N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED IN THE INTEREST OF MAKING AVAILABLE AS MUCH INFORMATION AS POSSIBLE

NASA CR-167409
AUG 26 1981

# AgRISTARS

SR-L1-04045
E82-10150

"Made available under NASA sponsorship in the interest of early and wide dissemination of Earth Resources Survey Program information and without liability for any use made thereof."

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

## Supporting Research

August 1981

# CLASSIFICATION OF CORN: BADHWAR PROFILE SIMILARITY TECHNIQUE

Willa W. Austin

(E82-10150) AGRISTARS: SUPPORTING RESEARCH. CLASSIFICATION OF CORN: BADHWAR PROFILE SIMILARITY TECHNIQUE (Lockheed Engineering and Management) 46 p HC A03/MF A01 CSCL 02C G3/43
N82-21676
Unclas
00150

Lockheed Engineering and Management Services Company, Inc.
1830 NASA Road 1, Houston, Texas 77058

Lyndon B. Johnson Space Center
Houston, Texas 77058

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| SR-L1-04045, JSC-17113 | | |

| 4. Title and Subtitle | 5. Report Date |
|---|---|
| Classification of Corn: Badhwar Profile Similarity Technique | August 1981 |
| | 6. Performing Organization Code |
| 7. Author(s) | 8. Performing Organization Report No. |
| Willa W. Austin<br>Lockheed Engineering and Management Services Company, Inc. | LEMSCO-16035 |
| 9. Performing Organization Name and Address | 10. Work Unit No. |
| Lockheed Engineering and Management Services Company, Inc.<br>1830 NASA Road 1<br>Houston, Texas 77058 | 11. Contract or Grant No. NAS 9-15800 |
| 12. Sponsoring Agency Name and Address | 13. Type of Report and Period Covered |
| National Aeronautics and Space Administration<br>Lyndon B. Johnson Space Center<br>Houston, Texas 77058 Technical Monitor: J. D. Erickson | Technical Report |
| | 14. Sponsoring Agency Code |

15. Supplementary Notes

The Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing program is a joint endeavor of the U.S. Departments of Agriculture, Commerce, and the Interior; the National Aeronautics and Space Administration; and the Agency for International Development.

16. Abstract

This report documents an extension of previous applications of the Profile Similarity Classification Technique developed by Dr. Gautam Badhwar of NASA. The same software programs used to classify spring wheat are applied to the classification of corn; numerical results of the acreage estimation are given. Potential problem areas defined in an earlier application are examined in detail.

| 17. Key Words (Suggested by Author(s)) | 18. Distribution Statement |
|---|---|
| Landsat data<br>Profile Similarity technique<br>Classification of corn<br>Spectral separability<br>Classification techniques | |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price* |
|---|---|---|---|
| Unclassified | Unclassified | 45 | |

*For sale by the National Technical Information Service, Springfield, Virginia 22161

JSC Form 1424 (Rev Nov 75)

NASA — JSC

PRECEDING PAGE BLANK NOT FILMED

SR-L1-04045
JSC-17113

# CLASSIFICATION OF CORN: BADHWAR PROFILE SIMILARITY TECHNIQUE

Job Order 71-306

This report describes classification activities of the Supporting Research project of the AgRISTARS program.


PREPARED BY

Willa W. Austin


APPROVED BY

R. K. Lennington, Supervisor
Techniques Development Section

T. C. Minter, Manager
Development and Evaluation Department


LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.

Under Contract NAS 9-15800

For

Earth Resources Research Division

Space and Life Sciences Directorate

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

August 1981

LEMSCO-16035

PRECEDING PAGE BLANK NOT FILMED

# PREFACE

The Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing is a multiyear program of research, development, evaluation, and application of aerospace remote sensing for agricultural resources, which began in fiscal year 1980. This program is a cooperative effort of the U.S. Department of Agriculture, the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration (U.S. Department of Commerce), the Agency for International Development (U.S. Department of State), and the U.S. Department of the Interior.

The work which is the subject of this document was performed within the Earth Resources Research Division, Space and Life Sciences Directorate, at the Lyndon B. Johnson Space Center, National Aeronautics and Space Administration. Under Contract NAS 9-15800, personnel of Lockheed Engineering and Management Services Company, Inc., performed the tasks which contributed to the completion of this research.

PRECEDING PAGE BLANK NOT FILMED

CONTENTS


Section Page

1. INTRODUCTION ..... 1-1

2. DATA SET ..... 2-1

2.1 SITE SELECTION ..... 2-1

2.2 ACQUISITION SELECTION ..... 2-1

2.3 TRAINING FIELD SELECTION ..... 2-4

3. SOFTWARE ..... 3-1

3.1 IMUNLD2A ..... 3-1

3.2 IMAPLT ..... 3-1

3.3 CLASFYT ..... 3-2

3.4 MISMAP ..... 3-3

3.5 MISMAP1 ..... 3-4

3.6 TAPEOUT ..... 3-4

4. PROCEDURE ..... 4-1

5. DISCUSSION OF POTENTIAL PROBLEM AREAS ..... 5-1

5.1 THE EFFECTIVENESS OF FOUR-ACQUISITION CLASSIFICATION COMPARED TO FIVE-ACQUISITION CLASSIFICATION ..... 5-1

5.1.1 METHOD USED TO ASSESS PROBLEM ..... 5-1

5.1.2 SUMMARY ..... 5-7

5.2 INADEQUATE ACQUISITION COVERAGE ..... 5-7

5.2.1 METHOD USED TO ASSESS PROBLEM ..... 5-7

5.2.2 SUMMARY ..... 5-8

5.3 APPARENT SENSITIVITY OF CLASSIFICATION RESULTS TO THE CHOICE OF TRAINING FIELD ..... 5-8

Section Page


5.3.1 DEFINITION.....5-8

5.3.2 METHOD USED TO ASSESS PROBLEM.....5-9

5.3.3 SUMMARY.....5-11

5.4 CONVERGENCE OF THE APPROXIMATING CURVE TO THE TRAINING FIELD DATA.....5-12

5.4.1 METHOD USED TO ASSESS PROBLEM.....5-12

5.4.2 SUMMARY.....5-12

5.5 THE EFFECTIVENESS OF THE GROWTH CYCLE ADJUSTMENT AND THE ADEQUACY OF THE RANGE SPECIFIED FOR VARIATION IN PLANTING DATE.....5-13

5.5.1 METHOD USED TO ASSESS PROBLEM.....5-13

5.5.2 SUMMARY.....5-14

5.6 IDENTIFICATION OF REASONS FOR GROUND-TRUTH IDENTIFIED CORN FAILING TO BE CLASSIFIED AS CORN.....5-14

5.6.1 METHODS USED TO ASSESS PROBLEM.....5-14

5.6.2 SUMMARY.....5-14

6. RESULTS.....6-1

7. RECOMMENDATIONS.....7-1

8. EVALUATION.....8-1

9. REFERENCES.....9-1

## TABLES


Table | Page

1 BASIC SITE DATA SET ..... 2-2

2 A COMPARISON OF FIVE- AND FOUR-ACQUISITION CLASSIFICATION ..... 5-3

3 SITES AND ACQUISITION SETS ..... 5-10

4 TRAINING FIELD SENSITIVITY ..... 5-10

5 CLASSIFICATION RESULTS ..... 6-2

# 1. INTRODUCTION

For a previous study (ref. 1), classification maps of corn or noncorn for 26 Large Area Crop Inventory Experiment (LACIE) corn and soybean sites were produced by using the corp profile classification method developed by Dr. G. Badhwar of the National Aeronautics and Space Administration (NASA) in Houston, Texas (refs. 2 and 3). This classification method incorporates the effects of crop-emergence date distribution and bases classification on the temporal spectral profile of the crop of interest. Spectral data is generated by the Landsat multispectral scanner (MSS); the crop profile is established for each site by local training of one input training field. Generation of results for so large a data set confirmed (1) that this classification method can be applied successfully to varying acquisition data distributions and training field signatures, and (2) that a method for producing an acceptable classification can be defined. Also, the accuracy of the classification results, when compared with the field-observation-based Accuracy Assessment (AA) digitized ground-truth maps, indicated that a more detailed analysis of results and evaluation of the Badhwar method should be done as an extension of the original work.

In this report, the results of an expanded study into the classification of corn using the Badhwar technique is documented. The purpose of this expanded study is (1) to define the probable incidence and impact of the potential problem areas indicated by the first study, and (2) to generate detailed information for use in further evaluation of this classification technique.

The data set, section 2, is the site data set used in the previous study with nine additional corn and soybean sites for which the Accuracy Assessment digitized ground truth became available. The software used for this study differed from that used for the first study. Specifically, the image data was not edited by the use of SCREEN (ref. 4) before classification, although SCREEN was applied to the training field data. In addition, a modification was made in the formula for calculation of the chi-square threshold values used in classification. Software programs are presented in section 3. Section 4 outlines the procedure used to fulfill the purposes of this study. Potential

problem areas are discussed in section 5 with an estimate of probable incidence and impact on classification results. Selected results of the classifications generated for this study are summarized in section 6. Recommendations for further application of this classification method are given in section 7. A brief evaluation of the Badhwar classification method applied to corn is given in section 8.

# 2. DATA SET

Thirty-four corn and soybean segments were classified using acquisitions for which the acquisition distribution over the corn growing season was adequate for crop profile definition in each MSS channel. The full segment (22 932 pixels) was classified.

## 2.1 SITE SELECTION

Geographical distribution of the segments used for this study is illustrated in figure 1. Twenty-six segments in the United States Corn Belt were successfully processed for evaluation as an aid to quality assurance for the Accuracy Assessment digitized ground-truth inventory maps (ref. 5). These segments were augmented by nine corn and soybean segments for which the digitized ground-truth maps recently became available. One segment in the data set (segment 123) could not be processed using the current software. The basic site data set used for this study is the 34 segments listed in table 1. This table includes LACIE sample segment numbers, locations, AA ground-truth percentages of corn, soybeans, sunflower, sorghum, and a comments section which identifies scene components (other than corn, soybeans, sunflowers, and sorghum) which comprise more than 10 percent of the scene. Scene percentages are computed only over the area of the segment identified by ground truth.

## 2.2 ACQUISITION SELECTION

The Badhwar classification method reported in this document requires four or five acquisitions in the postemergence to preharvest growth stages of corn. Five-acquisition classifications were made for as many segments as possible so that results could be compared with classifications generated on a four-acquisition subset. This dictated some relaxation of the criteria established in the first study, although generally these criteria were followed:

a. Acquisition distribution must be adequate to define a crop profile.

b. Acquisitions on cloudy or hazy days should be avoided. However, the effect of clouds or haze on crop profiles may be slight, and final judgment of acquisition usability was made with reference to graphs of the proposed training fields over the available acquisitions.

