E81-10207 CR-160970

SR-LO-00499
JSC-16826

NOV 0 6 1980

AgRISTARS

NASA CR-160970

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

Supporting Research

October 1980

CLASSIFICATION OF WHEAT: BADHWAR PROFILE SIMILARITY TECHNIQUE

Willa W. Austin

(E81-10207) CLASSIFICATION OF WHEAT: BADHWAR PROFILE SIMILARITY TECHNIQUE (Lockheed Engineering and Management) 32 p HC A03/MF A01 CSCL 02C

N81-31598

Unclas

G3/43 00207

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

NASA

Lyndon B. Johnson Space Center
Houston, Texas 77058

1. Report No. SR-LO-00499; JSC-16826	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Classification of Wheat: Badhwar Profile Similarity Technique	5. Report Date October 1980 6. Performing Organization Code
7. Author(s) Willa W. Austin Lockheed Engineering and Management Services Company, Inc.	8. Performing Organization Report No. LEMSCO-15305
9. Performing Organization Name and Address Lockheed Engineering and Management Services Company, Inc. 1830 NASA Road 1 Houston, Texas 77058	10. Work Unit No. 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 (Technical Monitor: J. D. Erickson)	13. Type of Report and Period Covered Technical Report 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The Badhwar Profile Similarity classification technique has been used successfully for classification of corn. This paper documents the results of applying the technique to spring wheat classification. 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 Profile Similarity classification technique Crop profile discrimination Spectral separability Spring wheat classification Classification techniques	18. Distribution Statement

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 32	22. Price*

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

NASA — JSC

SR-LO-00499
JSC-16826

CLASSIFICATION OF WHEAT: BADHWAR PROFILE SIMILARITY TECHNIQUE

Job Order 73-302

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

PREPARED BY

Willa W. Austin

APPROVED BY

T. C. Minter

T. C. Minter, Supervisor
Techniques Development Section

J. E. Wainwright

J. E. Wainwright, Manager
Development and Evaluation Department

LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.

Under Contract NAS 9-15800

For

Earth Observations Division

Space and Life Sciences Directorate

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

October 1980

LEMSCO-15305

PREFACE

The research which is the subject of this report was conducted to support the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing program. Under Contract NAS 9-15800, personnel of Lockheed Engineering and Management Services Company, Inc., completed this work for the Earth Observations Division, Space and Life Sciences Directorate, National Aeronautics and Space Administration, at the Lyndon B. Johnson Space Center.

CONTENTS

PRECEDING PAGE BLANK NOT FILMED

1. INTRODUCTION

The Profile Similarity technique for crop classification developed by Dr. Gautam Badhwar of the National Aeronautics and Space Administration (NASA), Lyndon B. Johnson Space Center (JSC), has proven effective for the classification of corn (ref. 1). This method incorporates the effects of emergence date distribution into the classification and bases classification on the temporal profile of the crop of interest (refs. 2 and 3). The classification method is for a specific crop. Results of applying the technique to the classification of spring wheat in the U.S. northern Great Plains are documented in this report.

The procedure used to apply software programs developed for the classification of corn to the classification of spring wheat is given, and numerical results are presented. The site data set is listed in section 2, along with the Accuracy Assessment (AA) ground-truth percentages for spring wheat, barley, and oats for each site. Section 3 describes the procedure followed for segment classification. Results of classification, in tabular form, are presented in section 4. Concluding remarks are given in section 5, and a recommended procedure for operational use of this technique is presented in section 6.

2. DATA SET

The segment data set for this study consists of 17 spring wheat sites for which the AA digitized ground-truth maps are available. This data set was used for the Analyst Labeling/Procedure M experiment (ref. 4).[1] The set has also been processed using Procedure 1 (ref. 5).

Geographical distribution of the sites is limited: 13 sites are in North Dakota, 3 are in Minnesota, and 1 is in Nebraska. Crop year 1978 acquisitions from Landsat-2 and Landsat-3 were used. The full segment, 22 932 picture elements (pixels), was classified.

