PCL (high means close to AI labeling accuracy) in the "undeveloped discriminator" for mid-season labeling analyses suggests that directed development of a mid-season LIST labeler (as opposed to a causal byproduct of an end-of-season LIST labeler) would yield a highly accurate operational labeling system.

The present Classification and Mensuration System (CAMS) procedural philosophy is for the AI to select imagery for a "reference" acquisition date and mentally adjust the registration discrepancies of other acquisitions to give accurate labels to the "real estate" represented in the reference film. It is becoming increasingly evident that LIST, and in fact any labeling procedure that relies on spectral aids (e.g., trajectories), is inherently based on a different philosophy. Since acquisitions are usually not registered identically, spectral values for a pixel across several acquisitions therefore represent the area about the "real estate" and not a precise pixel of one date. Boundary pixels and mixed pixels (across a boundary) have spurious spectral trajectories; i.e., the trajectory is not sampled from any category of interest but switches from one category to another. Such trajectories tend to confuse the labeling process and reflect a basic modeling error in image interpretation. LIST, on the other hand, labels what is represented by the trajectory, which, in this case, is the grid dot intersection on the PFC product. To make this more meaningful, LIST first filters out the boundary (and mixed) pixels and then treats these pixels as a nonlabelable class to be proportioned. In summary, LIST does not label real estate but does label film grid intersection pixels. This philosophical change is implied by the increased reliance on spectral trajectories.

The high accuracies in tables 4 and 6 demonstrate that the concept of a statistical discrimination approach to pixel labeling is a valid concept and, in particular, that the LIST procedure (appendix A) performed comparably with AI methods. In the restrictive environment of these test conditions, this is a highly successful result that confirms the efficacy of this LIST questionnaire. However, it can be easily and obviously improved through the development and training of the automated keys, and particularly green number ranges and trajectories.

The recommendations made by Abotteen and Pore in February 1978 (ref. 2) are still applicable and can be expanded as follows:

a. A stratified estimation procedure for using LIST in LACIE Procedure 1 area estimation, where permissible labels are small-grains, other, and boundary pixels, should be developed.

b. A set of suitable questions for discrimination of wheat from other small grains could be profitably developed.

c. An early-season technology of LIST labels could be developed easily from the LIST developments represented here.

d. A multicrop (corn/soybean) LIST technology is certainly indicated from the small-grain/other successes reported here.

e. Adaptation of this LIST to an interactive color console computer system would advance pixel labeling technology to a cybernetic (feedback) process that could increase accuracy and possibly decrease operational processing time.

5. REFERENCES

1. Pore, M. D.: Label Identification from Statistical Tabulation (LIST). LEC-10825, rev. A, NASA/JSC (Houston), Nov. 1977.

2. Abotteen, R. A.; and Pore, M. D.: Development of Semiautomated, Operational Label Identification from Statistical Tabulation (LIST). LEC-11457, NASA/JSC (Houston), Feb. 1978.

3. Register, D. T.; and Hocutt, W. T.: Proportion Assessment Using Systematic Sampling (PASS) Feasibility Study. LEC-8049, NASA/JSC (Houston), Mar. 1976

4. Austin, W. W.: Detailed Description of the Wheat Acreage Estimation Procedure Used in the Large Area Crop Inventory Experiment. LEC-11497, NASA/JSC (Houston), Feb. 1978.

5. Nie, N. H.; Hull, C. H.; Jenkins, J. G.; Steinbrenner, K.; and Bent, D. H.: Statistical Package for the Social Sciences. Second ed., McGraw-Hill Book Co., Inc. (N.Y.), 1975.

6. Thadani, S. G.: A Nonparametric Loss-Optimal Pattern Classification System. LEC-11451, NASA/JSC (Houston), Jan. 1978.

7. Ahlers, C. W.: "As-Built" Design Specification for the Patterson-Pitt-Thadani Minimum Loss Classifier. LEC-12285, NASA/JSC (Houston), May 1978.

Acknowledgment

The following LACIE AI's assisted the authors in the work which is the subject of this paper:

- Harley Dupuis, Barbara Tolbert, and Diana Youngs performed questionnaire interpretations.

- In addition to interpreting images, Wes Palmer directed the AI interpretation process and recommended system modification to make this type of test effective.