## TABLE 1.- BASIC SITE DATA SET

| Segment | County, state | Corn percentage | Soybean percentage | Sunflower percentage | Sorghum percentage | Comments |
|---|---|---|---|---|---|---|
| 127 | Montgomery, Ind. | 49.4 | 30.6 | 0.2 | - | - |
| 133 | Whitley, Ind. | 31.1 | 17.2 | - | - | 16% trees |
| 135 | Chickasaw, Iowa | 38.3 | 24.2 | - | - | - |
| 141 | Madison, Iowa | 24.1 | 18.9 | - | 0.1 | 26% pasture |
| 144 | Wapello, Iowa | 19.4 | 20.1 | - | 0.04 | 31% trees |
| 202 | Atchison, Mo. | 23.5 | 32.6 | - | 1.3 | 22% pasture |
| 205 | Clark, Mo. | 17.2 | 46.5 | - | 0.02 | 10% trees |
| 209 | Gentry, Mo. | 8.3 | 21.1 | - | 1.6 | 50% pasture and trees |
| 211 | Grundy, Mo. | 6.6 | 22.4 | - | 4.9 | 51% hay, trees, and pasture |
| 216 | Mercer, Mo. | 6.6 | 19.2 | - | 1.3 | 59% hay, trees, and pasture |
| 241 | Deuel, S. Dak. | 25.6 | 5.8 | 1.2 | 0.7 | 21% pasture |
| 800 | Clinton, Iowa | 53.8 | 27.3 | - | 0.02 | - |
| 804 | Marshall, Iowa | 46.0 | 28.6 | - | - | - |
| 809 | Ogle, Ill. | 53.2 | 12.7 | - | - | - |
| 824 | Iroquois, Ill. | 49.8 | 42.0 | - | 0.2 | - |
| 832 | Adams, Ind. | 21.6 | 39.0 | - | 0.1 | - |
| 837 | Benton, Ind. | 43.1 | 36.8 | - | - | - |
| 842 | Henry, Ind. | 42.6 | 28.5 | - | - | - |
| 843 | Henry, Ind. | 32.3 | 31.2 | - | - | 11% pasture |
| 852 | Randolph, Ind. | 27.0 | 30.7 | - | - | 22% pasture and trees |
| 853 | Randolph, Ind. | 34.9 | 30.3 | - | - | - |
| 854 | Tippecanoe, Ind. | 49.2 | 41.2 | - | 0.2 | - |
| 860 | Wells, Ind. | 28.2 | 31.3 | - | - | 13% nonagricultural |
| 864 | Crawford, Iowa | 45.2 | 11.7 | - | - | 17% pasture |
| 865 | Crawford, Iowa | 33.1 | 14.2 | - | 1.3 | 22% pasture |
| 877 | Ida, Iowa | 38.3 | 19.7 | - | - | 11% pasture (20% ground truth not identified) |
| 878 | Kossuth, Iowa | 43.0 | 42.5 | - | - | - |
| 880 | Monona, Iowa | 44.6 | 37.8 | - | 0.7 | - |
| 881 | Monona, Iowa | 43.5 | 7.9 | - | 0.1 | 22% pasture |
| 882 | Palo Alto, Iowa | 42.9 | 38.9 | - | 0.1 | - |
| 883 | Palo Alto, Iowa | 29.6 | 32.0 | - | 0.04 | 11% pasture (12% ground truth not identified) |
| 886 | Pottawatomie, Iowa | 46.8 | 25.5 | - | 0.2 | - |
| 891 | Shelby, Iowa | 46.4 | 16.8 | - | 0.03 | 13% oats |
| 892 | Shelby, Iowa | 30.1 | 14.3 | - | 0.2 | - |

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

Figure 1.- Geographical distribution of segments.

c. Acquisitions should occur in the growing season of the crop, although considerable variation seems to be acceptable in the definitions of postemergence and preharvest.

## 2.3 TRAINING FIELD SELECTION

Ground-truth cornfields were input to the program as training fields. For 17 of the segments, training fields were selected by Gary Gutcheski of NASA; training fields for the remaining sites were selected by the author. If available, training fields were used as defined for the first study.

Candidate training fields were selected on the production film converter (PFC) imagery products using these criteria:

a. Minimum training field size is 20 pixels.

b. Border and edge pixels are excluded from field delineation.

c. Fields are to be free of roads, drainage patterns, etc., if possible.

d. Field should be free of clouds and haze on the acquisitions used.

Since this study was directed toward problem definition, some of the selection criteria defined for the first study were relaxed; fields with unusual signatures were used as the basis for classification as well as fields which might be preemergent or harvested on the selected acquisitions. As expected, crop profile definition based on the training fields differed over a given set of acquisitions. The effect of this on classification results was assessed in connection with the discussion in section 5. Each segment had a maximum of five training fields defined; overall, each of 111 training fields was used as a base for classification.

# 3. SOFTWARE

All data processing used to generate the classifications was done on the programmed data processer model 11/45 (PDP-11/45). Several software programs were used, and these programs are described in this section. The software programs described in this section have also been used successfully for the classification of spring wheat (ref. 6).

## 3.1 IMUNLD2A

IMUNLD2A takes an image unload tape generated on the Earth Resources Interactive Processing System (ERIPS), edits it using SCREEN (ref. 4), adjusts the Landsat-3 acquisitions into a data range comparable to the data range of Landsat-2 acquisitions using the Wehmanen multiplicative factors (ref. 7), 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. 8) 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 in each channel with one standard deviation envelope; a curve is fitted through the mean values. Eight graphs, two for each MSS channel, are produced for a field over a set of acquisitions. Graphs are displayed on the Image-100 Tektronix screen, and hardcopies are made automatically. Listed on the first plot are the segment number, the acquisitions used, the coordinates of the field, the channel number, number of pixels in the field, and the mean and standard deviation on each acquisition. Listed on the second plot are 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), and the chi-square value for the fit of the approximating curve to the field data.

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

### 3.3 CLASFYT

CLASFYT (ref. 9) 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[1], 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, and channel 1 in succession, and to 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 corp profile (refs. 2 and 3). Accepted pixels are labeled corn, and rejected pixels are labeled noncorn.

- 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 corn or noncorn for each pixel in the segment; lineprinter sheet summarizing the following:
  a. The acquisitions used.
  b. The training field coordinates and the number of pixels in the field.
  c. The mean and standard deviation for each channel and each acquisition (field averages).

---

[1] As each image was unloaded from an ERIPS image unload tape onto a disk for processing on the PDP 11/45, it was edited using the ERIM program SCREEN, a procedure for automatically detecting garbled data, clouds, snow, cloud shadows, and water in MSS data. Pixels in the training field that failed to pass this edit step were excluded from processing and did not affect the crop profiles. However, screened pixels were restored before classification of the segment, so all 22 932 pixels are designated as corn or noncorn.

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

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

f. The 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 and removed from consideration as corn for exceeding the chi-square threshold in each channel.

i. The final numerical results: the number of pixels classified as corn, the number of pixels screened (always "0" in this study), and the number of pixels rejected as corn.

### 3.4 MISMAP

MISMAP (ref. 9) compares the classification file produced by CLASFYT with the AA digitized ground-truth inventory map for the segment. A numerical scene summary is given in confusion matrix form. A lineprinter map is generated with this code:

a. Ground-truth corn classified as corn appears as C.

b. Ground-truth noncorn rejected as corn is left blank.

c. Ground-truth noncorn classified as corn appears as +.

d. Ground-truth corn rejected as corn appears as -.

e. Pixels for which ground truth is not available but which are classified as corn appear as $.

f. Pixels for which ground truth is not available but which are rejected as corn appear as %.

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 CLASFYT 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.

### 3.5 MISMAP1

MISMAP1 was developed for use in this expanded study of the application of Badhwar's classification of corn. This program can be used to compare two classification files created by CLASFYT. A numerical comparison is given in confusion matrix form. A lineprinter map is generated with this code:

- Pixels classified as corn on both files appear as C.
- Pixels rejected as corn on both files are blank.
- Pixels classified as corn on one file but rejected as corn on the other appear either as + or -, depending upon the order of entry of the file names.

MISMAP1 was used to compare the classification files produced by two different training fields by using the same set of acquisitions, and it was also used to compare a five-acquisition classification with the four-acquisition classification that was generated by the same training field.

- Input: two classification files produced from CLASFYT.
- Output: full-scene lineprinter map comparing the classification maps and a confusion matrix numerical summary of results.

### 3.6 TAPEOUT

TAPEOUT (ref. 10) reads the data files produced by CLASFYT and creates universal format tapes. Black and white film product classification maps are produced on the PFC from these tapes. The scale used is the same as that of the PFC color imagery. The classification maps produced for this study are available for reference.

- Input: classification file from CLASFYT
- Output: black and white full-scene classification map on film.

## 4. PROCEDURE

Detailed description of the modeling used in Badhwar classification is given in "A Semi-Automatic Technique for Multitemporal Classification of a Given Crop" by G. Badhwar (ref. 2) and "Crop Emergence Data Determination from Spectral Data" by G. Badhwar (ref. 3). Implementation methods are explained in "Implementation of Badhwar Classification of Corn/Soybean Segments" by W. Austin (ref. 1). The software programs utilized have been presented in section 3. The procedure which was used to integrate these factors is outlined below.

a. LACIE corn/soybean segment images were unloaded to a PDP-11/45 disc using IMUNLD2A.

b. Acquisition sets were chosen and candidate training fields of corn (two to five per segment) were selected by an analyst using the PFC imagery products.

c. Candidate training fields were graphed over five[2] and over four acquisitions using IMAPLT.

d. Referring to the IMAPLT graphs, the analyst confirmed that (1) the acquisition set was adequate for crop profile definition, and (2) candidate training fields were larger than 20 pixels and were of reasonable homogeneity. For this study, no judgment of acquisition coverage relative to the training field growth cycle was made, although this was noted. A consistent set of acquisitions was used to classify all training fields so that the effect of using "early" or "late" (relative to the available acquisitions) fields for classification could be examined. Also, no restriction was placed on the chi-square measure of the fit of the approximating curve to the training data because the usefulness of this chi-square measure as a predictor of classification results was to be observed during this study.

e. Each candidate training field larger than 20 pixels and free of haze and clouds on the acquisition set was used as a base for a classification on

---

[2] Sample segments 800, 832, and 878 did not have five acceptable acquisitions available.

five acquisitions and on a four-acquisition subset of this. Acquisition sets usually conformed to those reported in reference 1. The restoration of pixels removed by SCREEN to the images before classification increased the available acquisition choice (some acquisitions were avoided in the first study because of excessive screening out of the pixels), so some differences do exist. In addition, some experimentation was done using early and postharvest dates. Training field coordinates and the selected acquisitons were put in a data file and processing was done "batch mode." Two hundred and twenty-five classifications were processed.

f. Each classification was compared with the digitized ground truth, using the software program MISMAP.

g. Classification tables were compiled for each segment. These tables provided the detailed information required for evaluation of this classification technique. For each training field and each acquisition set, this information was recorded:

   1. The number of pixels classified as corn.

   2. The ground-truth corn classified as corn (segment percentage, AA pure pixels only and all pixels).

   3. The ratio of ground-truth corn rejected as corn to the ground-truth noncorn classified as corn (in segment percentage, AA pure pixels only and all pixels).

