80-10060

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

JSC-16274
DEC 7 1979
NASA CR-160425

Ref: 642-7841
Contract NAS 9-15800
Job Order 73-302-01

(E80-10060) IMPLEMENTATION OF BADHWAR CLASSIFICATION OF CORN/SOYBEAN SEGMENTS (Lockheed Electronics Co.) 32 p HC A03/MF A01 CSCL 02C N80-18511 Unclas G3/43 00060

TECHNICAL MEMORANDUM

IMPLEMENTATION OF BADHWAR CLASSIFICATION OF CORN/SOYBEAN SEGMENTS

By

Willa W. Austin

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

October 1979

LEC-14064

DISTRIBUTION

Distribution of this document is limited to those people whose names are followed by an asterisk in the following list; all others receive an abstract (JSC Form 1424) only.†

JSC/G. Badhwar/SF3*
K. Baker/SF3
R. R. Baldwin/SF3
T. L. Barnett/SF3
R. M. Bizzell/SF4*
I. D. Browne/SF3
L. F. Childs/SF2
K. J. Demel/SF3
H. G. deVezin/FM8
J. W. Dietrich/SF3
J. L. Dragg/SF4*
R. B. Erb/SF1
J. D. Erickson/SF3*
J. G. Garcia/SF3
G. E. Graybeal/SF5*
F. G. Hall/SF1*
C. R. Hallum/SF4
K. E. Henderson/SF3
W. E. Hensley/SF2
R. P. Heydorn/SF3*
R. O. Hill/SF4*
A. G. Houston/SF4*
R. D. Juday/SF3
T. W. Pendleton/SF3*
D. E. Pitts/SF3*
D. R. Thompson/SF3*
M. C. Trichel/SF3*
V. S. Whitehead/SF3

USDA/G. O. Boatwright
A. D. Frank/SF6
R. E. Hatch/SA4
J. D. Murphy*
R. L. Packard/SA4*

NOAA/D. G. McCrary/SF4*

LEC/J. G. Baron*
M. L. Bertrand
B. L. Carroll*
P. L. Krumm
P. C. Swanzy
J. J. Vaccaro
Technical Library (5)*
Job Order File*

ERIM/Q. A. Holmes*
R. Horvath*
D. Rice

KSU/A. M. Feyerherm
E. T. Kanemasu

LARS/M. E. Bauer*
D. A. Landgrebe*
T. L. Phillips
P. H. Swain

TAMU/L. F. Guseman
J. C. Harlan
H. O. Hartley

UCB/R. N. Colwell
C. M. Hay*
R. W. Thomas

UH/H. P. Decell

ESCS/W. H. Wigton

---

†To obtain a copy of this document, contact one of the following:

J. D. Erickson — NASA/JSC Supporting Research Branch (SF3)

J. E. Wainwright — LEC/SSD Development and Evaluation Department (626-43, C09)

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
JSC-16274		

4. Title and Subtitle
Implementation of Badhwar Classification of Corn/Soybean Segments

5. Report Date
October 1979

6. Performing Organization Code

7. Author(s)
Willa W. Austin, Lockheed Electronics Co., Inc.

8. Performing Organization Report No.
LEC-14064

9. Performing Organization Name and Address
Lockheed Electronics Company, Inc.
Systems and Services Division
1830 NASA Road 1
Houston, Texas 77058

10. Work Unit No.
63-2137-3302

11. Contract or Grant No.
NAS 9-15800

12. Sponsoring Agency Name and Address
National Aeronautics and Space Administration
Lyndon B. Johnson Space Center
Houston, Texas 77058

13. Type of Report and Period Covered

14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This paper documents the results of applying the Badhwar classification method to a large data set of corn/soybean sites for crop year 1978. The software programs and the procedures used to generate full scene classifications are presented, and numerical results of the acreage estimations are given.

17. Key Words (Suggested by Author(s))
Badhwar classification, crop profile discrimination, spectral separability, planting date standardization, two-class classification, automatic crop inventory methods, SCREEN

18. Distribution Statement

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	34	

CONTENTS

TABLES

Table | Page

FIGURES

Figure | Page

1. INTRODUCTION

Dr. Gautam Badhwar of the National Aeronautics and Space Administration (NASA), Lyndon B. Johnson Space Center (JSC) has developed an approach to classification based on the hypothesis that a given crop, after emergence, has a unique spectral profile in time in the spectral region covered by the Land Satellite (Landsat) channels. A crop exhibits emergence at different times over a site. The result is displacement of the crop profile, but the profile retains the same shape (ref. 1). This classification method incorporates the effects of emergence data distribution and bases classification on the temporal profile of the crop (ref. 2). Local training, one input training field, is used to establish the crop profile for the site.[1]

The classification maps of 26 corn/soybean sites were produced using this method for evaluation as an aid to quality assurance for the Accuracy Assessment (AA) digitized ground-truth inventory maps (ref. 3). Since this method is in the development stage, it was expected that the generation of results for so large a data set would (1) test the applicability of the program to varying acquisition distributions and training field signatures and (2) define a tentative method for producing an acceptable classification as well as an initial set of criteria for evaluating the acceptability of the classification maps.

In this report, the software and the procedures used to apply Badhwar profile comparison classification are documented, and numerical results are presented. Section 2 defines the data set: site selection, acquisition choice, and editing of picture elements (pixels) using "SCREEN." Section 3 describes the available software programs. Section 4 discusses two aspects of the analysis: the analysis necessary to define the input to the program and the analysis used to evaluate the classification output. Section 5 presents numerical results in tabular form. Conclusions and recommendations, including a recommended data set for further program development, are given in section 6.

---

[1]Mr. Gary L. Gutchewski of NASA/JSC selected the training fields for this study.

2. DATA SET

Twenty-six corn/soybean segments were classified by using acquisitions for which the distribution over the corn-growing season was adequate to define a curve (crop profile) in each Landsat channel. The full segment (22 932 pixels), less those pixels removed by the SCREEN program developed by the Environmental Research Institute of Michigan (ERIM), was classified.

2.1 SITE SELECTION

Twenty-nine segments in the U.S. Corn Belt were used for the Multicrop Exploratory Experiment Test (ref. 4). These sites were located in agrophysical unit (APU) 14, 24, 25, or 28, had reasonable acquisition histories, and had available the aircraft photography ground truth. Final site selection for the multicrop experiment was determined by the procedural constraints on the acquisition requirements; of the 29 segments, procedure 1 (P1) estimates were made on 25 segments. Segments excluded from the test were sample segments 216, 878, 891, and 893 [sample segments 202 and 807 were not used in the followup Simulated Aggregation Test (ref. 5) because of poor data quality in crucial acquisitions].

For use of Badhwar classification, acquisition requirements are different from those for analyst labeling and for P1 classification. Hence, the entire set of 29 segments proposed for the multicrop test was taken as the basic site data set:

Sample segment	County, state	APU	Sample segment	County, state	APU
107	Boone, Ill.	25	202	Atchison, Mo.	14
123	Hamilton, Ind.	28	205	Clark, Mo.	25
127	Montgomery, Ind.	28	216	Mercer, Mo.	25
133	Whitley, Ind.	28	800	Clinton, Iowa	25
135	Chickasaw, Iowa	24	807	Henry, Ill.	25
141	Madison, Iowa	25	809	Ogle, Ill.	25
144	Wapello, Iowa	25	832	Adams, Ind.	28

Sample segment	County, state	APU	Sample segment	County, state	APU
837	Benton, Ind.	28	877	Ida, Iowa	14
842	Henry, Ind.	28	878	Kossuth, Iowa	24
843	Henry, Ind.	28	880	Monona, Iowa	14
852	Randolph, Ind.	28	881	Monona, Iowa	14
853	Randolph, Inc.	28	882	Palo Alto, Iowa	24
860	Wells, Ind.	28	891	Shelby, Iowa	14
864	Crawford, Iowa	14	893	Webster, Iowa	24
865	Crawford, Iowa	14			

2.2 ACQUISITION SELECTION

The Badhwar classification method reported in this document requires five acquisitions in the postemergence to preharvest growth stages of corn. Classification is successful, however, if four acquisitions are available in this period. Site selection depends upon the availability of a suitable set of acquisitions.

- Acquisition distribution *must* be adequate to define the crop profile.

  > For sample segment 107, available acquisitions sampled only the ripe to senescent crop stages (Julian days 208 to 305); hence, a proper crop profile could not be defined, and the segment could not be classified.

- Acquisitions on cloudy and hazy days should be avoided. The Badhwar program is not, however, overly sensitive to atmospheric conditions; cloud shadows and haze apparent on the film products of the Large Area Crop Inventory Experiment (LACIE) may not affect profiles or classifications.

  > Acquisitions available for sample segment 807 are these: 78137 (popcorn clouds), 78164 (haze and data drop), 78209 (haze obscured), 78218 (one-third haze obscured), 78272 (harvest beginning on corn), 78181, and 78290. This segment could not be classified because of poor data quality in the preharvest acquisitions.

- Acquisitions should occur in the growing season of the crop, although considerable robustness was exhibited by the present technique in the definitions of postemergence and preharvest. As with atmospheric condition effects, final judgment of the usable acquisitions was made with reference to graphs of the proposed training fields over the available acquisitions.

> The acquisitions available for sample segment 893 (78131, 78221, 78266, and 78293) clearly did not satisfy the constraint of a minimum of four acquisitions in the postemergence to preharvest period. This segment could not be classified.