Table 2-1 lists the sample segment numbers for the sites; the location of the sites; and the AA ground-truth percentages of spring wheat (including durum wheat), barley, and oats in each of the segments.

---

[1] The original data set contained 18 sites; however, because the AA digitized ground truth was not available for segment 1835 (Ottertail, Minnesota), this site was omitted from classification.

TABLE 2-1.- AA GROUND-TRUTH PERCENTAGES FOR THE SAMPLE SEGMENTS USED IN CLASSIFICATION

Sample segment	Location (county, state)	Spring wheat (including durum wheat)	Barley	Oats
1380	Kimball, Nebr.	7.00	0.05	1.81
1392	Benson, N. Dak.	26.24	5.37	1.14
1457	Ward, N. Dak.	42.04	1.22	2.67
1461	Pierce, N. Dak.	31.70	4.67	3.47
1467	Towner, N. Dak.	39.81	10.79	0.31
1473	Cass, N. Dak.	31.73	16.99	0.64
1518	Roseau, Minn.	22.19	2.79	7.53
1566	Red Lake, Minn.	17.70	5.19	5.07
1602	Mountrail, N. Dak.	26.47	1.08	1.90
1612	McHenry, N. Dak.	10.99	0.26	0.23
1619	Grand Forks, N. Dak.	35.72	0.41	0.30
1636	Stutsman, N. Dak.	36.76	2.24	3.90
1650	Hettinger, N. Dak.	16.30	0.91	4.43
1653	Burleigh, N. Dak.	14.64	0.40	3.71
1656	Morton, N. Dak.	3.75	0.47	2.85
1825	Norman, Minn.	12.83	4.88	8.45
1920	Sioux, N. Dak.	16.89	0.47	4.90

3. PROCEDURE

A detailed description of the modeling used in Badhwar Profile Similarity classification is given in references 2 and 3. Implementation methods are essentially the same as those explained in reference 1 for corn. All data processing used to generate the classifications was done on the programmed data processor, Model 11/45 (PDP 11/45). The software programs which are referred to in the procedure below are described in appendix A of this document.

a. Large Area Crop Inventory Experiment (LACIE) segment images for all available acquisitions were unloaded to a PDP 11/45 disk using the IMUNLD2A program.

b. The data quality of the imagery for all available acquisitions was noted using the production film converter (PFC) film products of these acquisitions.

c. Using the PFC products, at least four candidate training fields of spring wheat were defined with a tentative set of acquisitions. Reference to ground truth confirmed the field labels. If possible, AA special fields were included.

d. Candidate training fields were graphed over the tentative acquisition set using the IMAPLT software program. If necessary, the field definition was revised. Occasionally, signature abnormalities which could not be seen on the imagery eliminated a field from use, and new fields were defined. The acquisition choice was verified from the IMAPLT graphs; if necessary, alternate acquisitions were selected.

e. Two fields were selected as training fields; these fields were those which, based on the imagery and graphs, could be expected to produce acceptable classifications. The remaining fields were used as test fields. Since the study was conducted as a research and development effort, this field selection step sometimes involved several iterations using the IMAPLT program. The sites which were used exhibited a range of problems normal for LACIE segments in the U.S. northern Great Plains; e.g., low percentages of wheat, strip fields, inadequate and poorly distributed acquisitions, and poor data quality. The effects of these problems on classification were informally assessed; none precluded the generation of acceptable classification results. However, defining suitable training data, which is essential for good classification, was made more difficult by the problems evident in one-half of the sites used for this study.

f. For each of the training fields, a data file was established which contained the acquisition set, training field coordinates (line, pixel), and initial guesses (derived from the IMAPLT graphs) for the modeling constants for the training field.

g. Classification was done in the batch-processing mode. The required computer time varied from 20 to 60 minutes.

h. From the classification files[2] generated based on each training field, statistical summary sheets were output. The map formats of the classification file and the digitized ground-truth file were compared. These files were evaluated as described in reference 1. Areas of disagreement were examined; and, if possible, the reason for the disagreement was identified.

i. The test field classification was noted. This proved to be useful for assessing the classification results.

j. Classification results were compiled into the tables presented in section 4. Detailed records of the processing done for each segment are available.