   4. The final values of the constants (A, $\alpha$, $\beta$) used in the functions which define the crop profile in each channel.

$$P_v(t) = At^{\alpha}\exp(-\beta t^2)$$

   where

   $P_v(t)$ is the reflectance at time (t)

   5. The chi-square measure of curve fit to the training field data for each channel profile.

6. The chi-square cutoff threshold for each channel used for rejection of a pixel as corn with the number of pixels rejected by profile comparison in that channel.

7. Comments relevant to the classifications.

h. From the tables, the best available set of five and of four acquisitions and the training field were selected. These sets for the 34 segments are listed in table 2, section 5.1. A preferred acquisition set for each segment was selected.

i. A spot check comparison of disagreement was done on the MISMAP for the preferred classification of each segment. The largest percentage of classification disagreement with the ground truth is border and edge pixels; wherever an area was in disagreement, this was examined. This analysis was done to evaluate the performance of the program and to aid in problem definition.

j. The software program MISMAP1 was used to compare five- and four-acquisition classifications based on the same training field, and results were incorporated into section 5.1. MISMAP1 was also used to compare the preferred classification to the classifications based on other training fields over the same acquisitions as part of the training field sensitivity problem definition (section 5.3).

# 5. DISCUSSION OF POTENTIAL PROBLEM AREAS

Several potential problem areas were identified in the study documented in reference 1. Two problems were corrected before processing was done for this study.

1. Routine use of SCREEN caused excessive editing of the images and reduction of the number of pixels classified. For this study, pixels removed by SCREEN were restored before classification; hence, a classification was made for 22 932 pixels in every segment.
2. The chi-square threshold values in the individual channels are dependent upon the training field data. In the first study, if the training field data were approximated very well by the curve, an underclassification was likely to result. For this study, minimum threshold values were raised and the scaling factor in the formula used to calculate the chi-square cutoff was changed.

Problem areas discussed below are these:

1. The effectiveness of four-acquisition classification compared to classification on five acquisitions.
2. Inadequate acquisition coverage.
3. An apparent sensitivity of classification results to the choice of training field.
4. The convergence of the approximating curve to the training data.
5. The effectiveness of the growth cycle adjustment and the adequacy of the range specified for variation in planting data.
6. Identification of reasons for ground-truth corn failing to be classified as corn.

## 5.1 THE EFFECTIVENESS OF FOUR-ACQUISITION CLASSIFICATION COMPARED TO FIVE-ACQUISITION CLASSIFICATION

### 5.1.1 METHOD USED TO ASSESS PROBLEM

Classifications based on 102 training fields using five and four acquisitions were generated for this comparison. Selecting comparisons which are significant requires caution. If, to have five acquisitions, an acquisition of poor

data quality or a harvest date must be included, then classification on four acquisitions will be more accurate. This situation was checked using the software program MISMAP1; differences in classifications on five-acquisition and four-acquisition subsets with the same training field usually corresponded quite well to clouds or to harvested fields. If, however, five acquisitions of good data quality are available, the potential exists for better crop separability. Indeed, if a date of significance in crop separability is excluded, crop separability is lost for the comparable four-acquisition classification.

In the first study, five-acquisition classification was preferred; four-acquisition classification was done only if five acquisitions were not available or if no training field met the stated criteria using five acquisitions. Seventeen segments were classified on five acquisitions, and eight segments on four acquisitions. For this study, the number of acquisitions selected for the same data set was an analyst choice; six segments were selected on five-acquisition classification, and 19 were selected on four-acquisition classification. Accuracy was improved.

Table 2 provides a comparison of five-acquisition and four-acquisition classification. Eight of the segments listed (135, 144, 202, 800, 832, 878, 891, 892) are not suitable for use as a base in program development, but they are included so comparison could be made on segments with severe processing deficiencies. Percentages were calculated by the MISMAP program. This information is listed for each segment:

- Segment number and location.
- Training field coordinates (line, pixel) and training field sizes (pixels).
- A best-available five-acquisition set and a best-available four-acquisition subset of this; the preferred acquisition set is starred.
- Comparison of the classified proportion estimate to the ground-truth proportion estimate in segment percentage, AA pure pixels only. The proportion estimate does not include the areas which were not ground-truth identified.

# TABLE 2.- A COMPARISON OF FIVE- AND FOUR-ACQUISITION CLASSIFICATION

| Segment | Location (county, state) | Training field: Coordinates (line, pixel) | Training field: Size, pixels | Five-acquisition set / Four-acquisition set | Percentage estimation: corn/GT corn | GT** corn classified as noncorn/ GT noncorn classified as corn | Pixels classified as corn | A channels 1 | A channels 2 | A channels 3 | A channels 4 | α channels 1 | α channels 2 | α channels 3 | α channels 4 | β channels 1 | β channels 2 | β channels 3 | β channels 4 | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 127 | Montgomery, Ind. | (104,167)<br>(101,174)<br>(107,176)<br>(110,168) | 40 | 161,207,216,243,252 | 44.7/50.3 | 9.7/4.1 | 10 268 | 3.8 | 4.1 | 3.8 | 3.3 | -7.9 | -14.1 | 4.8 | 9.1 | -0.8 | -1.4 | 0.6 | 1.1 | Some haze on day 207; marginal acquisition history; day 161 is early; later final date would aid crop separability; 2% not identified ground truth; 19 287 AA pure pixels. |
| | | | | *161,216,243,252 | 40.7/50.3 | 5.2/6.6 | 11 383 | 3.8 | 4.2 | 3.7 | 3.3 | -11.3 | -17.8 | 3.4 | 8.9 | -1.2 | -1.9 | 0.4 | 1.0 | |
| 133 | Whitley, Ind. | (3,108)<br>(3,115)<br>(8,117)<br>(8,110) | 36 | *152,197,233,251,269 | 25.6/28.4 | 11.8/9.0 | 6 278 | 3.8 | 4.1 | 4.8 | 3.7 | -2.9 | -6.5 | 3.8 | 5.4 | -0.2 | -0.6 | 0.4 | 0.7 | Day 152 is early; harvest on day 260; small field size; 2.8% not identified ground truth; 18 786 AA pure pixels. |
| | | | | 152,197,233,251 | 20.7/28.4 | 14.2/6.5 | 5 879 | 3.8 | 4.1 | 4.0 | 3.7 | -4.3 | -6.9 | 2.8 | 4.9 | -0.6 | -0.6 | 0.3 | 0.6 | |
| 135 | Chickasaw, Iowa | (25,146)<br>(25,155)<br>(33,156)<br>(33,147) | 37 | 166,229,247,265,274 | 29.6/37.2 | 12.9/5.4 | 6 806 | 3.5 | 3.6 | 3.6 | 3.2 | -2.6 | -8.7 | 4.5 | 9.1 | -0.2 | -0.9 | 0.6 | 1.0 | Day 166 has clouds in upper left quadrant; some harvest evident on day 274; poor acquisition coverage; gap 166 to 229 in corn growing season; 2.9% not identified ground truth; 18 219 AA pure pixels. |
| | | | | *166,229,247,265 | 35.8/37.2 | 10.9/9.5 | 8 593 | 3.5 | 3.6 | 3.7 | 3.2 | -4.9 | -9.5 | 3.3 | 8.2 | -0.5 | -1.8 | 0.4 | 0.9 | |
| 141 | Madison, Iowa | (46,148)<br>(46,154)<br>(53,157)<br>(53,152) | 36 | 167,212,220,265,274 | 16.8/20.8 | 8.4/4.4 | 4 107 | 3.6 | 3.8 | 3.1 | 2.8 | -2.2 | -7.4 | 12.2 | 15.1 | -0.1 | -0.7 | 1.4 | 1.7 | Day 212 has three small clouds; some harvest evident on day 265, more on 274; .4% not identified ground truth; 19 363 AA pure pixels. |
| | | | | *167,212,220,265 | 20.8/20.8 | 5.4/4.6 | 4 913 | 3.7 | 3.8 | 3.1 | 2.7 | -3.8 | -8.1 | 12.7 | 16.3 | -0.2 | -0.8 | 1.5 | 1.9 | |
| 144 | Mapello, Iowa | (27,166)<br>(27,172)<br>(40,178)<br>(40,171) | 3 | 165,219,238,246,264 | 8.4/16.5 | 9.2/1.8 | 2 278 | 3.9 | 4.2 | 3.7 | 3.4 | -3.8 | -8.5 | 5.9 | 10.4 | -0.2 | -0.7 | 0.8 | 1.3 | Poor acquisition quality; 3.5% not identified ground truth; 19 384 AA pure pixels. |
| | | | | *165,219,246,264 | 13.2/16.5 | 6.5/3.1 | 3 483 | 3.9 | 4.1 | 3.7 | 3.4 | -3.7 | -9.4 | 5.8 | 10.1 | -0.2 | -0.8 | 0.8 | 1.2 | |
| 202 | Atchison, Mo. | (3,119)<br>(3,126)<br>(9,130)<br>(9,120) | 39 | 167,212,221,266,275 | 19.5/21.5 | 14.9/12.9 | 5 195 | 3.8 | 4.7 | 3.3 | 3.0 | -3.9 | -8.0 | 6.9 | 11.2 | -0.3 | -0.7 | 0.8 | 1.2 | Poor data quality; days 212 and 221 are cloudy, and day 275 is hazy; harvest is evident on days 266 and 275; 3.9% not identified ground truth; 19 439 AA pure pixels. |
| | | | | *167,212,221,266 | 18.6/21.5 | 13.9/11.0 | 4 794 | 3.8 | 3.9 | 3.1 | 2.9 | -2.6 | -6.1 | 8.8 | 12.3 | -0.2 | -0.5 | 0.9 | 1.4 | |
| 206 | Clark, Mo. | (40,169)<br>(40,173)<br>(55,175)<br>(55,172) | 31 | 155,164,219,246,272 | 11.3/15.0 | 10.8/7.1 | 3 405 | 3.6 | 3.7 | 3.7 | 3.4 | -3.8 | -7.5 | 7.3 | 8.9 | -0.3 | -0.8 | 0.9 | 1.1 | Day 164 is uncertain data quality, 8.3% not identified ground truth; 20 886 AA pure pixels. |
| | | | | *155,219,246,272 | 11.9/15.0 | 8.2/5.1 | 3 312 | 3.6 | 3.9 | 3.8 | 3.5 | -3.8 | -7.7 | 5.5 | 8.6 | -0.3 | -0.8 | 0.8 | 1.1 | |
| 209 | Gentry, Mo. | (11,187)<br>(11,193)<br>(13,193)<br>(13,187) | 21 | *167,221,247,266,274 | 7.9/6.5 | 2.7/4.1 | 2 820 | 3.7 | 3.9 | 3.3 | 2.8 | -0.8 | -4.6 | 8.7 | 10.9 | 0.0 | -0.3 | 0.9 | 1.2 | Image quality is good, day 274 corn is ripe; 5.8% not identified ground truth; 20 537 AA pure pixels. |
| | | | | 167,221,247,274 | 8.8/6.5 | 2.8/5.1 | 2 242 | 3.5 | 3.8 | 2.9 | 3.0 | -0.6 | -4.8 | 7.8 | 10.7 | 0.1 | -0.4 | 0.8 | 1.1 | |
| 211 | Grundy, Mo. | (50,122)<br>(50,128)<br>(55,128)<br>(57,122) | 13 | *166,220,247,265,274 | 6.7/5.3 | 2.9/6.3 | 2 317 | 3.8 | 3.9 | 3.4 | 3.1 | -3.4 | -9.0 | 6.5 | 9.3 | -0.2 | -0.7 | 0.8 | 1.1 | Poor data quality on days 166, 220, and 265, harvest is evident on days 265 and 274, 7.8% not identified ground truth, 26 457 AA pure pixels. |
| | | | | 166,220,247,274 | 16.0/5.3 | 2.5/13.2 | 4 049 | 3.8 | 3.9 | 3.4 | 3.1 | -3.6 | -8.7 | 6.5 | 9.2 | -0.2 | -0.7 | 0.8 | 1.0 | |