2.3 USE OF SCREEN

As each image was unloaded from an Earth Resources Interactive Processing System (ERIPS) image unload tape onto a disk for processing on the programmed data processor (PDP 11/45), it was edited using the ERIM program SCREEN (ref. 6) — a "procedure for automatically detecting garbled data, clouds, snow, cloud shadows, and water in MSS data."[2] Pixels failing to pass this edit step were excluded from processing. Since the number of pixels removed by SCREEN varied with segment and acquisition, percentages were normalized to 22 932 pixels for all segments for the presentation of results in section 5.

---

[2]A listing that tallied the number of pixels edited from each acquisition was available. This report affected acquisition choice in that overly screened acquisitions were avoided, if possible. If it was not possible to avoid using an acquisition with an excessive number of pixels removed, a program was available which could be used to omit an acquisition from editing (refs. 7 and 8). This was done on all acquisitions for sample segment 135 (agricultural land was removed by SCREEN as being cloud shadows) and on acquisition 78165 for sample segment 144 (10 858 pixels were removed by SCREEN as clouds although the acquisition is cloud free).

3. SOFTWARE AVAILABLE

All data processing used to generate the classifications was done on the PDP-11/45 image processor. Several software programs were used: IMUNLD, IMAPLT (to help define the analyst input training field and acquisition set), CLASFY, A2SGMAP, and MISMAP (to aid in analyst evaluation of results). This section describes each of these programs.

3.1 IMUNLD

IMUNLD takes an image unload tape generated on the Earth Resources Interactive Processing System (ERIPS), edits it using SCREEN, adjusts the Landsat-3 acquisitions into a data range comparable to the data range of Landsat-2 acquisitions using the Wehmanen multiplicative factors (ref. 9), and loads the images into a PDP-11/45 disk.

- Input: ERIPS image unload tape.
- Output: screened, Landsat-3 adjusted images on a PDP-11/45 disk.

3.2 IMAPLT

IMAPLT (ref. 10)[3] plots the individual pixels of a field, giving reflectance values versus time (i.e., the acquisition dates specified) for each channel. IMAPLT then plots the field mean values, each channel, with a one standard deviation envelope; a curve is fitted through the mean values. Eight graphs are produced for a field over a set of acquisitions. Graphs are displayed on the Image-100 Tektronix screen, and hardcopies are made automatically. The segment number, acquisitions used, coordinates of the field, channel number, the number of pixels in the field, and the mean and standard deviation on each acquisition are listed on the first plot. The constant values computed from the data for the model (with the estimated error), the estimated field planting date (with error), the values of the fitted curves at the specified acquisitions (which can be compared with the computed mean values of the data),

[3]Available reference is to TRJPLT, an early version of IMAPLT.

and the chi-square value for the fit of the approximating curve to the field data are presented on the second plot.

- Input: field coordinates (line, pixel) in order; acquisition set of four or of five acquisitions.
- Output: eight graphs as above, two for each Landsat channel.

3.3 CLASFY

CLASFY (ref. 11) computes the constants for the curves from the training field data, compares (with this crop profile in each channel) the values for each pixel in the segment, and rejects those pixels which are not within a specified chi-square measure of the profile. The technique for rejection is to compare pixel channel values with the profiles in channel 2, channel 3, channel 4, then channel 1 in succession and reject if the comparison in any single channel is inadequate. Variability of the time of planting/emergence is allowed for in the comparison of individual pixels with the crop profile (refs. 1 and 2). Accepted pixels are labeled "corn"; rejected pixels, non-corn.

- Input: five, or four, image files; coordinates of one crop-of-interest field to establish crop profiles; initial values for the function constants as computed in IMAPLT (to aid convergence of the approximating curve).
- Output: classification file on disk which has a designation of "screened," "corn," or "non-corn" for each pixel in the segment; lineprinter sheet summarizing the following:

  a. Acquisitions used.

  b. Training field coordinates and the number of pixels in the field.

  c. Mean and standard deviation for each channel and each acquisition (field averages).

  d. The input and the final constants (with error) for the model.

  e. Final chi-square values for each channel (training field data).

  f. Estimated planting date of the training field (with error) as derived for each channel.

g. The chi-square thresholds in each channel applied as cutoff values in classification.

h. The number of pixels cut for exceeding the chi-square threshold, hence removed from consideration as corn, in each channel.

i. The final numerical results: the number of pixels classified as corn, the number of pixels screened, and the number of pixels rejected as corn.

3.4 A2SGMAP

A2SGMAP provides a full classification map (22 932 pixels) of the results of CLASFY. The scale is the same as that used for the AA digitized ground-truth maps. Pixels classified as corn are designated "C," those screened are "T," and those rejected as corn are left as blank spaces on the map.

- Input: classification file from CLASFY.
- Output: lineprinter map, full classification.

3.5 MISMAP

MISMAP (ref. 11) compares the classification map presented by A2SGMAP with the AA digitized ground-truth inventory map for the segment. A lineprinter map with this code is generated:

- Pixels screened appear as T.
- Ground-truth corn classified as corn appears as C.
- Ground-truth non-corn rejected as corn is left blank.
- Ground-truth non-corn classified as corn appears as +.
- Ground-truth corn rejected as corn appears as -.
- Pixels for which ground truth is not available but which are classified as corn appear as $.
- Pixels for which ground truth is not available but which are rejected as corn appear as %.

A numerical scene summary is given in confusion matrix form.

MISMAP maps can be generated for all pixels or for pure (AA definition) pixels only. Pure pixels (AA) are those which on a subpixel level contain only one crop.

- Input: classification file from CLASFY and ground-truth inventory map file.
- Output: full scene lineprinter map comparing the classification map with the ground-truth map and a confusion matrix numerical summary of results.

4. DISCUSSION OF ANALYSIS

Use of Badhwar classification involves different analyst requirements from those necessary to implement a P1 classification (ref. 12). Four differences are as follows:

- Badhwar classification is trained on one class — the crop of interest (in this case, corn) as opposed to P1, in which training is done on both the crop and the non-crop classes.
- Input for Badhwar classification is one analyst-selected field as opposed to P1 analyst-labeled dots from a semipredetermined set of pixels that samples the major crop categories in the scene.
- In Badhwar classification, four or five acquisitions are needed to determine the crop profile; these, preferably, are well distributed in the post-emergence to preharvest period. In P1, a set of acquisitions, a maximum of four, is used; it includes, preferably, a pre-emergence and a postharvest acquisition.
- Output from Badhwar classification is a full classification map that must be evaluated as opposed to that of P1, the monitoring of a pixel-based bias correction applied to a statistical estimate.

Discussion of the analyst input and analysis will be divided into two parts: 4.1 — training field and acquisition selection and 4.2 — evaluation of classification results.

4.1 ANALYST INPUT TO CLASSIFICATION: TRAINING FIELD SELECTION AND CHOICE OF ACQUISITIONS

Initial assessment of the acquisition coverage for each segment was done by the analyst using the LACIE film products. For each segment, acquisitions available in the corn growth period were listed and comments presented on data quality and crop growth stage for each acquisition. Based on the list and the comments, a recommendation of "possible" or "reject" was given for the segment.

Criteria for "possible" acquisition coverage of a segment were as follows:

- A minimum of four acquisitions, reasonably haze-free and cloud-free, in the postemergence to preharvest growth stages of corn was necessary.
- Acquisitions were within LACIE acceptable registration error and were of reasonably good data quality.

For those segments with "possible" acquisition coverage, training fields of corn were defined *without reference to ground truth.* Since fields of corn vary in planting date and development, it was expected that the acquisitions available would characterize the crop profile with varying results between training fields. Approximately four candidate training fields were selected for each segment, and final field choice was determined by the profiles generated.

Criteria for training field selection from imagery are these:

- Training field size of 20 to 40 pixels.
- Exclusion of border and edge pixels.
- Avoidance of roads, drainage patterns, etc., in fields, if possible.
- No selection of fields with unusual signature (e.g., only irrigated field in dryland area).
- Training field cloud-free and haze-free on all acquisitions used, not harvested on final acquisition, and (preferably) some signs of emergence on first acquisition.