APPENDIX A

LIST QUESTIONNAIRE

APPENDIX A

LIST QUESTIONNAIRE

Line 1 of Keypunch Transmittal Form

Column	Entry
1-5	Segment number.
6-30	County, state, or country if not United States.
31-33	Universal strata number.
34	Segment type: 1 — Winter wheat. 2 — Mixed wheat. 3 — Spring wheat.
36-40	Acquisition date chosen by analyst as registration date (YDDD). (This is not necessarily the Goddard Space Flight Center reference segment.)
42-46 48-52 54-58 60-64	Interpretable acquisition dates (YDDD).

Line 2 of Keypunch Transmittal Form

Column	Entry
1-5	Segment number.
7-8	Sun angles for the respective acquisitions.
10-11 13-14 16-17 19-21	Robertson winter wheat biostages.

Column	Entry
23-25, 27-29, 31-33	Robertson winter wheat biostages (continued).
35-37, 39-41, 43-45, 47-49	Robertson spring wheat biostages.

Succeeding lines of Keypunch Transmittal Form

Column	Entry
1	Leave the first column blank.
2	1 — Pixel is in nonagricultural area. STOP; pixel is DO. Go to 9. 0 or blank — Agricultural area or indeterminate.
3	Is pixel registered with regard to analyst chosen registration date (i.e., in the same category)? 1 — No. STOP; pixel is not classifiable. Go to 9. 0 or blank — Yes or indeterminate.
4	Is pixel a mixed pixel (part of more than one field or boundary)? 1 — Yes. STOP; pixel is not classifiable. Go to 9. 0 or blank — No or indeterminate.
5	Is this an anomalous pixel (not representative of most of the other pixels within the field)? 1 — Yes. STOP; pixel is not classifiable. Go to 9. 2 — No.

Column	Entry
6-9	PFC vegetation canopy indication is ____________________. (Use all available imagery film types.)
	0 — No vegetation canopy.
	1 — Low-density green vegetation canopy.
	2 — Medium-density green vegetation canopy.
	3 — High-density vegetation canopy.
	4 — Senescing (turning) vegetation canopy.
	5 — Harvested canopy (stubble).
11-14	Is the vegetation indication of the pixel on PFC imagery valid for the Robertson biostage of wheat for the acquisition? (Check keys for partition.)
	1 — No.
	2 — Yes.
15	Pixel is:
	1 — Small grains.
	2 — Other.
16-18	Line (or row) number of pixel.
20-22	Column number of pixel.