*Preferred acquisition set.
**Ground-truth

ORIGINAL PAGE IS OF POOR QUALITY

TABLE 2.- Continued.

| Segment | Location (county, state) | Training field: Coordinates (line, pixel) | Training field: Size, pixels | Five-acquisition set / Four-acquisition set | Percentage estimation: corn/GT corn | GT** corn classified as noncorn/ GT noncorn classified as corn | Pixels classified as corn | α channels 1 | α channels 2 | α channels 3 | α channels 4 | β channels 1 | β channels 2 | β channels 3 | β channels 4 | γ channels 1 | γ channels 2 | γ channels 3 | γ channels 4 | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 216 | Mercer, Mo. | (82,63)<br>(80,74)<br>(86,75)<br>(86,65) | 38 | 106,220,230,247,274 | 44.9/5.2 | 1.1/40.7 | 11 216 | 3.7 | 4.4 | 3.5 | 2.8 | -1.1 | -7.8 | 7.0 | 8.4 | 0.0 | -0.6 | 0.8 | 0.9 | Day 230 is hazy; harvest on day 274; 6.3% not identified ground truth; 21 846 AA pure pixels. |
| | | | | *106,220,247,274 | 25.5/5.2 | 1.4/21.7 | 6 708 | 3.8 | 4.5 | 3.0 | 2.4 | -2.0 | -9.9 | 8.6 | 10.8 | -0.1 | -0.6 | 0.9 | 1.1 | |
| 241 | Douel, S. Dak. | (6,176)<br>(6,185)<br>(10,185)<br>(10,177) | 36 | *187,206,224,233,251 | 18.1/22.6 | 9.5/5.0 | 4 496 | 3.6 | 4.1 | 2.8 | 2.7 | -6.3 | -14.0 | 9.5 | 11.8 | -0.7 | -1.5 | 1.0 | 1.3 | Day 251 is final acquisition; two small clouds on day 187; 4.2% not identified ground truth; 19 666 AA pure pixels. |
| | | | | 187,206,224,251 | 18.4/22.6 | 9.5/5.3 | 4 501 | 3.7 | 4.1 | 2.7 | 2.7 | -7.6 | -14.6 | 9.9 | 12.5 | -0.8 | -1.6 | 1.1 | 1.3 | |
| 880 | Clinton, Iowa | (60,14)<br>(60,34)<br>(70,34)<br>(70,15) | 38 | Not available | | | | | | | | | | | | | | | | Day 201 has some clouds and harvest is evident; only four acquisitions available; 1.9% not identified ground truth; 18 547 AA pure pixels. |
| | | | | *164,219,246,281 | 47.7/55.9 | 10.3/8.1 | 10 846 | 3.6 | 2.8 | 3.6 | 3.4 | -4.1 | -8.5 | 5.2 | 9.4 | -0.4 | -0.8 | 0.7 | 1.1 | |
| 884 | Marshall, Iowa | (24,132)<br>(23,141)<br>(30,143)<br>(31,133) | 38 | 166,229,247,265,274 | 34.6/46.6 | 14.6/2.6 | 7 366 | 2.9 | 3.4 | 3.1 | 2.7 | 0.0 | -4.6 | 12.3 | 14.6 | -0.2 | -0.4 | 1.5 | 1.6 | Underclassification using five acquisitions corresponds to area harvested on 274, not harvested on 265; gap from day 166 to day 229; 2.5% not identified ground truth; 12 504 AA pure pixels. |
| | | | | *166,229,247,265 | 46.7/46.6 | 6.7/6.8 | 10 579 | 3.0 | 3.4 | 3.4 | 2.8 | -0.7 | -6.6 | 12.4 | 11.7 | -0.3 | -0.7 | 1.5 | 1.3 | |
| 809 | Ogle, Ill. | (43,6)<br>(40,17)<br>(47,20)<br>(50,6) | 38 | 164,218,244,262,271 | 45.5/56.0 | 16.1/5.0 | 10 163 | 3.7 | 4.0 | 3.5 | 3.2 | -4.1 | -10.4 | 6.9 | 9.3 | -0.6 | -1.0 | 0.8 | 1.1 | Day 262 has some haze and clouds; 6% not identified ground truth; 18 553 AA pure pixels. |
| | | | | *164,218,244,271 | 49.4/55.0 | 12.2/5.0 | 10 967 | 3.6 | 4.0 | 3.5 | 3.2 | -3.3 | -10.5 | 6.4 | 8.9 | -0.2 | -1.0 | 0.8 | 1.0 | |
| 824 | Iroquois, Ill. | (5,107)<br>(5,118)<br>(10,118)<br>(10,107) | 37 | *163,217,235,243,262 | 51.6/50.3 | 7.2/8.7 | 12 030 | 3.4 | 3.5 | 2.6 | 2.9 | -1.3 | -6.5 | 11.0 | 17.3 | -0.1 | -0.7 | 1.5 | 2.0 | Marginal data quality; day 163 has a few clouds, days 217, 236, and 262 are hazy; 2.6% not identified ground truth; 19 400 AA pure pixels. |
| | | | | 163,217,243,262 | 55.7/50.3 | 6.2/12.0 | 13 246 | 3.4 | 3.5 | 3.0 | 2.8 | -2.2 | -7.6 | 12.0 | 17.2 | -0.2 | -0.8 | 1.4 | 2.0 | |
| 832 | Adams, Ind. | (91,143)<br>(91,153)<br>(94,153)<br>(94,143) | 39 | Not available | | | | | | | | | | | | | | | | Only four acquisitions available; gap between day 160 and day 232, day 151 is early and some harvest evident on 268, 1.7% not identified ground truth; 19 146 AA pure pixels. |
| | | | | *151,166,232,268 | 35.9/17.8 | 4.4/22.5 | 9 154 | 3.6 | 3.9 | 4.0 | 3.6 | -4.7 | -10.1 | 3.3 | 6.1 | -0.4 | -1.0 | 0.4 | 0.7 | |
| 837 | Benton, Ind. | (58,35)<br>(98,47)<br>(101,40)<br>(104,37) | 39 | 100,198,216,234,252 | 56.2/42.9 | 7.0/20.3 | 13 099 | 4.1 | 4.5 | 3.6 | 3.1 | -12.1 | -20.7 | 6.5 | 11.8 | -1.2 | -2.1 | 0.8 | 1.3 | Good coverage, few clouds on day 216, 1.7% not identified ground truth, 19 172 AA pure pixels. |
| | | | | *100,198,216,252 | 42.9/42.9 | 9.7/9.7 | 9 912 | 4.1 | 4.6 | 3.4 | 2.8 | -11.6 | -20.5 | 8.4 | 15.0 | -1.2 | -2.1 | 1.0 | 1.6 | |
| 842 | Henry, Ind. | (3,133)<br>(3,145)<br>(6,147)<br>(9,137) | 40 | *160,170,232,250,268 | 38.7/42.1 | 11.2/7.9 | 9 116 | 3.8 | 4.0 | 3.7 | 3.1 | -4.1 | -10.3 | 5.0 | 8.2 | -0.3 | -0.9 | 0.6 | 0.9 | Day 160 is early, signs of harvest on day 268, 2% not identified ground truth; 18 556 AA pure pixels. |
| | | | | 170,232,250,268 | 53.3/42.1 | 8.1/19.3 | 12 835 | 4.2 | 4.8 | 3.7 | 3.0 | -8.1 | -16.9 | 6.4 | 9.8 | -0.7 | -1.5 | 0.8 | 1.0 | |