---

[2]The classification files generated for this report have been released to AA for comparison with results using the Procedure 1 classification technique.

4. CLASSIFICATION RESULTS

Classification results are presented in tabular form. Two classifications, based on different training fields, were produced for each segment. For some segments, the classification results are very different. These differences may be due to the extreme shortage of suitable training fields or to data quality problems. Field 1 is the preferred training choice for the segment. Results are given for AA pure pixels only and for all (22 932) pixels.[3]

All percentages are as calculated by the MISMAP program. The percentage of the segment not identified by ground truth plus the classified and rejected percentages will be 100 percent of the segment.

For this study, spring wheat is defined as spring wheat and durum wheat *only* (AA codes 95, 100, 120, and 125). Other small grains are considered misclassified if they classify as spring wheat.

For each segment, the following items are listed in table 4-1.

a. Sample segment number

b. Sample segment location

c. Acquisitions available (Julian date) for the segment [Consecutive-day acquisitions are omitted. Landsat-3 acquisitions are denoted by (3).]

d. Acquisitions used for the classification results presented

e. Coordinates of each of the fields used to train the classifier

f. The number of pixels in each training field

[3]The percentages based on pure pixels (defined by AA to be those pixels which on a subpixel level contain only one crop) appear above the dashed line in table 4-1; for all pixels, below the dashed line.

g. A confusion matrix of the classification in scene percentages:
(1) ground-truth spring wheat classified as spring wheat (S → S),
(2) ground-truth spring wheat not classified as spring wheat (S → N),
(3) ground-truth nonspring wheat classified as spring wheat (N → S), and
(4) ground-truth nonspring wheat not classified as spring wheat (N → N)

h. Proportion estimate comparison of the classified proportion of spring wheat and the ground-truth proportion (These proportions do not include the area which was not identified by ground truth.)

i. The percentage of the segment not identified by ground truth [This includes unknown fields (AA code 80), unidentified areas (AA code 164), and areas block-identified as strip fields of spring wheat (AA codes 170, 175, 220, and 225).]

j. Additional comments, including the number of AA pure pixels in the scene which was used to compute the percentages