The segments with fields defined by the analyst are those in which the program could be expected to be successful. These segments, indicated by an asterisk(*) in table 1, are the following sample segments:

107	837	864
127	842	865
133	843	877
141	852	881
216	853	882
809	860	

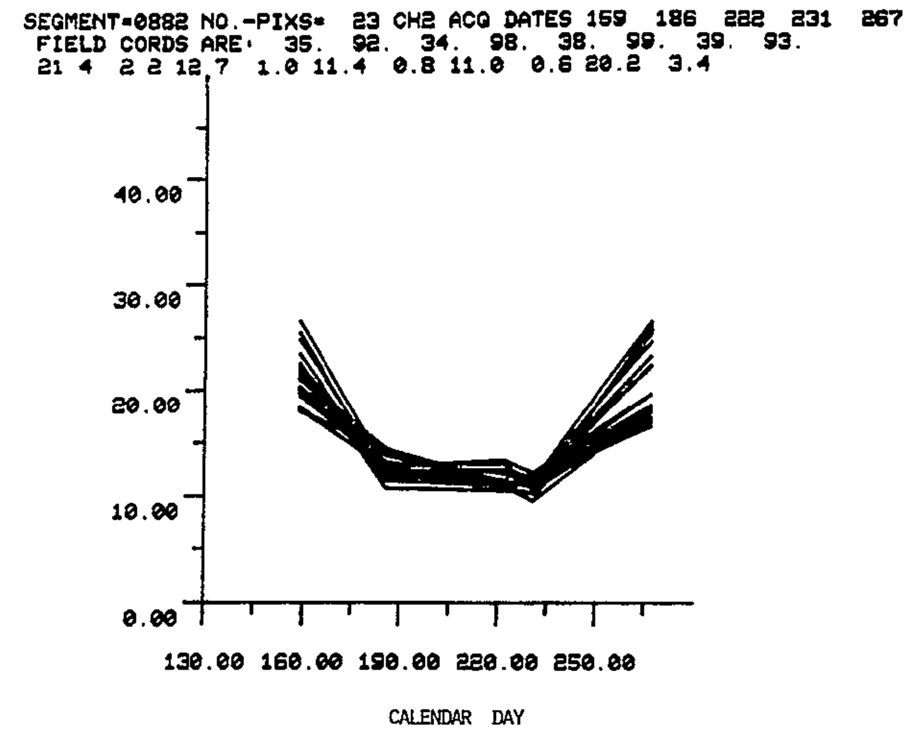
Processing was attempted for the remaining segments to "test the limits" of the program; the author defined two training fields, using ground truth, for these segments. Results for the following sample segments should be considered exploratory:

123	807
135	832
144	878
202	880
205	891
800	893

All training fields were graphed using IMAPLT. Five acquisitions were graphed if available. If a five-acquisition set was not available or if all training fields failed criteria using five acquisitions, four acquisitions were graphed.

Acquisition selection criteria are listed below:

a. Distribution of acquisitions relative to the training field crop development must be adequate to characterize the crop profile. As mentioned previously, it was expected that the training fields would differ in planting date and development. Hence, over a given set of acquisitions, the crop profile definition would differ between training fields. Figure 1 illustrates channel 2, training field 1, sample segment 882 (acquisitions 157, 186, 222, 231, 267); figure 2 shows training field 2 (same segment and same acquisitions). In development, training field 2 appears to be slightly later; for these acquisitions, field 1 is approximated better by the curve. Both of these fields are acceptable curves. Figure 3 illustrates channel 2 "curve" definition for sample segment 107 over available acquisitions (208, 226, 235, 244, 262); this definition is unacceptable as the field is "early" relative to all available acquisitions. Figures 4 and 5 illustrate an evident difference in crop development. Figure 5 shows a field which is later than that in figure 4 in exhibiting the characteristic dip associated with chlorophyll absorption. Both figures are for sample segment 860, channel 2 profile (acquisitions 160, 197, 232, 251, 268). Using these acquisitions, the curve generated for training field 2 (fig. 5) is not an adequate approximation of the data; training



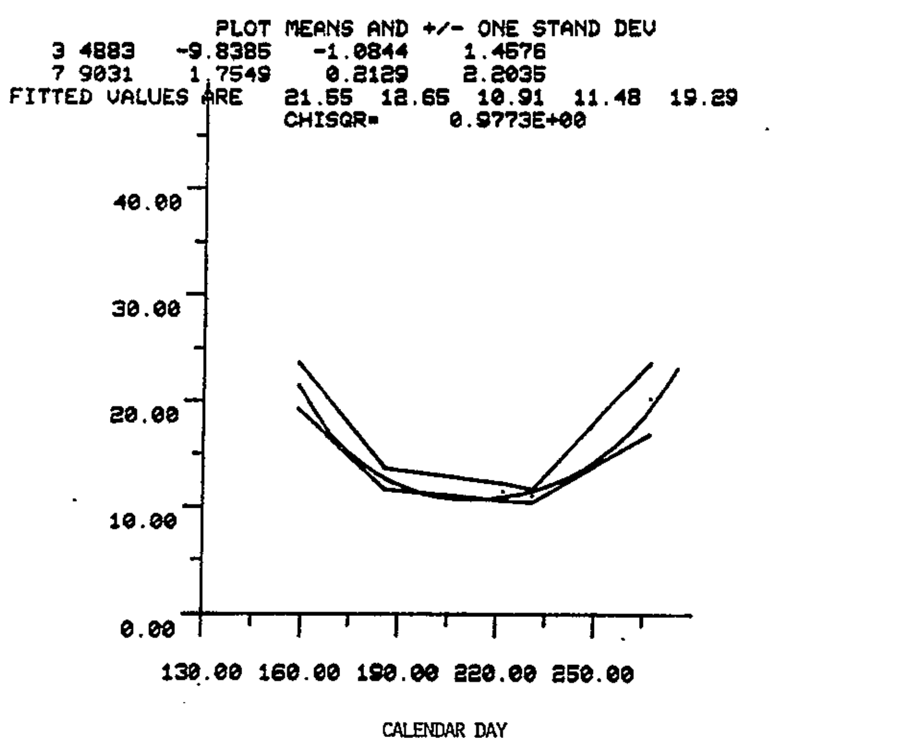
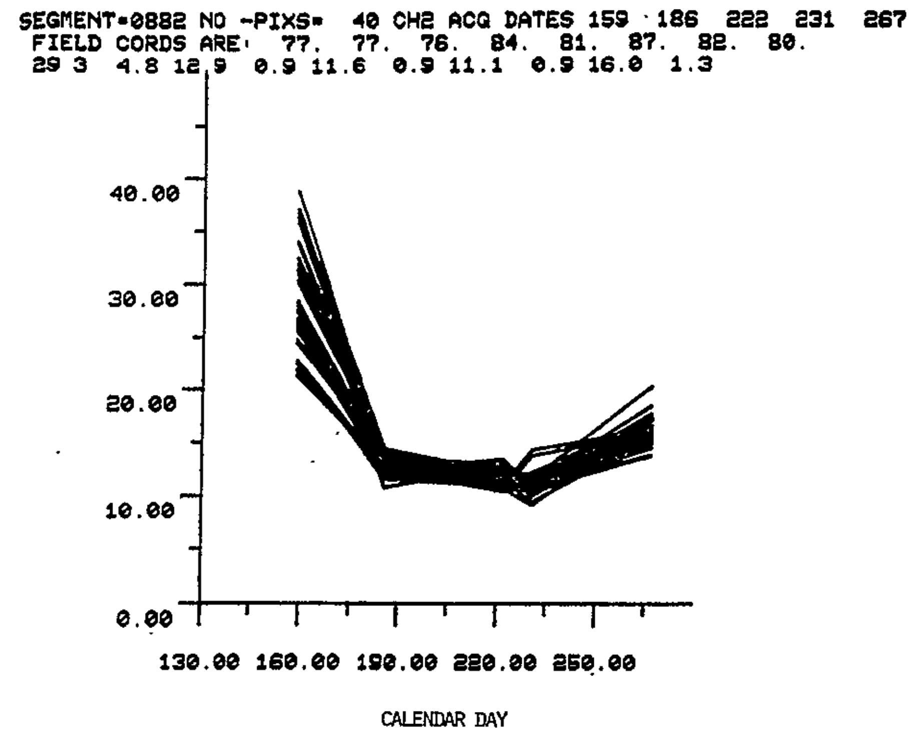


Figure 1.— Sample segment 882, training field 1.



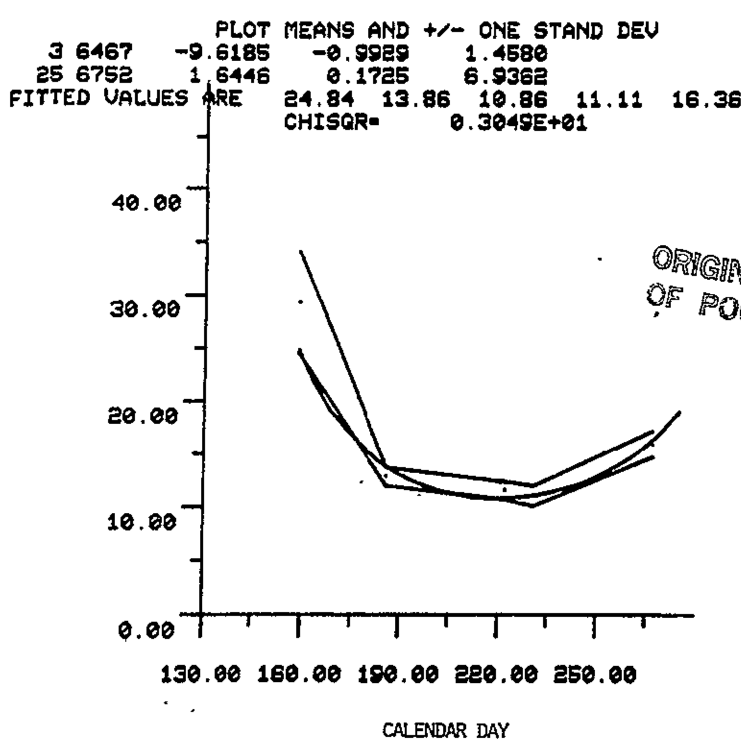
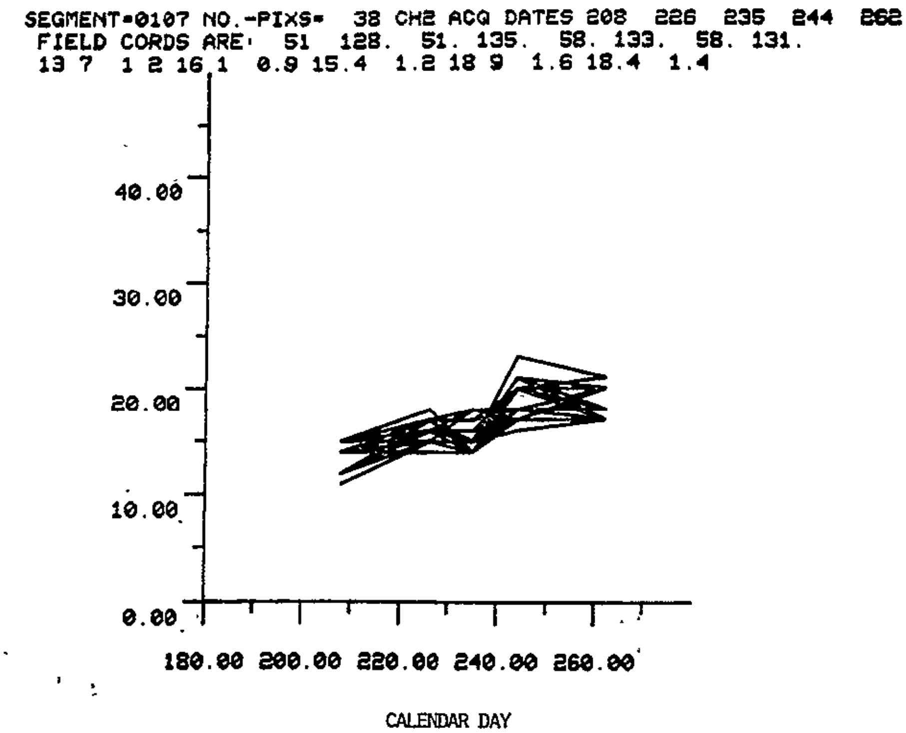