* Preferred acquisition set.
** Ground truth



ORIGINAL PAGE IS
OF POOR QUALITY

## TABLE 2.- Continued.

| Segment | Location (county, state) | Training field: Coordinates (line, pixel) | Training field: Size, pixels | Five-acquisition set / Four-acquisition set | Percentage estimation: corn/GT corn | GT** corn classified as noncorn/ GT noncorn classified as corn | Pixels classified as corn | A channels 1 | A channels 2 | A channels 3 | A channels 4 | α channels 1 | α channels 2 | α channels 3 | α channels 4 | β channels 1 | β channels 2 | β channels 3 | β channels 4 | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 843 | Henry, Ind. | (13,121)<br>(13,123)<br>(22,127)<br>(22,125) | 22 | 170,197,233,251,260 | 61.6/29.8 | 2.9/34.8 | 14 694 | 3.8 | 4.4 | 3.3 | 2.8 | -4.9 | -14.3 | 6.8 | 11.7 | -0.4 | -1.3 | 0.7 | 1.2 | Good acquisitions, harvest evident on day 260; corn signatures tend to "green up" this date; .8% not identified ground truth; 38 829 AA pure pixels. |
| | | | | *170,197,233,251 | 38.2/29.8 | 6.4/14.9 | 9 252 | 4.6 | 5.2 | 3.2 | 2.4 | -15.1 | -24.6 | 8.5 | 16.2 | -1.5 | -2.5 | 0.9 | 1.6 | |
| 852 | Randolph, Ind. | (49,36)<br>(49,44)<br>(51,46)<br>(53,37) | 32 | 160,170,232,250,260 | 16.5/24.4 | 11.7/3.8 | 3 867 | 3.8 | 4.2 | 3.4 | 2.5 | -3.6 | -5.7 | 4.2 | 7.8 | -0.3 | -0.4 | 0.4 | 0.7 | Harvest on day 260; 1% not identified ground truth; 19 179 AA pure pixels. |
| | | | | *170,232,250,260 | 19.1/24.4 | 9.9/4.6 | 4 469 | 4.2 | 4.9 | 3.4 | 2.3 | -8.4 | -16.9 | 4.4 | 10.6 | -0.7 | -1.6 | 0.4 | 1.0 | |
| 853 | Randolph, Ind. | (35,68)<br>(34,72)<br>(41,77)<br>(43,71) | 32 | *160,170,232,250,260 | 24.3/32.8 | 15.4/6.9 | 5 589 | 3.6 | 3.7 | 3.6 | 3.1 | -3.8 | -8.7 | 3.3 | 7.5 | -0.3 | -0.9 | 0.3 | 0.8 | 1% not identified ground truth; 16 500 AA pure pixels. |
| | | | | 160,170,232,260 | 14.8/32.8 | 21.4/3.5 | 3 683 | 3.6 | 3.8 | 3.6 | 3.1 | -4.6 | -10.0 | 4.4 | 7.2 | -0.4 | -1.0 | 0.5 | 0.7 | |
| 854 | Tippecanoe, Ind. | (73,148)<br>(73,155)<br>(80,155)<br>(80,148) | 38 | 161,207,234,243,251 | 43.6/49.4 | 10.4/4.6 | 10 871 | 3.6 | 3.9 | 3.8 | 3.6 | -4.9 | -11.4 | 5.3 | 6.4 | -0.5 | -1.1 | 0.6 | 0.7 | Good acquisition coverage; .1% not identified ground truth; 19 422 AA pure pixels. |
| | | | | *161,207,234,251 | 49.1/49.4 | 7.9/7.6 | 11 435 | 3.7 | 3.9 | 3.6 | 3.5 | -4.8 | -11.9 | 5.7 | 6.4 | -0.4 | -1.2 | 0.7 | 0.7 | |
| 860 | Wells, Ind. | (91,61)<br>(91,66)<br>(95,67)<br>(95,62) | 27 | *160,197,232,251,260 | 26.6/25.9 | 7.9/8.6 | 6 049 | 3.6 | 3.9 | 3.7 | 3.3 | -3.1 | -8.2 | 4.7 | 8.0 | -0.2 | -0.7 | 0.5 | 0.9 | Corn still vigorous on day 260; 6.1% not identified ground truth; 38 554 AA pure pixels. |
| | | | | *160,197,232,251 | 21.9/25.9 | 9.2/5.3 | 5 586 | 3.7 | 4.0 | 3.6 | 3.1 | -6.1 | -11.5 | 6.5 | 10.8 | -0.6 | -1.2 | 0.8 | 1.1 | |
| 864 | Crawford, Iowa | (65,116)<br>(65,120)<br>(69,120)<br>(69,116) | 24 | 159,186,222,231,267 | 29.8/44.8 | 19.6/3.7 | 6 598 | 3.7 | 4.0 | 4.0 | 3.6 | -5.0 | -12.6 | 4.1 | 8.0 | -0.4 | -1.3 | 0.6 | 1.0 | Day 222 has clouds; harvest began on day 267; 1.8% not identified ground truth, 18 282 AA pure pixels. |
| | | | | *159,186,231,267 | 39.5/44.8 | 10.3/5.0 | 9 022 | 3.7 | 4.1 | 3.9 | 3.5 | -5.0 | -13.0 | 4.5 | 9.4 | -0.4 | -1.3 | 0.6 | 1.2 | |
| 865 | Crawford, Iowa | (7,71)<br>(7,79)<br>(11,79)<br>(11,71) | 38 | 160,186,231,249,267 | 16.2/30.3 | 16.4/2.3 | 4 667 | 3.7 | 4.0 | 3.4 | 3.2 | -4.6 | -11.0 | 8.4 | 12.2 | -0.4 | -1.1 | 1.0 | 1.4 | Haze in lower right corner on day 249, harvest began on day 267, 4.5% not identified ground truth, 38 260 AA pure pixels. |
| | | | | *160,186,231,257 | 25.8/30.3 | 9.4/4.9 | 6 471 | 3.7 | 4.0 | 3.4 | 3.2 | -5.0 | -12.1 | 8.3 | 12.2 | -0.4 | -1.2 | 1.0 | 1.4 | |
| 877 | Ida, Iowa | (70,114)<br>(69,123)<br>(75,125)<br>(77,117) | 39 | 150,186,222,231,267 | 32.8/39.5 | 12.9/5.5 | 9 030 | 3.4 | 3.5 | 3.9 | 3.5 | -1.3 | -6.7 | 5.5 | 9.1 | -0.0 | -0.7 | 0.8 | 1.1 | Day 150 is early and has some cloud and shadow, some harvest on day 267, 20.1% not identified ground truth, 17 632 AA pure pixels. |
| | | | | *186,222,231,267 | 34.7/39.5 | 12.1/7.4 | 9 713 | 3.4 | 3.5 | 3.9 | 3.5 | -0.1 | -6.5 | 4.8 | 9.1 | 0.1 | -0.6 | 0.7 | 1.1 | |
| 878 | Kossuth, Iowa | (9,140)<br>(9,144)<br>(19,151)<br>(19,144) | 40 | Not available | | | | | | | | | | | | | | | | Only 3 acquisitions in the growing season of corn, much harvest on day 293, .7% not identified ground truth, 19 326 AA pure pixels. |
| | | | | *186,221,266,293 | 31.5/42.3 | 17.9/7.0 | 7 667 | 3.3 | 2.9 | 4.0 | 3.7 | -1.2 | -1.0 | 3.8 | 5.8 | -0.1 | -0.2 | 0.5 | 0.7 | |

* Preferred acquisition set.
** Ground-truth



ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 2.- Concluded.

| Segment | Location (county, state) | Training field: Coordinates (line, pixel) | Training field: Size, pixels | Five-acquisition set / Four-acquisition set | Percentage estimation: corn/GF corn | GF** corn classified as noncorn/ GF noncorn classified as corn | Pixels classified as corn | A channels 1 | A channels 2 | A channels 3 | A channels 4 | α channels 1 | α channels 2 | α channels 3 | α channels 4 | β channels 1 | β channels 2 | β channels 3 | β channels 4 | Comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 880 | Monona, Iowa | (46,123)<br>(46,133)<br>(54,136)<br>(54,127) | 36 | 150,186,222,231,267 | 33.5/44.9 | 16.3/4.9 | 7 704 | 3.1 | 3.2 | 3.5 | 3.1 | -0.8 | -4.1 | 8.6 | 13.2 | -0.0 | -0.5 | 1.1 | 1.5 | Small scattered clouds on day 222, shortage of acquisitions; day 150 is early; 3% not identified ground truth; 19 263 AA pure pixels. |
| | | | | *186,222,231,267 | 34.8/44.9 | 13.3/3.2 | 7 867 | 3.3 | 3.3 | 3.6 | 3.1 | -1.3 | -4.7 | 7.8 | 12.1 | -0.1 | -0.5 | 1.0 | 1.4 | |
| 881 | Monona, Iowa | (13,81)<br>(12,90)<br>(16,91)<br>(17,82) | 39 | *159,186,222,231,267 | 36.7/42.9 | 9.3/3.1 | 8 385 | 3.7 | 4.1 | 3.9 | 3.5 | -4.6 | -13.4 | 5.3 | 9.5 | -0.4 | -1.3 | 0.8 | 1.2 | Small cloud and shadow on day 150; little harvest on day 267; 1.3% not identified ground truth; 18 831 AA pure pixels. |
| | | | | 186,222,231,267 | 46.3/42.9 | 6.4/9.8 | 10 921 | 3.6 | 4.1 | 4.0 | 3.5 | -3.8 | -13.5 | 5.0 | 10.1 | -0.3 | -1.3 | 0.7 | 1.2 | |
| 882 | Palo Alto, Iowa | (77,77)<br>(76,84)<br>(81,87)<br>(82,80) | 36 | *159,186,222,231,267 | 40.3/42.8 | 6.6/3.5 | 9 695 | 3.4 | 3.6 | 3.7 | 3.5 | -3.7 | -9.5 | 4.6 | 7.8 | -0.3 | -1.0 | 0.6 | 1.0 | Good acquisition coverage; harvest on day 267; 5.3% not identified ground truth; 19 368 AA pure pixels. |
| | | | | 159,122,231,267 | 45.8/42.8 | 4.5/7.5 | 10 976 | 3.5 | 3.7 | 3.4 | 3.1 | -3.7 | -9.9 | 6.3 | 9.4 | -0.3 | -1.0 | 0.7 | 1.1 | |
| 883 | Palo Alto, Iowa | (22,109)<br>(22,118)<br>(29,120)<br>(29,111) | 39 | *186,204,213,221,267 | 30.9/28.4 | 6.2/0.7 | 8 060 | 3.1 | 3.1 | 3.0 | 3.6 | -0.0 | -4.7 | 5.2 | 7.9 | 0.0 | -0.6 | 0.7 | 1.5 | Small cloud on day 204, slight haze on day 213; harvest began on day 267; day 186 is late for first date; 30.3% not identified ground truth; 16 936 AA pure pixels. |
| | | | | 186,204,221,267 | 35.5/28.4 | 4.7/11.0 | 9 497 | 3.6 | 3.4 | 3.0 | 3.6 | -4.0 | -7.6 | 4.9 | 8.4 | -0.4 | -0.8 | 0.7 | 1.0 | |
| 886 | Pottawattamie, Iowa | (102,77)<br>(102,86)<br>(110,91)<br>(110,81) | 37 | 167,186,204,231,249 | 25.4/47.1 | 23.6/1.9 | 5 870 | 4.0 | 4.4 | 3.7 | 2.8 | -12.4 | -21.3 | 4.5 | 11.1 | -1.3 | -2.2 | 0.5 | 1.2 | Data quality is poor; days 167, 186, and 249 are hazy; day 204 has scattered small clouds; 3.5% not identified ground truth; 18 762 AA pure pixels. |
| | | | | *167,204,231,249 | 37.2/47.1 | 13.6/3.7 | 8 426 | 4.0 | 4.4 | 3.7 | 2.9 | -13.6 | -21.9 | 4.2 | 11.3 | -1.4 | -2.3 | 0.5 | 1.3 | |
| 891 | Shelby, Iowa | (90,5)<br>(90,13)<br>(97,13)<br>(97,5) | 37 | 168,186,204,250,267 | 35.7/47.0 | 18.0/7.5 | 8 252 | 3.7 | 4.0 | 3.5 | 3.0 | -6.5 | -14.2 | 8.5 | 13.1 | -0.6 | -1.4 | 1.0 | 1.5 | Poor data quality during corn growing season; clouds on days 168 and 204, and clouds and haze on day 250, 2.9% not identified ground truth; 17 378 AA pure pixels. |
| | | | | *168,186,204,267 | 38.9/47.0 | 16.6/8.5 | 8 795 | 3.7 | 4.0 | 3.5 | 3.0 | -6.4 | -14.3 | 8.5 | 12.6 | -0.6 | -1.4 | 1.0 | 1.4 | |
| 892 | Shelby, Iowa | (16,50)<br>(15,59)<br>(19,60)<br>(19,53) | 34 | 167,204,221,231,266 | 43.2/51.5 | 10.9/2.5 | 9 302 | 3.0 | 4.3 | 3.0 | 3.3 | -6.5 | -14.3 | 5.4 | 10.2 | -0.6 | -1.4 | 0.7 | 1.2 | Inadequate acquisitions of good data quality; day 167 has shadow, and day 204 has small clouds, 3.7% not identified ground truth; 17 830 AA pure pixels. |
| | | | | *167,204,221,266 | 50.2/51.5 | 7.3/5.9 | 11 242 | 3.0 | 4.3 | 3.0 | 3.4 | -6.0 | -14.0 | 5.5 | 10.0 | -0.6 | -1.4 | 0.7 | 1.2 | |