TABLE 4-1.- CLASSIFICATION RESULTS

Segment	Location (county, state)	Acquisitions: Available, Julian date	Acquisitions: Used in classification[a]	Training field: Coordinates (line, pixel)	Training field: Size, pixels	Confusion matrix, %: S→S / N→S	Confusion matrix, %: S→N / N→N	Wheat proportion estimate: classified/ground truth, % (classified)	Wheat proportion estimate: classified/ground truth, % (ground truth)	Percentage of segment not identified by ground truth	Comments
1380	Kimball,	115	115	(110, 130)	18	2.5	2.3	5.0		8.7	According to ground truth, only 4.7% of this segment is wheat; however, it is a good corn and soybean segment. Training fields were homogeneous, but fields of acceptable minimum size could not be defined for training. (20 440 AA pure pixels are in the scene.)
	Nebr.	169	169	(110, 134)		2.5	84.2		4.7		
		196 (3)	196	(113, 134)		2.9	4.0	5.7		10.0	
		205	205	(113, 130)		2.8	80.4		6.9		
		222	222	(76, 68)	18	2.2	2.6	5.6		8.7	
		232 (3)		(76, 72)		3.4	83.1		4.7		
		241		(79, 72)		2.7	4.2	6.5		10.0	
		249 (3)		(79, 68)		3.8	79.4		6.9		
		263 (3)									
1392	Benson,	136	136	(15, 53)	30	18.5	5.5	37.7		0.8	Acquisition coverage for this segment is inadequate. Fields tend to be small. (20 143 AA pure pixels are in the scene.)
	N. Dak.	154	154	(14, 62)		19.2	56.0		24.0		
		190	190	(17, 62)		19.7	6.4	39.9		1.1	
		208	217	(17, 53)		20.2	52.6		26.0		
		217 (3)		(93, 65)	36	8.3	15.7	13.4		0.8	
				(93, 70)		5.1	70.1		24.0		
				(100, 73)		8.4	17.7	13.9		1.1	
				(100, 68)		5.5	67.3		26.0		
1457	Ward,	156	174	(34, 170)	39	21.5	15.3	36.0		0.7	The area is dotted with small lakes. Spring wheat exhibits two distinct growth cycles: (1) early — vigorous on 174 and harvested on 246 and (2) late — barely emergent on 174 and ripe on 264. Characterization of the late wheat profile is better, but classification tends to miss early wheat and to misclassify summer crops such as sunflowers into wheat. (19 558 AA pure pixels are in the scene.)
	N. Dak.	174	228	(34, 175)		14.5	47.7		36.8		
		228	246	(40, 175)		21.6	16.8	36.6		0.9	
		246	264	(40, 170)		15.0	45.5		38.4		
		264		(85, 22)	23	21.5	15.4	36.3		0.7	
		273 (3)		(85, 27)		14.8	47.7		36.8		
				(88, 27)		21.4	16.9	36.7		0.9	
				(88, 22)		15.[illegible]	4[illegible]		38.4		

[a]Two classifications, based on different training fields, were produced for each segment. Results for the first classification appear above the solid horizontal line; for the second classification, below the solid line. Percentages based on AA pure pixels appear above the dashed line; for all pixels, below the dashed line.

TABLE 4-1.- Continued

Segment	Location (county, state)	Acquisitions: Available, Julian date	Acquisitions: Used in classification	Training field: Coordinates (line, pixel)	Training field: Size, pixels	Confusion matrix, %: S→S / N→S	Confusion matrix, %: S→N / N→N	Wheat proportion estimate: classified/ ground truth, %	Percentage of segment not identified by ground truth	Comments
1461	Pierce, N. Dak.	118	155	(7, 159)	38	20.1	10.0	31.2	1.7	This segment suffered hail damage between acquisitions 190 and 199. The damage is most visible on the imagery for day 208 (omitted for this classification). (20 039 AA pure pixels are in the scene.)
		136	190	(7, 166)		11.1	57.1	30.1		
		155	199	(11, 166)		20.1	11.8	31.5	2.0	
		190	217	(11, 159)		11.4	54.7	31.8		
		199 (3)		(84, 60)	38	20.9	9.2	34.9	1.7	
		208		(84, 68)		14.0	54.2	30.1		
		217 (3)		(90, 68)		21.0	10.8	35.5	2.0	
		236 (3)		(90, 60)		14.5	51.7	31.8		
1467	Towner, N. Dak.	137	154	(82, 60)	38	22.9	16.0	38.8	3.0	Data quality is poor for days 190, 191, 200, and 208. Days 199, 200, and 217 (Landsat-3 acquisitions designated as Landsat-2) failed to receive the Landsat-3 adjustment, and the data could not be well approximated by a curve. Days 190 and 199 were overscreened by SCREEN;[b] this program had to be removed before fields could be selected. Misclassification reflects these problems. (19 207 AA pure pixels are in the scene.)
		154	190	(82, 65)		15.9	42.2	38.9		
		190	199	(89, 65)		22.6	17.2	39.2	3.5	
		199 (3)	217	(89, 60)		16.6	40.1	39.8		
		208		(12, 3)	37	21.2	17.7	35.9	3.0	
		217 (3)		(12, 10)		14.7	43.4	38.9		
				(18, 10)		20.9	18.9	36.3	3.5	
				(18, 3)		15.4	41.3	39.8		
1473	Cass, N. Dak.	116	116	(97, 182)	29	16.2	14.3	26.6	4.1	Acquisition coverage was very poor; no acquisitions were in the green-up phase. Use of a preemergent early date produced a data pattern which could not be well approximated by a curve. (19 884 AA pure pixels are in the scene.)
		197 (3)	197	(94, 194)		10.4	55.1	30.4		
		207	207	(96, 194)		16.3	15.4	27.5	4.9	
		224	224	(99, 182)		11.2	52.0	31.8		
		242		(47, 164)	39	12.7	17.7	24.8	4.1	
		251 (3)		(47, 171)		12.1	53.4	30.4		
		269 (3)		(53, 171)		12.7	19.1	25.0	4.9	
				(53, 166)		12.3	51.0	31.8		