AUTOMATED LIST QUESTIONS FOR SMALL-GRAINS CLASSIFICATION

1. Green number of pixel is __________________. (Corrected to 60° latitude.)

2. Is the green number of the pixel within the range for small grains?

   Yes

   No

3. Brightness number of pixel is ___________________.

4. The winter principal component greenness (PCG) statistic is ____________.

5. The spring PCG statistic is ___________________.

6. Is the vegetation indication of the pixel valid for the Robertson biostage of wheat for the acquisition?

   Yes

   No

7. Does the pixel follow a small-grains spectral development pattern?

   Yes

   No

APPENDIX B

SPSS PROGRAM LISTING FOR SPRING WHEAT SITES

ORIGINAL PAGE IS
OF POOR QUALITY

APPENDIX B

SPSS PROGRAM LISTING FOR SPRING WHEAT SITES

```
RUN NAME        LIST TEST ON MIXED SEGMENTS
FILE NAME       LIST
PRINT BACK      CONTROL
VARIABLE LIST   SEGMENT,NAME,STRATA,TYPE,ACQ1,ACQ2,ACQ3,ACQ4,
                ANG1,ANG2,ANG3,ANG4,WBIO1,WBIO2,WBIO3,WBIO4,
                SBIO1,SBIO2,SBIO3,SBIO4,AI,TR,SEGNUM,
                AINUM,NONAG,REG,MIX,ANOM,CANOPYA,CANOPYB,
                CANOPYC,CANOPYD,VALIDA,VALIDB,VALIDC,VALIDD,
                AICLASS,ROW,COLM,TRUTH,G1,B1,G2,B2,G3,B3,G4,B4/
INPUT MEDIUM    DISK
INPUT FORMAT    FIXED(1X,F4.0,1X,A4,21X,F2.0,F1.0,7X,4(F5.0,1X)/5X,4(F2.0,1X),
                4(F3.1,1X),5X,4F2.0/1X,8F1.0,1X,5F1.0,2(F3.0,1X)/
                1X,A2,1X,4(F6.1,1X,F6.1,1X))
N OF CASES      UNKNOWN
MISSING VALUES  TRUTH('  ','K ')/
RECODE          ANOM (2=0) / TYPE (3=2)
COMMENT
COMMENT         SUN ANGLE CORRECTION
COMMENT
DO REPEAT       XG=G1,G2,G3,G4/XB=B1,B2,B3,B4/XANG=ANG1 TO ANG4/
COMPUTE         XG=(XG * .8660254)/SIN(XANG * .0174533)
COMPUTE         XB=(XB * .8660254)/SIN(XANG * .0174533)
END REPEAT
COMMENT
COMMENT         PCG STATISTICS FOR WINTER
COMMENT
COMPUTE         MUL1=0
COMPUTE         MUL2=0
COMPUTE         MUL3=0
COMPUTE         MUL4=0
DO REPEAT       XBIO= WBIO1 TO WBIO4/XMUL=MUL1 TO MUL4
IF              (XBIO LE 2.75) XMUL = (XBIO * .969)- 2.081
IF              ((XBIO GT 2.75)AND(XBIO LE 3.4)) XMUL= (XBIO * .275)- .172
IF              ((XBIO GT 3.4) AND(XBIO LE 4.15))
                XMUL = 2.479 - (XBIO * .505)
IF              ((XBIO GT 4.15) AND (XBIO LE 4.75))
                XMUL = 1.073 - (XBIO * .166)
IF              (XBIO GT 4.75) XMUL = 1.811 - (XBIO * .321)
END REPEAT
COMPUTE         PCGW = (G1 * MUL1)+(G2 * MUL2)+(G3 * MUL3)+(G4 * MUL4)
COMMENT
COMMENT         PCG STATISTICS FOR SPRING
COMMENT
DO REPEAT       XBIO = SBIO1 TO SBIO4/XMUL = MUL1 TO MUL4/
IF              (XBIO LE 3.5) XMUL = (XBIO * .36) - .65
IF              ((XBIO GT 3.5) AND (XBIO LE 4.5)) XMUL = (XBIO * .19)- .05
IF              ((XBIO GT 4.5)AND (XBIO LE 5.5)) XMUL = 1.73 - (XBIO * .21)
IF              (XBIO GT 5.5) XMUL = 2.96 - (XBIO * .43)
END REPEAT
COMPUTE         PCGS = (G1 * MUL1)+(G2 * MUL2)+(G3 * MUL3)+(G4 * MUL4)
COMMENT
COMMENT         KEYS FOR AI CANOPY ANSWERS
COMMENT
COMPUTE         WSUM=0
COMPUTE         WKEY1=0
COMPUTE         WKEY2=0
COMPUTE         WKEY3=0
COMPUTE         WKEY4=0
DO REPEAT       XBIO= WBIO1 TO WBIO4/XKEY= WKEY1 TO WKEY4/
                XCANOPY=CANOPYA TO CANOPYD/
IF              ((XBIO LE 2.0)AND (XCANOPY GE 1)AND(XCANOPY LE 3))XKEY=1
IF              ((XBIO LE 5.)AND (XCANOPY GE 4))XKEY=10
IF              ((XBIO LE 3.)AND (XBIO GT 2.5)AND (XCANOPY EQ 0))XKEY=1
IF              ((XBIO LE 6.)AND (XBIO GT 5.5)AND (XCANOPY LE 3)AND
                (XCANOPY GE 1)) XKEY=1
IF              ((XBIO GE 7.) AND (XCANOPY EQ 4))XKEY=1
IF              ((XBIO GT 3.) AND (XCANOPY EQ 0)) XKEY=10
IF              ((XBIO GT 6.) AND (XCANOPY LE 3)) XKEY=10
END REPEAT
COMPUTE         WSUM= WKEY1 + WKEY2 + WKEY3 + WKEY4
COMMENT
COMPUTE         SUM=0
COMPUTE         KEY1=0
COMPUTE         KEY2=0
COMPUTE         KEY3=0
COMPUTE         KEY4=0
DO REPEAT       XBIO= SBIO1 TO SBIO4/XKEY=KEY1 TO KEY4/
                XCANOPY=CANOPYA TO CANOPYD/
IF              ((XBIO LE 2.0)AND (XCANOPY GE 1)AND(XCANOPY LE 3))XKEY=1
```

```
IF                ((XBIO LE 5.)AND (XCANOPY GE 4))XKEY=10
IF                ((XBIO LE 3.)AND (XBIO GT 2.5)AND (XCANOPY EQ 0))XKEY=1
IF                ((XBIO LE 6.)AND (XBIO GT 5.5)AND (XCANOPY LE 3)AND
                  (XCANOPY GE 1)) XKEY=1
IF                ((XBIO GE 7.) AND (XCANOPY EQ 4))XKEY=1
IF                ((XBIO GT 3.) AND (XCANOPY EQ 0)) XKEY=10
IF                ((XBIO GT 6.) AND (XCANOPY LE 3)) XKEY=10
END REPEAT
COMPUTE           SUM= KEY1 + KEY2 + KEY3 + KEY4
DO REPEAT         XKEY=KEY1 TO KEY4/XW = WKEY1 TO WKEY4/
IF                (TYPE EQ 1) XKEY = XW