* Preferred acquisition set.
** Ground-truth



ORIGINAL PAGE IS
OF POOR QUALITY

- Ratio of the ground-truth corn classified noncorn to the ground-truth noncorn classified corn. AA pure pixels only, segment percentages.
- The total number of pixels classified as corn.
- The final value of the constants (A, $\alpha$, $\beta$) in each channel used to model the crop profile. This provides a comparison of the curve approximations of the data.
- Comments, including the percentage of the segment not identified ground truth, and the number of AA pure pixels upon which calculation of the scene percentages is based.

### 5.1.2 SUMMARY

If the crop profile, adequately representative of the data and sufficiently well defined to provide crop separabililty, can be generated on four acquisitions, classification results seem to be equally good using a set of four acquisitions or five acquisitions. The distribution of acquisitions relative to the crop growth stage of the training field and the data quality of the acquisitions is more important than the number of acquisitions used.

Obtaining Landsat acquisitions of good data quality is a problem. A requirement of five acquisitions for application of Badhwar classification would restrict the number of segments that could be classified. Comparable results using fewer acquisitions will permit the classification of more segments.

## 5.2 INADEQUATE ACQUISITION COVERAGE

### 5.2.1 METHOD USED TO ASSESS PROBLEM

Inadequate acquisition coverage was assessed by effect. Thirty-eight corn and soybean segments were examined for application of Badhwar classification; two were rejected for inadequate acquisition coverage, one for an allied problem (inadequate acquisitions of usable data quality), and one more for processing problems related to data quality. Of the remaining 34 segments used in this study, eight are not recommended for continued use in program development because of inadequate acquisitions of acceptable data quality and three segments did not have five acceptable acquisitions available for processing.

Adequate acquisition coverage of good data quality is a continuing and severe problem in the use of Landsat data. Maximum coverage is every 9 days (using data from both satellites) and much of this is so adversely affected by atmospheric conditions that it cannot be used in classification.

Badhwar classification can currently be applied to segments where there are a minimum of four acquisitions, including one in the green-up and one in the senescent growth stage of the crop; this is comparable to the requirements of other classification methods. Badhwar classification tends to be successful even with marginal or poor acquisition coverage. For the eight segments listed in section 5.1 as rejected for use in future research, misclassification averages 21.3 percent for an approximate accuracy of 78.7 percent. From section 5.1, it can be inferred that improved results were generated for 11 segments where acquisition requirements were relaxed from five to four. Similarly, if requirements could be reduced even further to the absolute minimum for curve definition--three acquisitons--improved results would be expected for 10 additional segments.

### 5.2.2 SUMMARY

Inadequate acquisition coverage of good data quality exists as an important limitation for any classification method which uses Landsat data. The Badhwar method is not affected more severely than other methods; accuracy appears to remain satisfactory even when using very poor acquisition sets. However, if the program could be modified to relax restrictions on the number of acquisitions (and their positions relative to the crop growth stage), classification results could be expected to improve since more stringent requirements on data quality could be maintained.

## 5.3 APPARENT SENSITIVITY OF CLASSIFICATION RESULTS TO THE CHOICE OF TRAINING FIELD

### 5.3.1 DEFINITION

Training field sensitivity refers to a situation where classifications based on different training fields differ by more than 2000 pixels classified as corn, although the training field signatures are similar, and similar classification results would be expected.

### 5.3.2 METHOD USED TO ASSESS PROBLEM

Nine sites were selected as suitable for investigation of this potential problem. These sites had good acquisition coverage and included sites for which alternate training field selection had been used effectively as a rework tool during the first study. Segment 809 was included as an example of training field insensitivity. The 10 sites with acquisition sets are listed on table 3.

The procedure used was described in section 4, with some additional steps.

- General data quality was noted for each of the acquisitions used. Data quality could cause classification differences and would preclude attributing these differences to training field sensitivity.
- Training field signatures on the sequence of PFC imagery were compared in detail. Graphs of the training fields were compared in each channel, and differences (data dispersion, convergence of the curve to the data, etc.) were noted. These differences would be expected to produce differences in classification results.
- Training fields were used as test fields if they were not used to define the crop profile and the classified results of these fields were noted. This proved to be an effective way to assess overall classification results.
- MISMAP1 was used to compare the classifications based on different training fields with classifications based on the training fields listed on table 2. This comparison map of the scene plus the confusion matrix numerical summary provided both an area and a statistical comparison of differences.

Of the 10 segments, the problem could be defined on the segments listed in table 4 for the stated acquisition sets and as affecting the listed training fields. The ratio of the number of training field combinations exhibiting sensitivity to the number of combinations examined for the segment is also given.

Thus, of the 53 combinations examined on these segments, the defined problem could be isolated only 17 times. This does not mean that training field

TABLE 3.- SITES AND ACQUISITION SETS.

| Segment | County, state | Five-acquisition set | Four-acquisition set |
|---|---|---|---|
| 127 | Montgomery, Ind. | 161, 207, 216, 243, 252 | 161, 216, 243, 252 |
| 141 | Madison, Iowa | 167, 212, 220, 265, 274 | 167, 212, 220, 265 |
| 809 | Ogle, Ill. | 164, 218, 244, 262, 271 | 164, 218, 244, 271 |
| 860 | Wells, Ind. | 160, 197, 232, 251, 268 | 160, 197, 232, 251 |
| 864 | Crawford, Iowa | 159, 186, 222, 231, 267 | 159, 186, 231, 267 |
| 865 | Crawford, Iowa | 168, 186, 231, 249, 267 | 168, 186, 231, 267 |
| 877 | Ida, Iowa | 150, 186, 222, 231, 267 | 186, 222, 231, 267 |
| 880 | Monona, Iowa | 150, 186, 222, 231, 267 | 186, 222, 231, 267 |
| 881 | Monona, Iowa | 159, 186, 222, 231, 267 | 186, 222, 231, 267 |
| 882 | Palo Alto, Iowa | 159, 186, 222, 231, 267 | 159, 222, 231, 267 |

TABLE 4.- TRAINING FIELD SENSITIVITY

| Segment | Acquisition set | Training fields affected | Ratio |
|---|---|---|---|
| 127 | (161, 207, 216, 254, 252)<br>(161, 216, 254, 252) | #1, #2, #4, #5 | 3/20 |
| 141 | (167, 212, 220, 265, 274)<br>(167, 212, 220, 265 | #1, #4 | 2/6 |
| 864 | (159, 186, 231, 267) | #2, #3, #4, #5 | 5/11 |
| 865 | (168, 186, 231, 267) | #1, #2, #3, #4, #5 | 4/13 |
| 881 | (159, 186, 222, 231, 267) | #1, #2, #3 | 3/3 |

sensitivity does not exist, only that it is very difficult to isolate. An examination of 281 combinations over the 34 segments located 18 potential incidences of training field sensitivity, only one more than defined on the 10 segments.

For this study, training fields which were obviously affected by haze or cloud cover on any one of the acquisitions were not used as a base for classification. However, classifications were performed on training fields for which the acquisition distribution relative to the field crop development was poor (or other problems existed) and on training fields of varying signatures on the imagery and as graphed. Acquisition sets which included hazy or cloudy acquisitions were used for classification. However, these factors were considered legitimate bases for classification result differences, and training field sensitivity could not be defined as the source of differences in results when these factors were present. In summary, nine of the 34 sites were selected because they indicated a sensitivity of classification results to training field choice existed; sensitivity of results to training field choice is defined only if results are expected to be similar based on examination of training field signature.

### 5.3.3 SUMMARY

Training field sensitivity cannot be isolated and defined using a multipurpose data set such as this one. For this data set, there is too much variation in training field signature. Training fields which have similar signatures should be defined for a set of sites with good acquisition coverage. These fields should also conform to some criteria for the fit of the approximating curve to the training data. Then, if training field choice affects classification results, this will be significant. Sample segment 881 is an example of a segment suitable for use in such a study.

The incidence of this problem is indeterminate, but probably acceptable since the impact is minor. An unacceptable level of underclassification or overclassification is clearly indicated by a loss of field pattern and an increase in scattering on the classification map. Test fields could be defined as an aid to

classification evaluation; a film product classification map would facilitate comparison with the PFC imagery.

Alternate training field selection as a rework technique is effective, clear, and simple. Change of training data to improve classification results has been used with success during all of the LACIE project.

## 5.4 CONVERGENCE OF THE APPROXIMATING CURVE TO THE TRAINING FIELD DATA

### 5.4.1 METHOD USED TO ASSESS PROBLEM

All training fields for this study were graphed at least twice using IMAPLT. Examination of these graphs and comparison with the chi-square measure of the fit of the approximating curve to the Landsat data was used to evaluate this problem. Initial definition of this problem, "convergence of the curve to the training field data appears incomplete," resolved to "the curve fit to the data was marginal because the data is difficult to fit." This redefined problem occurred with more frequency than anticipated; many of the segments exhibited data patterns which could not be closely approximated by the modeling function. However, in these segments curve fit may be adequate for crop profile definition and crop separation, and it may produce a good classification even when the chi-square measure of the curve fit to the training data was large and the curve approximation aesthetically displeasing. The quality of the curve is important, however. Poor curves, when the data defined shallow curves or even lines, tended to result in overclassifications even if these were well fitted to the data.