[b]A procedure developed by the Environmental Research Institute of Michigan (ERIM) for automatically detecting garbled data, as well as clouds, snow, cloud shadows, and water in multispectral scanner data.

TABLE 4-1.- Continued.

Segment	Location (county, state)	Acquisitions: Available, Julian date	Acquisitions: Used in classification	Training field: Coordinates (line, pixel)	Training field: Size, pixels	Confusion matrix, %: S→S S→N / N→S N→N	Wheat proportion estimate: classified/ ground truth, %	Percentage of segment not identified by ground truth	Comments
1518	Roseau,	116	153	(60, 40)	35	11.2 9.3	17.4	5.5	Data quality is a problem: acquisitions 188, 206, and 224 have some haze and/or clouds. A final date of 243 can be substituted, but it is a harvest date. (20 213 AA pure pixels are in the scene.)
	Minn.	135	188	(60, 45)		6.2 67.9	20.5		
		153	206	(65, 45)		11.2 10.9	17.6	6.4	
		188	224	(65, 40)		6.4 64.9	22.2		
		206		(35, 14)	39	17.4 3.1	28.3	5.5	
		224		(35, 22)		10.9 63.1	20.5		
		243		(39, 22)		17.7 4.5	29.1	6.4	
		251 (3)		(39, 14)		11.4 59.9	22.2		
		260							
1566	Red Lake,	115	133	(51, 180)	24	8.7 7.3	13.6	14.5	The ERIM SCREEN program removed all pixels from acquisitions 196 and 232, so SCREEN had to be removed before fields could be defined; these acquisitions are of acceptable quality. Inadequate acquisitions resulted in the use of a harvest date in classification. (19 417 AA pure pixels are in the scene.)
	Minn.	133	169	(51, 185)		4.9 64.7	16.0		
		169	196	(55, 186)		8.7 9.1	13.7	16.0	
		196 (3)	232	(55, 181)		5.0 61.4	17.7		
		232 (3)		(77, 135)	37	9.1 6.9	14.7	14.5	
				(77, 142)		5.6 63.8	16.0		
				(84, 142)		9.2 8.5	15.0	16.0	
				(84, 135)		5.8 60.5	17.7		
1602	Mountrail,	174	174	(18, 65)	38	12.1 10.6	18.8	0.2	This segment has insufficient acquisitions and strip fields. Planting dates are wide spread, hence a variety of curve characterizations over the available acquisitions. The training choice is extremely limited. (19 165 AA pure pixels are in the scene.)
	N. Dak.	211	211	(18, 71)		6.7 70.4	22.8		
		228	228	(27, 73)		13.0 13.4	20.9	0.3	
		264	264	(27, 69)		7.9 65.4	26.4		
				(74, 187)	39	16.8 5.9	41.7	0.2	
				(75, 192)		24.9 52.0	22.8		
				(80, 192)		18.5 7.9	44.2	0.3	
				(80, 185)		25.7 47.4	26.4		
1612	McHenry,	118	155	(94, 101)	38	3.3 4.8	14.6	0.6	This segment has a low wheat proportion and strip fields; only one satisfactory training field could be defined. Acquisition coverage is marginal; therefore, a harvest acquisition had to be used in classification. (21 107 AA pure pixels are in the scene.)
	N. Dak.	137	199	(94, 108)		11.3 80.0	8.1		
		155	218	(98, 108)		4.1 6.8	15.5	0.8	
		199 (3)	236	(98, 101)		11.4 76.9	10.9		
		218 (3)		(92, 109)	26	5.5 2.6	55.8	0.6	
		236 (3)		(92, 115)		50.3 41.0	8.1		
				(96, 116)		7.4 3.5	56.3	0.8	
				(96, 112)		48.9 39.4	10.9		