Figure 2.— Sample segment 882, training field 2.



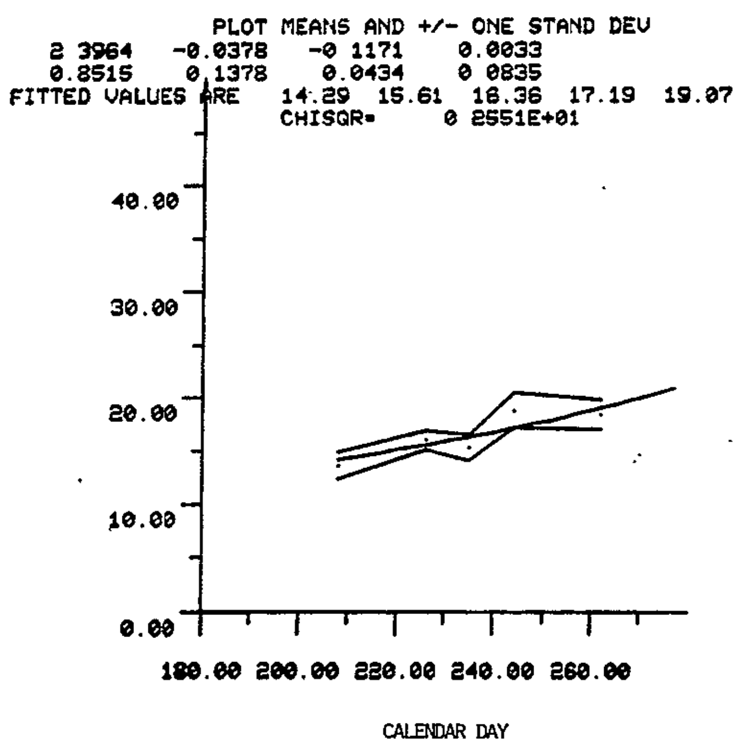
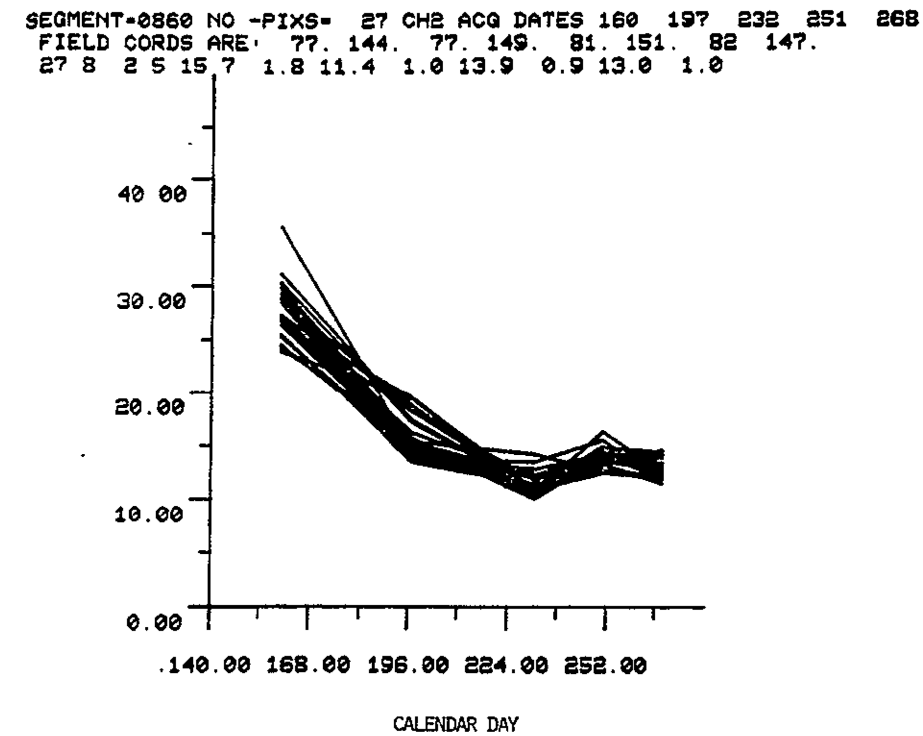


Figure 3.— Sample segment 107, training field 3.



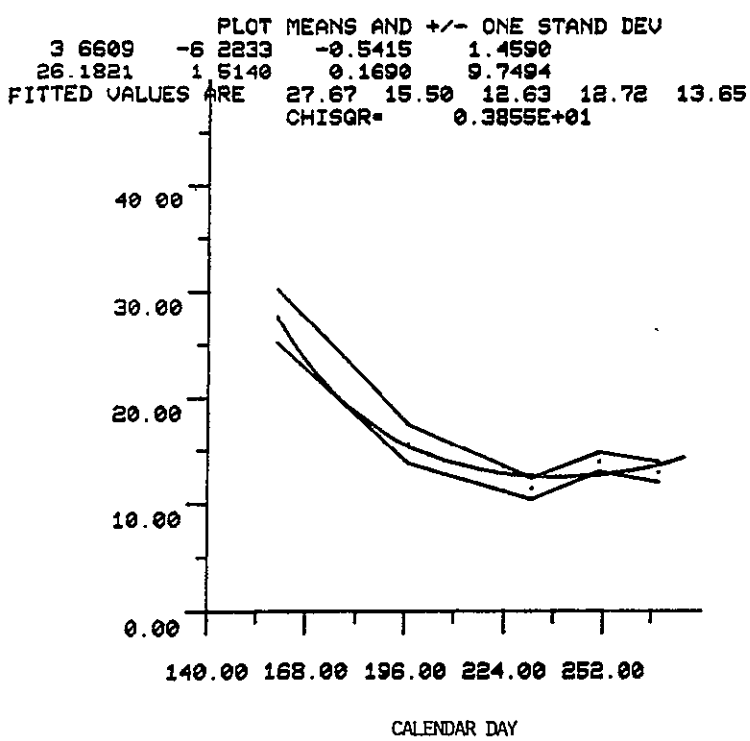
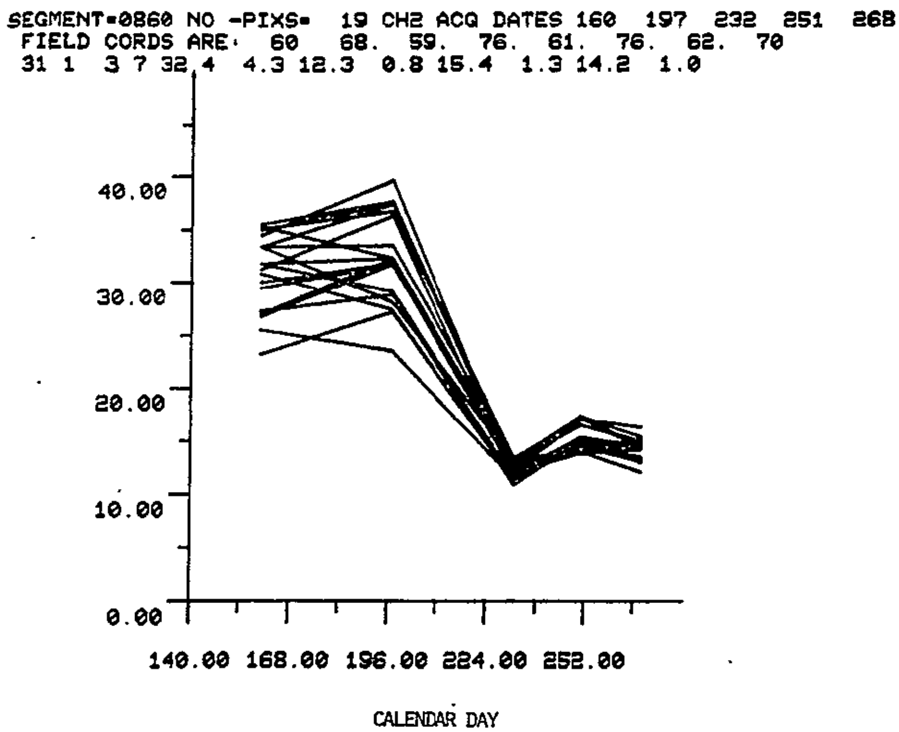


Figure 4.— Sample segment 860, training field 1.