### 5.4.2 SUMMARY

Much of the Landsat training field data cannot be closely fitted by a curve; atmospheric conditions, cropping practices, and the normal variation in data often produce lumpy crop profiles. The local training used in this method is sensitive to local conditions, which is an advantage since the lumps tend to have significance. The classification method seems to be able to tolerate considerable smoothing; i.e., poor fit of the approximating curve to the training data. The quality of the curve (the definiteness of the curvature) is important for crop separability.

In classification, the use of the chi-square measure is very effective. However, observation of the chi-square measure of the fit of the curve to the training data, an intermediate step, does not seem significant for predicting classification results. A low chi-square measure can occur from (a) a good curve fit to the training data, (b) disperse training data, or (c) a line approximation of the data. A large chi-square measure can occur from (a) a smoothed and adequate approximation of the training data or (b) a poor data approximation. This chi-square measure must be used with the graphs of the data and the crop profile curve; alone, it is not significant.

It was seldom that all the training fields that were defined for a segment exhibited poor fit of the approximating curve to the data, even in the experimental mode used for this study. This classification method performs very well. When a distinctive crop profile is defined even if training field data cannot be closely approximated by a curve, results tend to be good.

## 5.5 THE EFFECTIVENESS OF THE GROWTH CYCLE ADJUSTMENT AND THE ADEQUACY OF THE RANGE SPECIFIED FOR VARIATION IN PLANTING DATE

### 5.5.1 METHOD USED TO ASSESS PROBLEM

Training fields not in use to define the crop profile were used as test fields in 10 segments. Classification of these fields provided some indication of the effectiveness of the growth cycle adjustment, because the fields tended to exhibit different growth stages relative to the acquisition set. For the preferred classification of each segment, areas of disagreement with the ground truth were examined in detail. This examination located corn which failed to be classified as corn and identified a possible reason for the omission, including an atypical growth cycle for the omission.

The growth cycle adjustment performed very well. If a training field were selected for which the available acquisitions were well distributed for good approximation of the data by a curve, this crop profile curve produced a successful classification of fields for which the available acquisitions did not--because of differences in planting dates--define a good crop profile.

### 5.5.2 SUMMARY

For 28 of the segments, no areas of corn which failed to be classified corn could be attributed to failure of the growth cycle adjustment or planting date range specified in the program. Six segments, two in Missouri and four in Indiana, exhibited signature differences for corn which might have been classified more successfully if the allowed range for planting data had been extended. Problems with the growth stage adjustment were rare and seem to be limited to a geographic region.

## 5.6 IDENTIFICATION OF REASONS FOR GROUND-TRUTH-IDENTIFIED CORN FAILING TO BE CLASSIFIED AS CORN

### 5.6.1 METHOD USED TO ASSESS PROBLEM

The software program MISMAP creates a lineprinter map comparison of a classification with the digitized ground truth. The preferred classification for every segment was examined in detail. The amount of misclassification using this method is small, and the majority of the misclassification is border and edge pixels. The areas of ground-truth-identified corn which failed to be classified as corn were checked by sampling approximately 500 pixels per segment.

### 5.6.2 SUMMARY

Omission errors could usually be traced to unusual corn signatures on the film products or graphs or to an apparent inconsistency between the ground-truth label and the imagery. Sometimes the source of atypical signatures could be traced to agricultural or meteorological causes; for instance, fields were cut early for silage, damaged by hail, or contained pumpkins or soybeans which confused the late acquisition. The sensitivity of the channel crop profiles to episodic events was impressive. It was very easy to tell from the graphs that the crop had been affected; tracing the probable agricultural or meteorological cause was very difficult. Much more detailed correlation of agricultural events to reflectance changes in the Landsat channels needs to be done.

Misclassification, excluding border and edge pixels, is very small using this technique, and, in addition, each segment exhibited different reasons for misclassification. No consistent pattern could be established.

## 6. RESULTS

Classification results are summarized in table 5 in this section. These results are for the sample segments of table 2, the preferred (starred) acquisition set. Training fields are as listed on table 2. Percentages are computed based on all pixels in the segment, regardless of purity, and the base for all percentage calculations is 22 932 pixels.

Table 5 gives the following:

a. Segment number and location.

b. Acquisitions available in the growing season of corn, taken as Julian dates 130 to 300. Landsat-3 acquisitions are denoted by (3). The presence of clouds is indicated by (C), and haze is indicated by (h). Other factors affecting data quality are denoted (a) and explained. Acquisitions used to produce the classification results are underlined.

c. A confusion matrix of the classification is given: (1) ground-truth corn classified as corn; (2) ground-truth corn classified noncorn; (3) ground truth noncorn classified corn; (4) ground-truth noncorn classified as noncorn. These percentages are as calculated by the MISMAP program.

d. The percentage of the segment not ground-truth identified.

e. Comments on factors which might affect classification quality, problem areas encountered in processing the segment, and (for some segments) an assessment of the usefulness of the segment in program development.

Classification agreement with the ground truth improves as the purity of the pixels included in the results improves. Average disagreement for the preferred acquisition set all segments is 17 percent using A.A. pure pixels only; approximate accuracy, then, is 83 percent. Average disagreement is 21 percent over all pixels; approximate accuracy is 79 percent.

# TABLE 5.- CLASSIFICATION RESULTS

| Segment | Location (county, state) | Julian dates of available acquisitions | Confusion matrix, % | | Comments |
|---|---|---|---|---|---|
| | | | C→C<br>N→C | C→N<br>N→C | |
| 127 | Montgomery, Ind. | 152 (3), 161, 197 (C), 207 (3, h), 216, 243 (3), 252, 269 | 39.8<br>8.4 | 9.8<br>39.7 | Acquisition history marginal due to data quality of 197, 207; evidence of training field sensitivity; 2.3% of segment not identified ground truth |
| 133 | Whitley, Ind. | 152 (3), 197, 233, 251, 260 (3,a), 269<br>(a): heavy ground moisture or shadow. | 16.5<br>9.8 | 14.5<br>55.6 | Small fields; marginal curve fit to data since data is difficult to approximate; 3.6% of segment not identified ground truth. |
| 135 | Chickasaw, Iowa | 130, 166 (C), 229 (h), 247 (3), 265 (3), 274 (a), 283 (C, h, 3), 292<br>(a): two-line data drop. | 25.3<br>11.1 | 12.9<br>47.0 | Poor acquisition coverage; segment is not recommended for use in program development; 3.8% of segment not identified ground truth. |
| 141 | Madison, Iowa | 130, 166, 167, 212 (3,C), 220 (may be misregistered), 221 (h), 256 (C,h), 265 (3), 266 (3), 274, 292, 293 | 15.7<br>5.6 | 8.4<br>69.5 | .8% not identified ground truth |
| 144 | Wapello, Iowa | 130, 165 (a), 183 (C) 219, 220, 238, 246 (3), 264 (3), 274, 292, 300 (3)<br>(a): banded, poor data quality. | 10.4<br>3.7 | 9.1<br>72.1 | Poor acquisition coverage and distribution; heavy editing by SCREEN on day 265 led to small training field size; segment not recommended for program development; 4.7% not identified ground truth. |
| 202 | Atchison, Mo. | 167, 212 (3, C), 221 (h), 266 (3), 275 (h), 284 (3), 293 | 8.0<br>11.2 | 15.5<br>59.7 | Poor quality data; segment is not recommended for use in program development; 5.7% of segment not identified ground truth. |
| 205 | Clark, Mo. | 138 (3), 155 (3), 156 (3), 164 (a), 209 (3, C), 218, 219, 246 (3), 272 (h), 282 (3), 290<br>(a): banded, data drop | 7.2<br>5.6 | 9.9<br>68.0 | Wide range of planting dates evident; 9.3% of segment not identified ground truth. |

TABLE 5.- Continued.

| Segment | Location (county, state) | Julian dates of available acquisitions | Confusion matrix, % | | Comments |
|---|---|---|---|---|---|
| | | | C→C<br>N→C | C→N<br>N→C | |
| 209 | Gentry, Mo. | 167, 185, 212 (3), 220, 221, 238, 247 (3), 266 (3), 274, 292 | 4.0<br>4.2 | 4.2<br>80.1 | Low percentage of corn in segment; definition of suitable training field is difficult; 7.5% of segment not identified ground truth. |
| 211 | Grundy, Mo. | 166 ($a^1$), 184, 202, 220 (c), 238 ($a^2$), 247 (3), 265 (3, C), 274, 292<br>($a^1$): banded, ($a^2$) misregistered | 2.6<br>6.5 | 4.1<br>77.5 | Low percentage of corn in segment, and data quality is a problem; 9.3% of segment not identified ground truth. |
| 216 | Mercer, Mo. | 130, 184, 202 (a), 220, 238, 247 (3), 265 (3, h), 274, 292<br>(a): banded | 4.6<br>21.8 | 2.0<br>64.1 | Low percentage of corn in segment; overclassification of pasture and soybeans as corn; 7.6% of segment not identified ground truth. |
| 241 | Deuel, S. Dak. | 133, 134, 169, 187, 196 205, 224, 233 (3), 241, 251 (3), 296 | 13.1<br>5.7 | 12.5<br>64.0 | Short acquisition sequence available; 4.7% of segment not identified ground truth. |
| 800 | Clinton, Iowa | 138 (3), 164, 218, 219 246 (3), 272 (h), 281 (3), 290, 300 (3) | 36.5<br>9.7 | 17.5<br>33.7 | Only four useable acquisitions in corn growing season, segment is not recommended for use in program development; 2.7% of segment not identified ground truth. |
| 804 | Marshall, Iowa | 130, 166, 220 (C), 229 (3) 247 (3), 265 (3), 274, 292 | 36.6<br>8.1 | 9.5<br>42.0 | Only five acquisitions available in growing season of corn; segment is not recommended for use in program development; 3.8% of segment not identified ground truth. |
| 809 | Ogle, Ill. | 163 (C), 164, 209 (C, h), 218, 244 (3), 262 (3), 271, 272, 281 (3) | 35.8<br>6.6 | 17.6<br>32.6 | Channel 1 data could not be approximated well by a curve; 7.3% of segment not identified ground truth. |

TABLE 5.- Continued.