TABLE 4-1.- Continued.

Segment	Location (county, state)	Acquisitions: Available, Julian date	Acquisitions: Used in classification	Training field: Coordinates (line, pixel)	Training field: Size, pixels	Confusion matrix, %: S→S / N→S	Confusion matrix, %: S→N / N→N	Wheat proportion estimate: classified/ ground truth, %	Percentage of segment not identified by ground truth	Comments
1619	Grand Forks, N. Dak.	135	135	(32, 136)	37	24.0	10.5	37.6	0.7	Acquisition coverage is inadequate, and the distribution is poor. No coverage was in the green-up period. (20 152 AA pure pixels are in the scene.)
		198 (3)	198	(32, 143)		13.6	51.2	34.5		
		207	207	(38, 145)		23.6	12.2	37.5	0.9	
		216 (3)	216	(38, 139)		13.9	49.4	35.8		
		243		(3, 85)	37	31.4	3.1	49.5	0.7	
		252 (3)		(3, 90)		18.1	46.7	34.5		
		270 (3)		(10, 92)		31.6	4.2	50.8	0.9	
				(10, 85)		19.2	44.0	35.8		
1636	Stutsman, N. Dak.	117	136	(27, 90)	28	16.9	18.5	22.4	1.6	This segment has a wide range of growth patterns for ground-truthed wheat. Poor data quality was a problem, and a cloudy acquisition had to be used in classification. (19 774 AA pure pixels are in the scene.)
		135	190	(26, 96)		5.5	57.5	35.4		
		154	207	(30, 96)		16.4	20.5	22.5	1.3	
		190	216	(30, 90)		6.1	55.2	36.9		
		207		(34, 182)	20	16.6	18.8	20.3	1.6	
		216 (3)		(34, 186)		3.7	59.3	35.4		
		226		(39, 186)		15.9	21.0	20.1	1.8	
		243		(39, 184)		4.2	57.2	36.9		
		270 (3)								
1650	Hettinger, N. Dak.	137	155	(17, 3)	40	5.3	7.3	9.1	17.0	This segment contains strip fields, and 17% has been block ground truthed so comparisons cannot be made in these areas. Strip fields separated very well in the classification. (18 847 AA pure pixels are in the scene.)
		155	191	(17, 10)		3.8	66.7	12.6		
		191	209	(22, 10)		5.5	10.8	9.6	16.8	
		209	218	(22, 3)		4.1	62.8	16.3		
		218 (3)	228	(108, 76)	21	4.4	8.2	7.1	17.0	
		228		(108, 182)		2.7	67.7	12.6		
		236 (3)		(110, 182)		4.5	11.8	7.3	16.8	
		246		(110, 176)		2.8	64.0	16.3		
		264								
		273 (3)								
1653	Burleigh, N. Dak.	101	154	(93, 48)	38	8.3	3.3	15.3	2.3	This segment contains strip fields. Classification results are very good, and strip fields separated well in the classification. (20 272 AA pure pixels are in the scene.)
		119	191	(93, 55)		7.0	79.1	11.6		
		136	208	(98, 57)		9.5	5.1	17.3	3.0	
		154	217	(98, 49)		7.8	74.5	14.7		
		191		(3, 141)	40	8.0	3.5	18.3	2.3	
		199 (3)		(3, 158)		10.3	75.8	11.6		
		208		(5, 158)		9.2	5.5	19.9	3.0	
		217 (3)		(5, 142)		10.7	71.6	14.7		

TABLE 4-1.- Concluded.