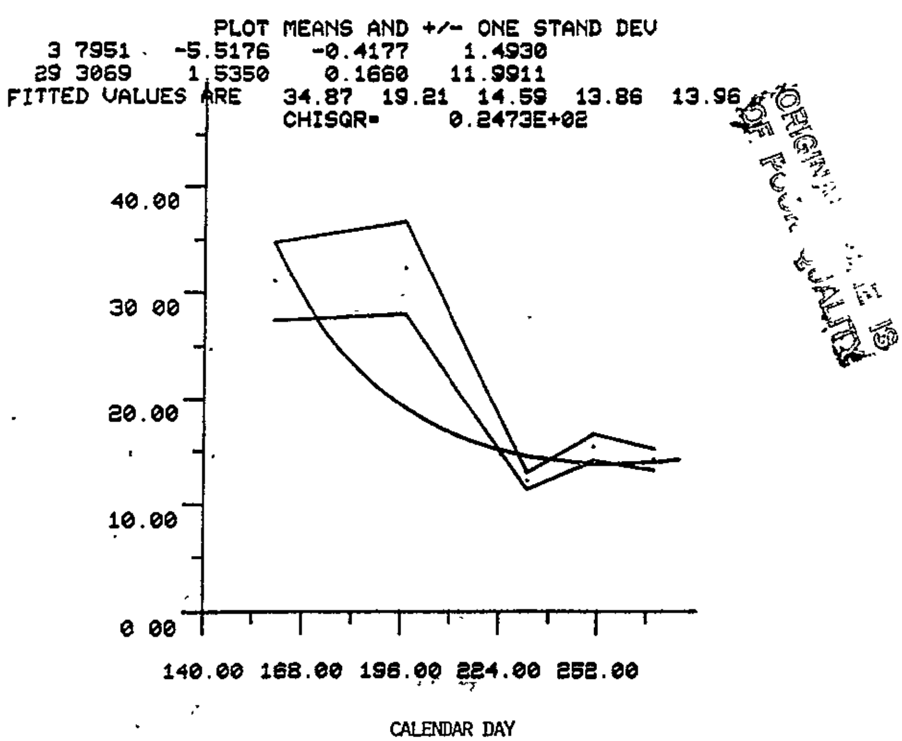


Figure 5.— Sample segment 860, training field 2.

field 2 is not acceptable as a basis for classification. The data in training field 1 (fig. 4) is adequately approximated by the curve; training field 1 is acceptable for use in classification. By proper choice of a training field and an acquisition set, a representative crop profile was defined in each channel for use in classification.

b. Chi-square fit values in each channel should be less than 10. This criterion is dependent upon the acquisition selection criteria just described. Notice that the chi-square value in figure 3(b) is 2.5 and 3.0 in figure 2(b). A misleadingly good chi-square fit is also generated when the data are dispersed; many different curves fit well for dispersed data with large standard deviations. Chi-square values give an estimation of the adequacy of the crop profile curve as an approximation of the training field data. It is sufficient *only* if (1) a curve is defined and (2) the data are compact enough that representation by a single curve is sensible.

c. Estimated planting date for the training field as generated on each channel should be the same within the estimated planting date error. This criterion was seldom violated since the estimated planting date error tended to be quite large.

As mentioned in section 2.2, data quality, atmospheric effects, and crop growth stages were assessed by using the film products; but for the Badhwar classification method, final judgment of usability depended upon the curve definition. For sample segment 144, day 165, the imagery showed poor data quality. However, this is the only date usable for curve definition in early growth stages, and it proved to be adequate. Atmospheric effects were also evaluated from the graphs. If an apparently hazy, but necessary, acquisition graphed successfully into the crop profile, it was accepted for use. If use of a cloudy acquisition could not be avoided, training fields were selected free of clouds; and resultant screening-out or misclassification of cloudy corn fields was accepted as a penalty. Very early and very late acquisitions, when corn was apt to be pre-emergent or harvested, were not used. There is, however, no way of being absolutely sure that these conditions have been avoided. Successful generation of crop profile confirmed the imagery growth stage assessment.

Acquisitions marred by poor data quality, atmospheric effects, unusual crop conditions (such as hail damage or fields cut for silage) or too early or too late crop growth stage can be used. The training field will be chosen free of these conditions and hence will be unrepresentative of some corn in the segment. The corn affected by the conditions will tend to be misclassified; this is programmatically correct.

4.2 ANALYST EVALUATION OF OUTPUT CLASSIFICATION RESULTS

Input to CLASFY are the line-pixel coordinates of the training field selected as a basis for classification, an acquisition set of sufficient temporal distribution to characterize the crop profile in the four Landsat channels, and initial values for the constants in the modeling function. Outputs are a summary sheet of input values, calculated parameters and pixel classification, and a full scene classification file, which is translated to a lineprinter map of the classification and to a misclassification map of the classification file and the ground-truth file.

- Figure 6 is a copy of the lineprinter summary sheet for sample segment 882.
- Figure 7 is a copy of the first quadrant of the lineprinter classification map.
- Figure 8 is a copy of the misclassification map of AA pure pixels only.
- Figure 9 is a copy of the misclassification map of all pixels.

The classification was evaluated using several criteria. The classification map should be "clean." Field patterns should be evident and fields well filled out with a minimum of blank spaces (pixels rejected as corn) on the interiors. Blank areas should also be clear and be reasonably free of scattered pixels classified as corn.

Classification comparison with the ground truth should have a statistical agreement greater than 70 percent and areas of disagreement should be examined. In evaluating areas of disagreement, these points are important.

1. The training field is selected with the constraints that the field should not be haze or cloud covered on any acquisition, should not be harvested

```
CLASSIFICATION FILE = J821[131,3]C88279232.DC1        RUN TIME =  92.46 MINUTES
ACCURACY ASSESSMENT QUALITY ASSURANCE CLASSIFICATION

PROCESSING DATE =  29-AUG-79 AT 16:37:47

SEGMENT NUMBER =  882     CROP OF INTEREST = CORN

IMAGE FILES USED IN CLASSIFICATION = L821[131,3]088278159.IH3
                                     L821[131,3]088278186.IH2
                                     L821[131,3]088278222.IH2
                                     L821[131,3]088278231.IH3
                                     L821[131,3]088278267.IH3

TRAINING FIELD =  LINE NO.     SAMP. NO.

                    77.0          77.0
                    76.0          84.0
                    81.0          87.0
                    82.0          80.0

MEANS AND STD. DEV. FOR TRAINING FIELD BASED ON 36 PIXELS =

   CHANNEL                 ----------ACQUISITION DATES----------
   NUMBER                  78159   78186   78222   78231   78267

      1     MEAN           27.46   20.20   17.06   16.98   17.41
            STD. DEV.       3.54    0.98    1.18    1.34    1.35

      2     MEAN           29.18   12.92   11.61   11.01   15.83
            STD. DEV.       4.85    0.87    0.86    0.81    1.07

      3     MEAN           34.93   59.14   49.50   45.93   30.97
            STD. DEV.       5.61    3.21    1.65    2.25    1.00

      4     MEAN           29.95   57.95   53.44   49.43   28.66
            STD. DEV.       4.59    2.41    1.27    2.34    1.84

CONSTANTS FOR MODEL =

   CHANNEL
   NUMBER              A               ALPHA             BETA              T0               CHISQ

      1     INITIAL   3.48            -3.65            -0.31             1.50
            FINAL     3.45+- 2.21     -3.69+- 1.42     -0.32+- 0.15      1.48+- 1.44       0.09

      2     INITIAL   3.64            -9.61            -0.90             1.50
            FINAL     3.56+- 2.44     -9.48+- 1.53     -0.98+- 0.16      1.48+- 0.69       3.00

      3     INITIAL   3.71             4.70             0.62             1.50
            FINAL     3.73+- 4.33      4.56+- 1.06      0.61+- 0.10      1.48+- 3.39       6.07

      4     INITIAL   3.44             7.98             0.96             1.50
            FINAL     3.46+- 1.27      7.84+- 1.05      0.96+- 0.11      1.49+- 0.53       5.04

CHISQ THRESHOLD =  CHANNEL     1      2      3      4

                 THRESHOLD   4.62   6.36   8.50   8.60

CLASSIFICATION RESULTS =   PIXELS CLASSIFIED CORN      =  7128
                           PIXELS SCREENED             =  2576
                           PIXELS CLASSIFIED NON-CORN  = 13228    CUT BY CH2 =  3508
                                                                  CUT BY CH3 =  7724
                                                                  CUT BY CH4 =  1996
```


Figure 6.— Lineprinter summary sheet for sample segment 882.