| Segment | Location (county, state) | Julian dates of available acquisitions | Confusion matrix, % | | Comments |
|---|---|---|---|---|---|
| | | | C→C<br>N→C | C→N<br>N→C | |
| 824 | Iroquois, Ill. | 163 (C), 217 (h), 235 (h), 243 (3), 244 (3), 262 (3, h) 271 | 40.8<br>10.3 | 9.1<br>36.9 | Data quality is marginal and there are only five acquisitions available in the corn growing season; segment is marginal for program development; 2.9% of segment not identified ground truth. |
| 832 | Adams, Ind. | 151 (3), 160, 178 (C, h), 232, 268 | 15.6<br>23.0 | 5.9<br>53.1 | Only four acquisitions of useable data quality; segment is not recommended for use in program development; 2.4% of segment not identified ground truth. |
| 837 | Benton, Ind. | 180, 198, 207 (3), 216 (C), 225 (3), 234, 243 (3), 251, 252, 270 | 31.4<br>11.1 | 11.7<br>43.7 | Two separate planting dates for corn; 2.2% of segment not identified ground truth. |
| 842 | Henry, Ind. | 160, 178, 232, 250, 268 | 29.1<br>9.9 | 13.7<br>44.9 | Data is difficult to approximate by a curve; 2.4% of segment not identified ground truth. |
| 843 | Henry, Ind. | 151 (3), 152 (3), 160, 178, 197, 232, 233, 251, 268, 269 | 23.6<br>16.1 | 8.6<br>50.4 | 1.3% of segment not identified ground truth. |
| 852 | Randolph, Ind. | 151 (3), 160, 178, 232, 250, 268 | 14.1<br>5.2 | 13.1<br>65.9 | Data cannot be well approximated by a curve; 1.8% of segment not identified ground truth. |
| 853 | Randolph, Ind. | 151 (3), 160, 178, 232, 250, 268 | 17.4<br>8.4 | 17.3<br>55.3 | Two signatures of corn; some problems with fit of approximating curve using five acquisitions; 1.6% of segment not identified ground truth. |

TABLE 5.- Continued.

| Segment | Location (county, state) | Julian dates of available acquisitions | Confusion matrix, % | | Comments |
|---|---|---|---|---|---|
| | | | C→C<br>N→C | C→N<br>N→C | |
| 854 | Tippecanoe, Ind. | 152 (3, C), 161, 197 (C), 207 (3), 216 (C), 233, 234, 243 (3), 251, 252, 269, 270 | 40.0<br>9.5 | 9.2<br>40.9 | Unusual crop signatures; .3% of segment not identified ground truth. |
| 860 | Wells, Ind. | 151 (3), 152 (3), 160, 161, 178, 197, 232, 233, 251, 268, 269 | 18.3<br>9.9 | 10.0<br>54.5 | Corn still vigorous on day 268; misregistration is a problem; evidence of training field sensitivity; 7.3% of segment not identified as ground truth. |
| 864 | Crawford, Iowa | 141 (3), 150, 159 (3), 186, 222 (C), 231 (3), 249 (3), 258 (C), 267 (3), 294 (C) | 32.0<br>6.9 | 13.1<br>45.9 | Data quality is a problem; acquisition coverage is marginal (gap days, 186-231); evidence of training field sensitivity; 2.1% of segment not identified as ground truth. |
| 865 | Crawford, Iowa | 141 (3), 150, 159 (3), 168, 186, 231 (3), 249 (3), 267 (3), 294 | 20.3<br>6.3 | 12.8<br>56.0 | 4.6% of segment not identified as ground truth; evidence of training field sensitivity. |
| 877 | Ida, Iowa | 150 (h), 186, 222, 231 (3), 267 (3) | 24.5<br>8.3 | 13.7<br>32.2 | Marginal for use in program development since 21.4% of segment is not ground truth identified; data difficult to approximate by curve. |
| 878 | Kossuth, Iowa | 131, 186, 221, 266 (3), 293 | 24.4<br>8.7 | 18.7<br>47.3 | Shortage of acquisitions; segment should not be used in program development; .9% of segment not identified ground truth. |
| 880 | Monona, Iowa | 150, 186, 204 (C), 222, 231 (3), 249 (3), 267 (3), 294 | 29.1<br>4.7 | 15.7<br>48.8 | 1.6% of segment not identified ground truth. |

TABLE 5.- Concluded.

| Segment | Location (county, state) | Julian dates of available acquisitions | Confusion matrix, % | | Comments |
|---|---|---|---|---|---|
| | | | C→C<br>N→C | C→N<br>N→C | |
| 881 | Monona, Iowa | 141 (3), 159 (3), 186<br>213 (3), 222, 231 (3),<br>249 (a), 267 (3)<br>(a): much of image appears dark. | 31.2<br>4.8 | 12.5<br>49.9 | Evidence of training field sensitivity; 1.7% of segment not identified ground truth. |
| 882 | Palo Alto, Iowa | 150, 159 (3), 186,<br>213 (3), 222, 231 (3),<br>258 (C), 267 (3), 293 | 34.7<br>5.0 | 8.3<br>45.9 | 6.1% of segment not identified ground truth. |
| 883 | Palo Alto, Iowa | 150, 186, 204 (C), 213 (3),<br>221, 222, 231 (3),<br>258 (C), 267 (3), 293 | 21.2<br>9.6 | 8.5<br>49.0 | Acquisition coverage marginal; 11.6% of segment not identified ground truth. |
| 886 | Pottawatomie, Iowa | 167, 186 (h), 204 (C),<br>212 (C, 3), 231 (3), 249 (3),<br>258, 267 (3), 293 | 30.4<br>5.4 | 16.6<br>43.7 | Poor data quality; 4% of segment not identified ground truth. |
| 891 | Shelby, Iowa | 159 (3), 168, 186,<br>204 (C), 249 (3, h),<br>258 (h), 267 (3) | 27.8<br>9.7 | 18.8<br>39.9 | Data quality of acquisitions in growing season is very poor; segment is not recommended for use in program development; 3.9% of segment not identified ground truth. |
| 892 | Shelby, Iowa | 167 (a), 204 (C), 212 (3, h),<br>221, 231 (3), 240 (C),<br>249 (3, C, h), 266 (3),<br>267 (3), 293<br>(a): shadow | 40.1<br>8.1 | 10.4<br>36.9 | Inadequate acquisitions of good data quality in early season of corn; segment is not recommended for use in program development; 4.5% of segment not identified ground truth. |

Matrices of average values are given below for AA pure pixels (A) and for all pixels (B).

$$(A)\ \begin{pmatrix} 24.8 & 9.3 \\ 7.7 & 54.6 \end{pmatrix} \qquad (B)\ \begin{pmatrix} 23.7 & 12.3 \\ 8.7 & 50.9 \end{pmatrix}$$

Pure pixels only, 34 segments  All pixels, 34 segments

$$\begin{pmatrix} C \rightarrow C & C \rightarrow N \\ N \rightarrow C & N \rightarrow N \end{pmatrix} \qquad \begin{pmatrix} C \rightarrow C & C \rightarrow N \\ N \rightarrow C & N \rightarrow N \end{pmatrix}$$

Omitting the eight segments with processing deficiencies listed on page 5-2 did not have much effect on average disagreement. For pure pixels only, disagreement averages 17 percent and approximate accuracy is 83 percent (matrix C). For all pixels, disagreement averages 19.8 percent for an approximate average accuracy of 80.2 percent (matrix D).

$$(C)\ \begin{pmatrix} 24.4 & 8.7 \\ 8.3 & 53.9 \end{pmatrix} \qquad (D)\ \begin{pmatrix} 23.2 & 11.7 \\ 8.1 & 52.0 \end{pmatrix}$$

Pure pixels only, 26 segments  All pixels, 26 segments

# 7. RECOMMENDATIONS

- More research should be done into the selection process for training fields which are the basis of successful classification results. It is recommended that this research have a double purpose:

  1. Establishment of criteria for training field selection. If possible, criteria should be based on the use of the PFC film products (or interactive console) alone; this would make it possible to omit the IMAPLT step used in the procedure. Graphs of the training field could be presented with the classification results, and IMAPLT could be reserved for use in rework.

  2. Establishment of a set of training fields for a segment which are similar in graphic and film product signature. A data set of several such segments could be used to decide this question: Do comparable training fields generate comparable classification results, or are classification results sensitive to the choice of the training field?

- This method of classification should be extended for multicrop use by sequential classification of pixels. Those pixels classified as corn should be edited from the image, then a crop profile of soybeans could be used to classify the remaining pixels as soybeans or as rejected. A data set suitable for testing this application could be selected from the data set presented in this report. Sites selected should have good acquisition coverage in the growing season.

- Inadequate acquisitions of good data quality are a continuing problem in use of Landsat data for any classification method. For application of Badhwar classification, it is recommended that two potential paths be assessed for improved use of the available acquisition coverage.

  1. A computer program using three well-distributed acquistions should be developed for establishing the crop profile. Results of four-acquisition classification were comparable to those based on five acquisitions; perhaps a further reduction to three acquisitions is feasible.

  2. More research should be done into the use of preemergence and postharvest acquisitions. In this study, such acquisitions were used when crop -

profile definition was poor (i.e., curves tended to be flat) without apparent penalty. Relaxation of the postemeergence to preharvest range for acquisitions would increase the number of acquisitions available for Badhwar classification.

- A decision to increase the range allowed for the estimated planting date in Missouri and Indiana should be assessed by correlating ground-truth cornfields exhibiting unusual signatures with the local cropping practices.
- Similarly, research should be conducted into the correlation of unusual crop profiles in a channel to agricultural and meteorological events. If significance can be defined for the observed episodic events, usefulness of the channel crop profiles will be increased.
- The order in which the chi-square cuts in individual channels are applied should be examined for possible effect on classification.

# 8. EVALUATION

The Badhwar classification method applied to corn is very successful. For this report, unsuitable training data were processed in an attempt to exaggerate problem areas. The accuracy of the method, even under adverse conditions such as inadequate acquisition coverage, frustrated this approach and confirmed the stability of the method.

Recommendations for extending the application of the program arose from this study, but no recommendations can be suggested for improving the performance of the program. This program should be considered operational for the classification of corn.

## 9. REFERENCES


1. Austin, W. W.: Implementation of Badhwar Classification of Corn/Soybean Segments. LEC-14064, October 1979.

2. Badhwar, G.: A Semi-Automatic Technique for Multitemporal Classification of a Given Crop. JSC-16381, July 31, 1979.

3. Badhwar, G.: Crop Emergence Data Determination from Spectral Data. Photogrammetric Engineering and Remote Sensing, vol. 46, no. 3, March 1980, pp. 369-377.

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

5. 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, September 1979.

6. Austin, W. W.: Classification of Wheat: Badhwar Profile Similarity Technique. LEC-15305, September 1980.

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

8. Badhwar, G.: A Program Called TRJPLT for Use on the PDP-11/45 Image Processor. SF3/6531, Feb. 5, 1979. (This is an early version of IMAPLT.)

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

10. Hocutt, W. J.: A.A. Application System Improvements: Letter to A. G. Houston (SF4), ref. 644-1618, April 1980.