IF                ((TYPE EQ 2) AND (SUM LT WSUM)) XKEY = XW
RECODE            XKEY(0 THRU 9=0)
END REPEAT
RECODE            SUM (0 THRU 2=0)(3 THRU HIGHEST=1)
VALUE LABELS      SUM (1)NOT SM  GR (0)SM GR /KEY1 TO KEY4
                  (0)VALID AT BIOSTAGE(10)NOT VALID
COMMENT
COMMENT           GREEN NUMBER IN SMALL GRAIN RANGE
COMMENT
COMPUTE           GREEN1=0
COMPUTE           GREEN2=0
COMPUTE           GREEN3=0
COMPUTE           GREEN4=0
DO REPEAT         XBIO=WBIO1 TO WBIO4/XGREEN= GREEN1 TO GREEN4/XG=G1,G2,G3,G4/
IF                ((XBIO LE 2.8)AND(XG LE ((XBIO * 28.34) - 56.13)AND
                  LE (1.82 - (XBIO * .3)))) XGREEN = 1
IF                ((XBIO GT 2.8 AND LE 3.25)AND(XG GE ((XBIO * 5.64)
                  - 14.82)AND LE ((6.27 * XBIO) + 5.67))) XGREEN = 1
IF                ((XBIO GT 3.25 AND LE 3.75)AND(XG GE (8.65 - (XBIO * 1.58))
                  AND LE (2.65 + (XBIO * 7.38)))) XGREEN = 1
IF                ((XBIO GT 3.75 AND LE 4.25)AND(XG GE (.71 + (XBIO * .54))
                  AND LE (.71 + (XBIO * 7.74)))) XGREEN = 1
IF                ((XBIO GT 4.25 AND LE 5.0)AND(XG GE (3.62 - (XBIO * .15))
                  AND LE (75.02 - (XBIO * 9.75)))) XGREEN = 1
IF                ((XBIO GT 5.0)AND(XG GT (10.65 - (XBIO * 1.55)) AND LE
                  (61.25 - (XBIO * 6.99)))) XGREEN = 1
END REPEAT
COMPUTE           SGREEN1 = 0
COMPUTE           SGREEN2 = 0
COMPUTE           SGREEN3 = 0
COMPUTE            SGREEN4 = 0
DO REPEAT          XBIO = SBIO1 TO SBIO4/XGREEN = SGREEN1 TO SGREEN4/
                  XG = G1,G2,G3,G4/
IF                ((XBIO LE 2.25) AND (XGREEN GT ((XBIO * 3.39) - 3.38))
                  AND (XGREEN LE ((XBIO * 7.21) - 3.02))) XGREEN = 1
IF                ((XBIO GT 2.25) AND (XBIO LE 2.75) AND (XGREEN GT
                  ((XBIO * 1.29) + 5.35)) AND (XGREEN LE ((XBIO * 23.9)
                  - 40.57))) XGREEN = 1
IF                ((XBIO GT 2.75) AND (XBIO LE 3.25) AND (XGREEN GT
                  ((XBIO * 4.54) - 3.86)) AND (XGREEN LE ((XBIO * 2.64) +
                  17.9))) XGREEN = 1
IF                ((XBIO GT 3.25) AND (XBIO LE 3.75) AND (XGREEN GT
                  ((XBIO * 4.6) - 3.73)) AND (XGREEN LE ((XBIO * 2.6) +
                  18.03))) XGREEN = 1
IF                ((XBIO GT 3.75) AND (XBIO LE 4.5) AND (XGREEN GT
                  ((XBIO * .97) + 9.87)) AND (XGREEN LE (32.23 -
                  (XBIO * 1.19)))) XGREEN = 1
IF                ((XBIO GT 4.5) AND (XBIO LE 5.5) AND (XGREEN GT
                  (45.84 - (XBIO * 7.03))) AND (XGREEN LE (76.71 -
                  (XBIO * 11.07)))) XGREEN = 1
IF                ((XBIO GT 5.5) AND (XGREEN GT (21.63 - (XBIO * 2.62)))
                  AND (XGREEN LE (37.38 - (XBIO * 3.42)))) XGREEN = 1
END REPEAT
COMPUTE           GEX = GREEN1 + GREEN2 + GREEN3 + GREEN4
COMPUTE           SGE = SGREEN1 + SGREEN2 + SGREEN3 + SGREEN4
DO REPEAT         XW = GREEN1 TO GREEN4/ XS = SGREEN1 TO SGREEN4/
IF                (SGE GT GEX) XW = XS
END REPEAT
COMPUTE           GW1 = G1 * WBIO1
COMPUTE           GW2 = G2 * WBIO2
COMPUTE           GW3 = G3 * WBIO3
COMPUTE           GW4 = G4 * WBIO4
COMPUTE           GS1 = G1 * SBIO1
COMPUTE           GS2 = G2 * SBIO2
COMPUTE           GS3 = G3 * SBIO3
COMPUTE           GS4 = G4 * SBIO4
COMPUTE           BW1 = B1 * WBIO1
COMPUTE           BW2 = B2 * WBIO2
```

```
COMPUTE          HW3 = H3 * WRI03
COMPUTE          HW4 = H4 * WRI04
COMPUTE          HS1 = H1 * SRI01
COMPUTE          HS2 = H2 * SRI02
COMPUTE          HS3 = H3 * SRI03
COMPUTE          HS4 = H4 * SRI04
VALUE LABELS     GREEN1 TO GREEN4(0)NOT SM GR (1)SM GR
PRINT FORMAT     NAME(A)
RECODE           TRUTH(' O','OO','RO','OR',' R'='O')(' W','WW',
                 ' A','WA',' F','AW'= 'W')
                 ('R ','OA','OR','RA'='OA')('HO','AO','WR'='WO')
                 (' B','BB'='B')
RECODE            TRUTH('W'=1)('B'=0)
SELECT IF         ((RES EQ 0 AND MIX EQ 0 AND ANOM EQ 0 AND
                 NOMAS EQ 0) AND ((TYPE EQ 2 AND TR EQ 1 AND (TRUTH EQ
                 0 OR 1)) OR (TYPE EQ 2 AND TR EQ 2 AND TRUTH EQ 0)))
IF               ((TR EQ 2) TRUTH = -1
MISSING VALUES   TRUTH(-1)
VALUE LABELS     TRUTH (0) BARLEY (1) WHEAT /
DISCRIMINANT       GROUPS = TRUTH (0,1) /
                 VARIABLES = CANOPYA TO CANOPYD,SUM TO KEY4,
                 GS1 TO GS4,HS1 TO HS4/
                 METHOD = MINRESID/PRIORS = SIZE/
                 ANALYSIS = CANOPYA TO CANOPYD,SUM TO KEY4,
                 GS1 TO GS4,HS1 TO HS4/
                 PRIORS = SIZE/
OPTIONS          2,5,6,11,12
STATISTICS       ALL
READ INPUT DATA
FINISH
```