Segment	Location (county, state)	Acquisitions: Available, Julian date	Acquisitions: Used in classification	Training field: Coordinates (line, pixel)	Training field: Size, pixels	Confusion matrix, %: S→S / N→S	Confusion matrix, %: S→N / N→N	Wheat proportion estimate: classified	Wheat proportion estimate: ground truth, %	Percentage of segment not identified by ground truth	Comments
1656	Morton,	101	137	(93, 10)	27	0.4	1.5	2.8		1.4	This segment has far too low a wheat proportion (2%) or small-grain proportion (5%) to process. Pure training samples simply cannot be defined when there is so little wheat in the segment. Acquisition 218 was not flagged as Landsat-3. (21 595 AA pure pixels are in the scene.)
	N. Dak.	137	155	(93, 18)		2.4	94.3		1.9		
		155	209	(95, 18)		0.7	2.9	3.3		2.1	
		191	263	(95, 10)		2.6	91.7		3.6		
		209		(114, 180)	39	1.6	0.3	81.2		1.4	
		218 (3)		(113, 190)		79.6	17.1		1.9		
		263		(117, 190)		3.0	0.6	80.5		2.1	
				(117, 180)		77.5	16.7		3.6		
1825	Norman,	97	133	(75, 141)	40	7.1	5.1	11.8		24.1	This segment contains much small-grain confusion cropland which separates well from the spring wheat. Ground-truth identification is insufficient. (18 975 AA pure pixels are in the scene.)
	Minn.	133	169	(75, 148)		4.7	59.0		12.2		
		169	196	(80, 150)		6.7	6.1	11.6		25.8	
		196 (3)	206	(80, 141)		4.9	56.5		12.8		
		206	223	(77, 161)	25	5.1	7.1	10.2		24.1	
		223		(77, 164)		5.1	58.5		12.2		
		232 (3)		(84, 166)		5.0	7.9	10.4		25.8	
		242		(84, 163)		5.4	56.0		12.8		
		250 (3)									
1920	Sioux,	101	136	(78, 28)	39	8.0	5.9	20.4		0.4	This segment has strip fields; pasture and hay are confusion crops. (20 821 AA pure pixels are in the scene.)
	N. Dak.	136	199	(78, 37)		12.4	73.4		13.8		
		199 (3)	209	(82, 37)		8.7	8.0	21.1		0.6	
		209	217	(82, 28)		12.4	70.2		16.8		
		217 (3)	236	(49, 180)	35	9.5	4.4	26.9		0.4	
		236 (3)		(49, 184)		17.4	68.4		13.8		
		271 (3)		(57, 187)		10.6	6.2	28.1		0.6	
				(57, 184)		17.5	65.1		16.8		

5. CONCLUSIONS

The average percentages of misclassification in the ground-truth-identified area of each segment using field 1 are 17.6 percent for pure pixels only and 19.8 percent for all pixels. Using field 2, misclassifications average 25.6 percent for pure pixels and 27.4 percent for all pixels. Matrices of average misclassification are shown below.

Training field	Pure pixels	All pixels
1	$\begin{pmatrix} 12.2 & 8.7 \\ 8.9 & 65.1 \end{pmatrix}$	$\begin{pmatrix} 12.4 & 10.4 \\ 9.4 & 62.0 \end{pmatrix}$
2	$\begin{pmatrix} 12.5 & 8.4 \\ 17.2 & 56.8 \end{pmatrix}$	$\begin{pmatrix} 12.9 & 10.0 \\ 17.4 & 54.0 \end{pmatrix}$

A scene accuracy of 75 to 80 percent is a reasonable expectation for the classification of spring wheat using the Badhwar Profile Similarity technique.