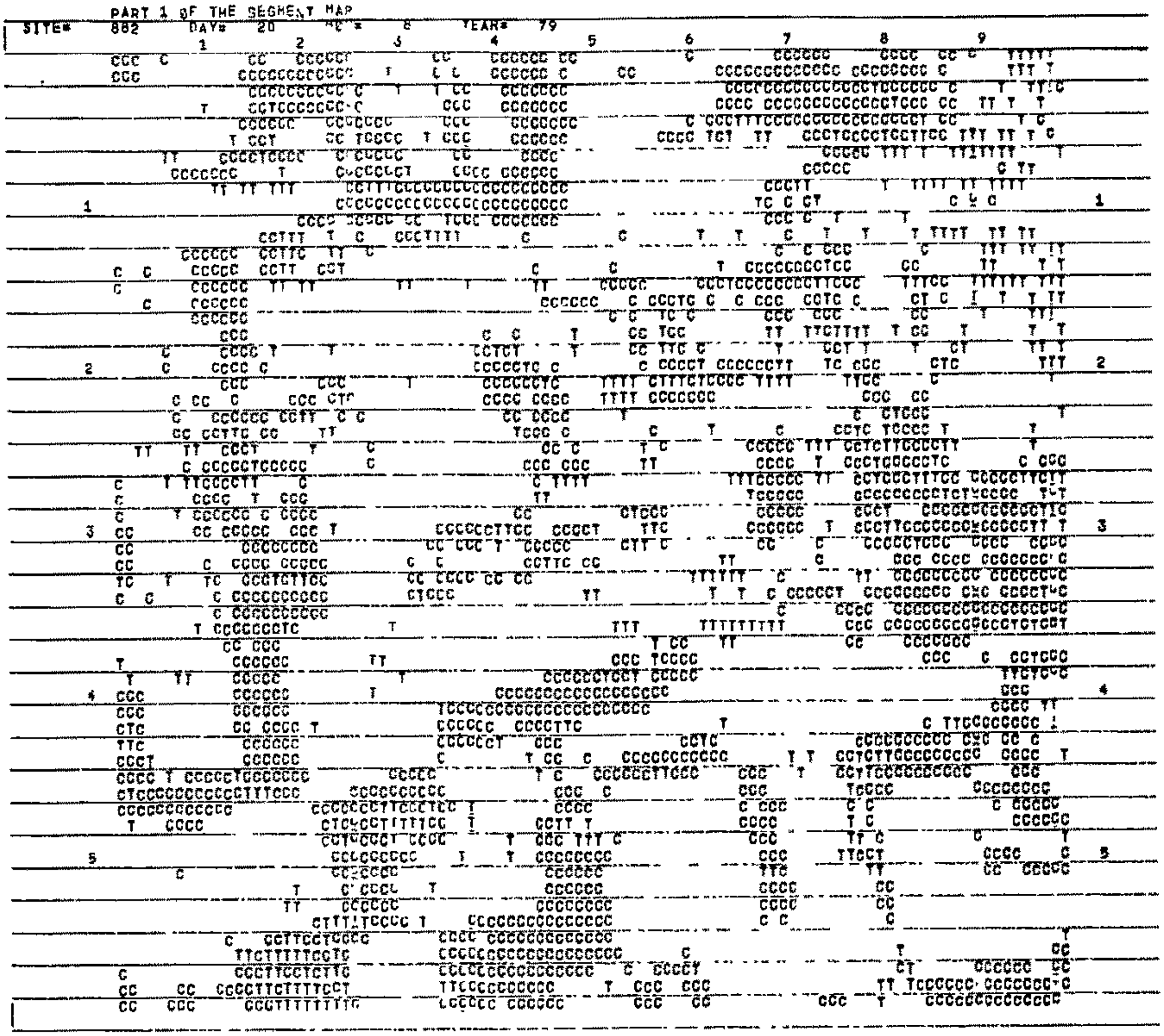


Figure 7.— First quadrant of the lineprinter classification map.

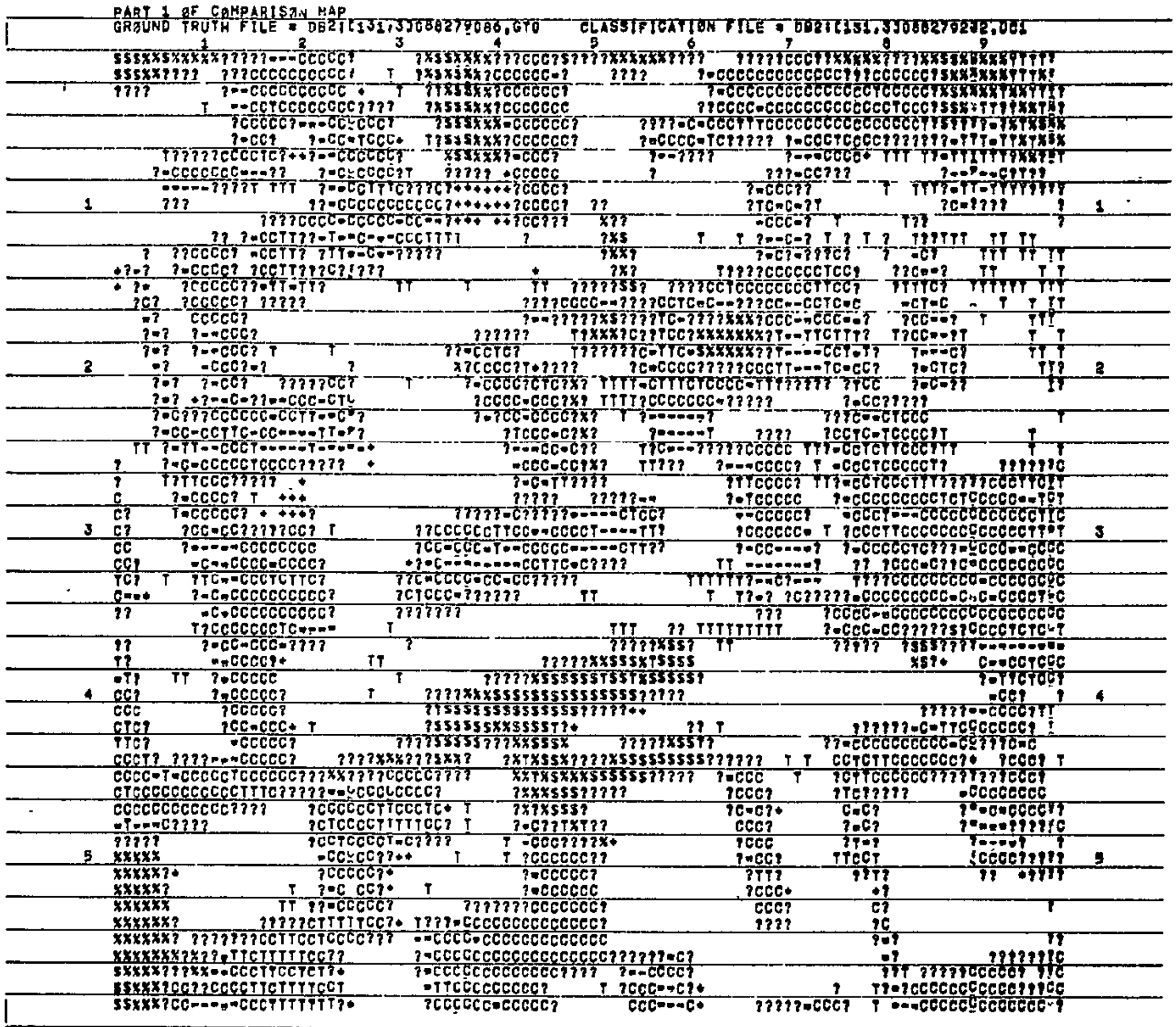


Figure 8.— Misclassification map, AA pure pixels only.