APPENDIX C

VARIABLE DEFINITIONS FOR ANALYSES

APPENDIX C

VARIABLE DEFINITIONS FOR ANALYSES[a]

Variable	Definition
B-G-BIO	Brightness, greenness, and biostage interaction
BIO1, BIO2, BIO3, BIO4 or WBIO1 through WBIO4	Winter wheat Robertson biostages for the respective acquisitions.
SBIO1 through SBIO4	Spring wheat biostages.
G1, G2, G3, G4	Green numbers.
B1, B2, B3, B4	Brightness numbers.
GREEN1 through GREEN4	Yes/No answer: Is green number in the small-grain range?
KEY1 through KEY4	Yes/No answer: Is canopy in the small-grain range?
Canopy trajectory	Yes/No answer: Is canopy trajectory acceptable for small grains?
PCGW, PCGS	Principal component greenness statistic for winter and spring wheat, respectively.
GW1 through GW4	Products of $G_i \times WBIO_i$ for $i = 1,2,3,4$.
GS1 through GS4	Products of $G_i \times SBIO_i$ for $i = 1,2,3,4$.
BW1 through BW4	Products of $B_i \times WBIO_i$ for $i = 1,2,3,4$.
BS1 through BS4	Products of $B_i \times SBIO_i$ for $i = 1,2,3,4$.

[a]See Abotteen and Pore (ref. 2) for the numerical derivations.