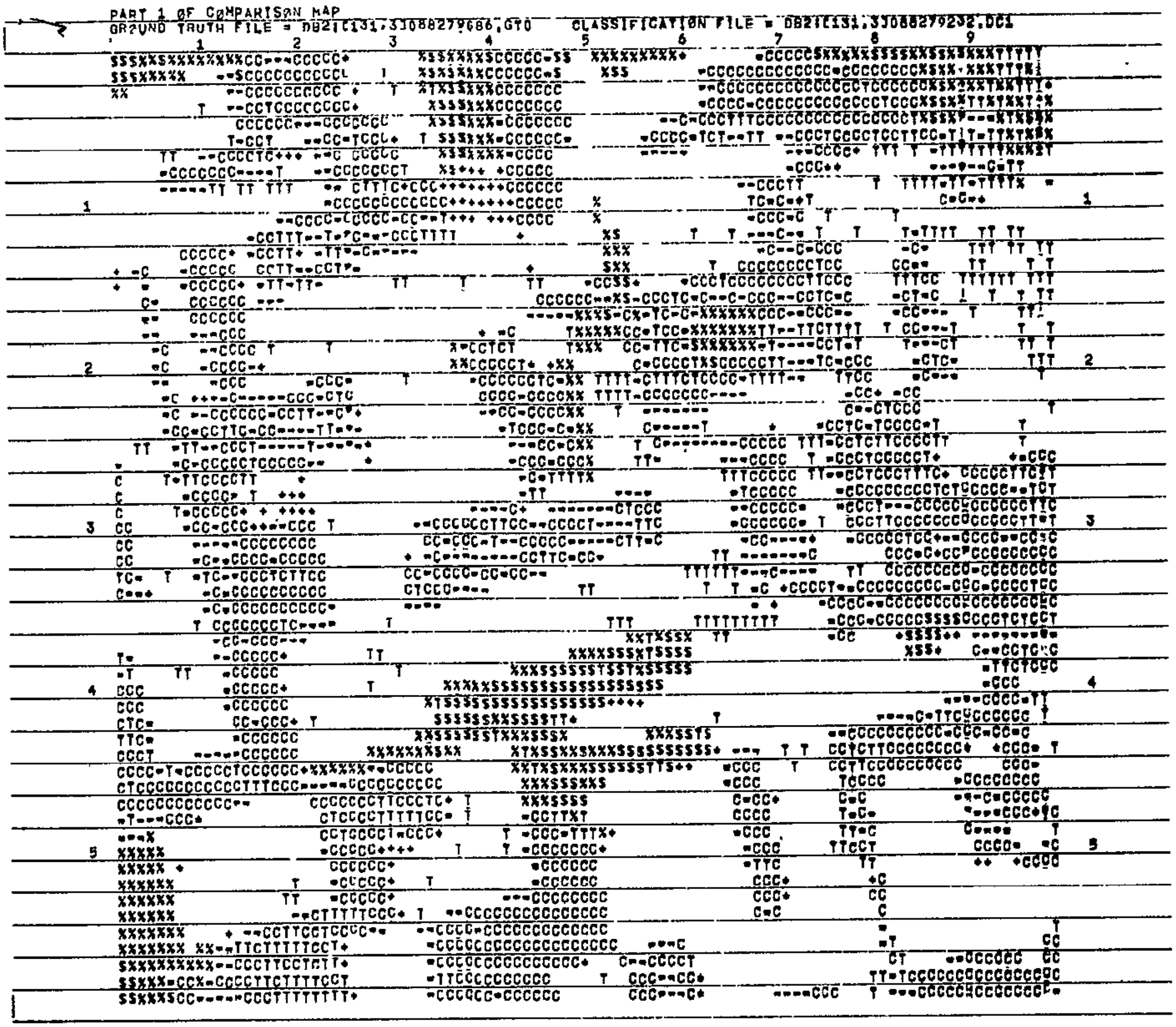


Figure 9.— Misclassification map, all pixels.

on the final date, or should not represent an abnormal crop condition (such as hail-damaged or cut for silage). Consequently, this method may not classify fields that are harvested on the final date or that are abnormal in development into the crop-of-interest category. This is "programmatically correct."

2. Field borders — mixed pixels in general — will tend to be rejected from the crop-of-interest category. Again, this is "programmatically correct"; the effect is more pronounced in a two-class classifier, such as this, than in a maximum likelihood classifier. In this study, by far the greatest area of disagreement between the classifications and the ground-truth inventory maps was border pixels. Some indication of the number of border pixels in an individual segment can be derived from the number of AA pure pixels compared to the total number in a Landsat segment; this is, however, a subset of the number of pixels which must be defined as mixed when multitemporal classification of any type is used.

Rework techniques were limited to choice of a different set of acquisitions, use of four-acquisition classification instead of five, or selection of a different training field as the basis of classification. With a satisfactory set of acquisitions, training field selection was the most effective rework tool. A few classifications were done with relaxed training field criteria (this is noted on the tables). The effect of a relaxed chi-square criterion seemed to be less distinct field patterns and a more speckled classification map. Relaxing the quality-of-the-curve requirements seemed to produce an overclassification or an underclassification.

5. RESULTS

The values presented in table 1 must be considered *initial* results and should be viewed as a basis for program development. Evaluation of Badhwar classification as a potential quality assurance check in AA led to the definition of a large sample of corn/soybean segments to which this method could be applied. As noted in previous sections, within the tentative guidelines, considerable experimentation, testing, and exploration of limits and guideline validity were conducted. No attempt was made to optimize the procedure based on the results.

Twenty-six of the 29 segments defined for the Multicrop Exploratory Experiment Test were classified. Results could not be generated for sample segments 107 (Boone County, Illinois), 807 (Henry County, Indiana), and 893 (Webster County, Iowa).

For each segment, the following items are listed in table 1:

a. Sample segment number and location are noted. An asterisk (*) indicates that the training fields were defined by the analyst.

b. Acquisitions available are listed. Consecutive-day acquisitions are omitted unless used for classification, and acquisitions outside the growing season for corn (Julian days 130 to 300) are omitted.

c. Acquisitions used for the classification results presented in table 1 are noted. Landsat-3 acquisitions are denoted by (3).

d. The coordinates of the training field used for classification and the number of pixels in the field are listed.

e. The percentage of the scene correctly classified (PSCC) based on AA pure pixels only is given.

f. The number of pixels used to compute the percentage above is listed. For PSCC based on pure pixels, the base is the number of AA pure pixels in the scene minus the number edited by SCREEN minus the number for which there is no ground truth.

g. A confusion matrix of the classification in scene percentages is given: (1) ground-truth corn classified as corn, (2) ground-truth corn classified as non-corn, (3) ground-truth non-corn classified as corn, and (4) ground-truth non-corn classified as non-corn. These percentages are based on calculations by the MISMAP program.

h. The PSCC based on all pixels, mixed and AA pure, is given.

i. The number of pixels used to compute the "all pixel" percentages is listed. For PSCC "all pixels," the base is 22 932 pixels minus the number edited by SCREEN minus the number for which there is no ground truth.

j. A confusion matrix of the classification of all pixels in scene percentages is given.

k. Additional comments are given in the last column of the table.

TABLE 1.— CLASSIFICATION RESULTS

Segment and location	Acquisitions available	Acquisition used	Training field coordinates	PSCC (a)	Base (a)	Confusion matrices, % (C → C / N → C)	(C → N / N → N)	Comments
123 Hamilton, Ind.	152, 161, 197, 233, 269	152(3), 161, 197, 233, 269	(3,58), (3,64), (8,66), (8,60), 33 pixels	86.8 - - - 82.7	12 897 - - - 15 777	24.8 6.3 - - - 24.3 7.7	6.8 62.0 - - - 9.4 58.4	Acquisition coverage marginal. Severe screening problems.
127* Montgomery, Ind.	152, 161, 197, 207, 216, 243, 252, 269	161, 207(3), 216, 243(3), 252	(66,123), (66,133), (73,136), (73,126), 56 pixels	85.9 - - - 82.5	16 817 - - - 19 860	43.2 5.4 - - - 40.8 7.2	8.6 42.7 - - - 10.4 41.7	Problems with data quality; additional early acquisition (198) on order.
133* Whitley, Ind.	152, 197, 233, 251, 260, 269	152(3), 197, 233, 251, 269	(3,108), (3,115), (8,117), (8,110), 38 pixels	79.8 - - - 75.6	16 440 - - - 20 684	15.1 5.9 - - - 15.0 6.7	14.8 64.7 - - - 17 7 60.6	Small Indiana fields; misregistration.
135 Chickasaw, Iowa	130, 166, 229, 247, 265, 274, 283, 292	166, 229, 247(3), 265(3), 274	(25,146), (25,155), (33,156), (33,147), 36 pixels	79.0 - - - 75.0	17 599 - - - 21 969	20.4 3.0 - - - 18.7 4.1	18.0 58.7 - - - 21.0 56.3	SCREEN removed agricultural fields; so classification is done without any screening. Marginal acquisition coverage and data quality. Training field exceeded criteria in channel 3 chi square.
141* Madison, Iowa	130, 167, 212, 220, 256, 265, 274, 292	167, 212(3), 220, 265(3), 274 274	(14,161), (13,169), (17,172), (18,163), 39 pixels	83.6 - - - 79.8	17 138 - - - 19 722	14.8 8.9 - - - 15.7 9.7	7.6 68.8 - - - 10.8 64.1	
144 Wapello, Iowa	130, 165, 183, 219, 238, 246, 264, 274, 292, 300	165, 219, 238, 246(3), 264(3)	(27,166), (27,172), (40,178), (40,171), 39 pixels	90 - - - 86.4	15 270 - - - 17 864	9.5 1.3 - - - 9.5 1.9	8.3 80.5 - - - 11.6 76.9	Screening removed on acquisition 165. Poor data quality; marginal acquisition coverage.
202 Atchison, Mo.	167, 212, 221, 266, 275, 284, 293	167, 212(3), 221, 266(3)	(17,9), (17,16), (20,18), (20,9), 33 pixels	83.5 - - - 80.6	18 709 - - - 20 914	10.6 6.6 - - - 10.9 7.0	9.7 72.9 - - - 12.4 69.7	Acquisitions inadequate. Clouds, haze affect days 212, 221, and 275.
205 Clark, Mo.	137, 155, 164, 209, 218, 219, 246, 272, 282, 290	155(3), 218, 219, 246(3), 272	(28,182), (28,191), (33,192), (33,181), 37 pixels	84.7 - - - 82.5	16 545 - - - 18 736	5.9 3.9 - - - 6.5 4.3	11.4 78.8 - - - 13.3 76.0	Marginal acquisition coverage (164 is one-third haze, clouds and cannot be used).
216* Mercer, Mo.	130, 184, 202, 220, 238, 247, 265, 274, 292	184, 220, 247(3), 274	(31,184), (31,187), (37,187), (37,184), 22 pixels	87.7 - - - 86.6	19 582 - - - 20 341	2.6 9.4 - - - 3.2 9.7	2.7 85.1 - - - 3.7 83.4	Low corn segment, small fields, excessive screening on days 202 and 265. Chi square = 0.0 in channel 1; missed some ground-truth corn fields.
800 Clinton, Iowa	130, 164, 218, 219, 246, 272, 281, 290, 300	164, 219, 246(3), 281(3)	(68,14), (68,34), (70,34), (70,15), 38 pixels	67.2 - - - 65.3	16 899 - - - 20 616	26.7 3.0 - - - 24.4 3.9	29.7 40.5 - - - 30.8 40.9	Corn crop was hail-damaged before acquisition 246; considerable harvest evident on day 281.
809* Ogle, Ill.	164, 209, 218, 244, 262, 271, 281, 289	164, 218, 244(3), 262(3), 271	(20,18), (19,23), (27,26), (28,22), 26 pixels	76.8 - - - 73.8	14 063 - - - 17 130	46.3 6.9 - - - 42.2 7.6	16.4 30.5 - - - 18.5 31.6	Segment more than 50% corn. Training field exceeded chi-square criteria in channel 1.
832 Adams, Ind.	151, 160, 178, 232, 268, 304	151(3), 160, 232, 268	(91,143), (91,153), (94,153), (94,143), 39 pixels	75.5 - - - 72.5	17 335 - - - 20 662	12.2 17.8 - - - 13.9 18.6	6.6 63.3 - - - 8.9 58.6	Inadequate acquisition coverage; day 178 is cloudy.
837* Benton, Ind.	180, 198, 207, 216, 225, 234, 243, 252, 270	180, 198, 216, 234, 252	(98,35), (96,47), (101,48), (104,37), 39 pixels	80.0 - - - 76.5	17 204 - - - 20 409	37.5 13.2 - - - 35.7 14.5	6.8 42.5 - - - 8.9 40.8	Training field exceeds chi-square criteria in channels 3 and 4; good acquisition coverage.
842* Henry, Ind.	160, 176, 232, 250, 268	178, 232, 250, 268	(40,156), (40,162), (45,164), (47,160), 32 pixels	77.2 - - - 73.0	14 436 - - - 17 589	29.7 10.3 - - - 28.4 12.1	12.3 47.5 - - - 14.9 44.6	Training field exceeds chi-square criteria in channel 1, more screening than apparent data quality merits.

[a]The PSCC, the base, and the confusion matrices for "pure pixels" appear above the dashed line; for "all pixels," below the dashed line.

TABLE 1.— Concluded.

Segment and location	Acquisitions available	Acquisition used	Training field coordinates	PSCC (a)	Base (a)	Confusion matrices, % (C → C) (N → C)	(C → N) (N → N)	Comments
843* Henry, Ind.	152, 160, 178, 197, 233, 251, 268	178, 197, 233, 251	(13,121), (13,123), (22,127), (22,125), 22 pixels	80.5	17 605	17.6 6.3	13.0 62.9	
				77.0	21 327	17.4 7.2	15.9 59.6	
852* Randolph, Ind	151, 160, 178, 232, 250, 268	178, 232, 250, 268	(49,36), (49,44), (51,46), (53,37), 32 pixels	82.8	16 974	9.1 1.9	15.4 73.7	Training field exceeds chi-square criteria in channels 1 and 2, some misregistration, wide range of planting dates and much of the corn failed to keep signature sequence through day 178.
				72.6	19 492	1.8 1.8	25.6 70.8	
853* Randolph, Ind.	151, 160, 178, 232, 250, 268	160, 178, 232, 250, 268	(35,68), (34,72), (41,77), (43,71), 32 pixels	75.6	15 135	13.2 3.4	20.7 62.4	Training field exceeds criteria in channels 1, 2, and 4; severe problems with chi-square criteria on this segment
				72.3	18 552	13.1 4.4	23.1 59.2	
860* Wells, Ind.	151, 160, 178, 197, 232, 251, 268	160, 197, 232, 251, 268	(91,61), (91,66), (95,67), (95,62), 27 pixels	82.8	16 309	16.7 5.5	11.7 66.1	Misregistration.
				78.8	19 928	16.8 6.7	14.6 62.0	
864* Crawford, Iowa	150, 159, 186, 222, 231, 249, 258, 267, 294	159(3), 186 222, 231(3), 267(3)	(65,116), (65,120), (69,120), (69,116), 24 pixels	77.5	15 430	24.8 2.4	20.0 52.7	
				73.8	19 308	22.4 3.3	22.8 51.4	
865* Crawford, Iowa	150, 159, 168, 186, 231, 249, 267, 294	168, 186, 231(3), 249(3), 267(3),	(65,31), (65,35), (72,38), (73,35), 32 pixels	82.3	17 562	22.3 6.4	10.4 61.0	
				79.0	21 877	22.0 7.9	14 0 57 0	
877* Ida, Iowa	150, 186, 222, 231, 267	186, 222, 231(3), 267(3)	(30,61), (28,67), (33,69), (36,63), 38 pixels	72.3	13 896	26.1 4.8	23.2 46.2	Segment has large areas for which ground truth was not available.
				69.3	17 818	23.2 5.7	25.1 46.1	
878 Kossuth, Iowa	131, 186, 221, 266, 293	186, 221, 266(3), 293	(9,140), (9,148), (19,151), (19,144), 40 pixels	71.6	18 360	18.5 3.0	25.5 53.1	Segment processed as an experiment; day 293 too late to be used.
				69.4	21 923	18.0 4.0	26.7 51.4	
880 Monona, Iowa	150, 186, 204, 222, 231, 249, 267, 294	186, 222, 231(3), 267(3)	(46,123), (46,133), (54,136), (54,127), 35 pixels	77.4	17 934	23.2 0.9	21.8 54 2	
				74.3	21 189	20.9 1.4	24.4 53.4	
881* Monona, Iowa	159, 186, 213, 222, 231, 249, 267	159(3), 186, 222, 231(3), 267(3)	(13,81), (12,90), (16,91), (17,82), 39 pixels	83.6	17 916	29.1 1.3	14.9 54.5	
				79.1	21 740	26.6 2.4	18.5 52.5	
882* Palo Alto, Iowa	150, 159, 186, 213, 222, 231, 258, 267, 293	159(3), 186, 222, 231(3), 267(3)	(77,77), (76,84), (81,87), (82,80), 36 pixels	87.3	15 967	34.1 2.0	10.7 53.2	
				83.3	19 057	31.9 3.2	13.6 51.4	
891 Shelby, Iowa	159, 168, 186, 204, 249, 258, 267	168, 186, 204, 258, 267(3)	(90,5), (90,13), (97,13), (97,5), 37 pixels	73.4	14 828	25.7 4.5	21.4 47.7	Training fields exceeded chi-square criteria in channels 3 and 4; data quality marginal.
				70.0	19 010	23.9 5.9	23.9 46.1	

[a]The PSCC, the base, and the confusion matrices for "pure pixels" appear above the dashed line; for "all pixels," below the dashed line.

6. CONCLUSIONS AND RECOMMENDATIONS

Badhwar classification provides an approach to classification different from that of the classification methods currently in use. Basic data requirements are similar to those of current methods, however; and Badhwar classification is applicable to a very large percentage of LACIE sample segments. This paper documents initial results as a base for program development. Accuracy averaged over the 26 segments is 80.4 percent for AA pure pixels only and 76.6 percent for all pixels.

Further research using Badhwar classification should be conducted on a subset of the site data set used for this study. The following set of sample segments is recommended:

127	852
133	853
135	860
141	864
205	865
216	877
809	880
837	881
842	882
843	

The available acquisitions for these particular sites seem to be well suited for meeting the criteria of data quality, acquisition distribution, and coverage of the corn growth cycle.

Further evaluation of this classification technique is recommended.

7. REFERENCES

1. Badhwar, G.: Crop Emergence Date Determination from Spectral Data. In press (Photogram. Eng. Remote Sensing).

2. Badhwar, G.: A Semi-Automatic Technique for Multitemporal Classificatio of a Given Crop. July 31, 1979.

3. Carnes, J. G.: Results of the Accuracy Assessment Quality Assurance Evaluation for Corn/Soybean Verification Test Sites. Presentation to NASA and LEC, ref. 644-1345, Sept. 1979.

4. Abotteen, K. M.; and Dailey, C. L.: Procedures Control Report, Multicrop Exploratory Experiment Test. Letter to L. M. Flores (LEC), ref. 644-121(May 21, 1979.

5. Dailey, C. L.; and Abotteen, K. M.: Procedures Control Report, Simulated Aggregation Test. Letter to R. O. Hill (SF4), July 13, 1979.

6. Lambeck, P. F.: Signature Extension Preprocessing for Landsat MSS Data. ERIM 122700-32-F, Nov. 1977.

7. Hocutt, W. T.: Use of SCREEN Algorithm to Flag Certain Questionable Pixels in Landsat Images. Letter to A. G. Houston (SF4), ref. 642-7746, July 1979.

8. Hocutt, W. T.: Alternative Screening Technique for AA Image Data Base. Letter to A. G. Houston (SF4), ref. 644-1315, Aug. 1979.

9. Wehmanen, O. A.: Landsat-3 Calibration Factors. Letter to R. O. Hill (NASA/JSC), ref. 644-1044, Sept. 13, 1978.

10. Badhwar, G.: A Program Called TRJPLT for Use on the PDP-11/45 Image Processor. SF3/6541, Feb. 5, 1979.

11. Carnes, J. G.: Quality Control Classification Algorithm Development. Letter to A. G. Houston (SF4), ref. 642-7673, May 17, 1979.

12. Austin, Willa W.: Detailed Description of the Wheat Acreage Estimation Procedure Used in the Large Area Crop Inventory Experiment. LEC-11497, Feb. 1978.