This method of classification can be applied effectively to segments with a very low percentage of wheat. Finding suitable training fields in these segments can be difficult; however, if a good crop profile is defined, classification results are good.

The chi-square value gives an estimation of the adequacy of the crop profile curve as an approximation of the training field data. It must be used with visual examination of the training field data to assure that (1) a curve is defined and (2) the field data are compact. Chi-square values are dependent upon the standard deviations of the data and must be monitored. This is cumbersome. More objective criteria should be provided for assessing training field data approximation.

The use of test fields is an efficient and effective aid to assessing classification results.

For North Dakota, the acquisition coverage and the range allowed for estimated planting date in the program (currently ±15 days from the estimated planting date of the training field) should be extended. In North Dakota, spring wheat has an early and a late planting. If acquisition coverage of the segment is cut off in August, after the early planted wheat is harvested, the late planted wheat tends to be misidentified as a summer crop. The option to remove the restriction on the planting date range should be provided.

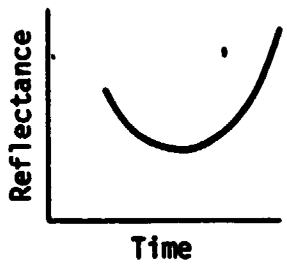
6. RECOMMENDATION

Spring wheat in these segments in the U.S. northern Great Plains was classified using the same program — with no changes — used for the classification of corn in the Corn Belt.[4] The overall accuracy of the results confirms the adaptability of the Badhwar Profile Similarity classification technique to a variety of crops. This accuracy also indicates that the software program used for classification, CLASFYT, should be considered operational for spring wheat, as well as for corn.

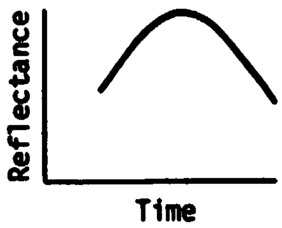
The following procedure is recommended for operational use of this classification technique.

a. Choice of training field and test fields:

Wheat is distinguished from other vegetation by its growth cycle over a time interval. In the Landsat bands covering the visible spectral regions, channels 1 and 2, this cycle defines a curve similar to that shown below in reflectance versus time.



In the near infrared spectral regions, channels 3 and 4, the curve is similar to the following:



[4]This work is being documented by the author at the present time.

On the PFC imagery, where channels 1 and 2 are blue and green and channel 4 is red, this relative channel reflectance change depicts a time signature change of gray (bare soil; approximately equal spectral response and color components) to red (vegetation; since the infrared value rises as the visible value decreases) and back to yellow or gray tones (stubble and bare soil). An appropriate time interval can be determined from ancillary data; i.e., crop calendars and regional statistics.

- Choose four candidate fields which are (1) of a 20-pixel minimum size; (2) free of roads, nonagricultural components, etc.; and (3) free of clouds and haze. These fields should exhibit a continuous gray to red to brown/gray/green signature sequence in the proper time interval. *Use PFC product 3.*
- Determine field borders; border and edge pixels must be avoided in field definition. Define field coordinates at least two pixels inside field borders on all available acquisitions. *Use PFC products 1 and 2.*
- Designate one field as the training field on which the classification will be based. The remaining fields will be used as test fields to aid in classification evaluation.

b. <u>Choice of acquisition set</u>:

Acquisitions for classification are chosen to characterize the wheat growth cycle; the set selected should be well distributed over this cycle. Cloudy or hazy acquisitions, as well as those which are preemergent for wheat or which exhibit appreciable amounts of wheat harvest, should be avoided. Classification can be done on a set of four or five acquisitions. A five-acquisition classification increases crop separability, but this many suitable acquisitions may not be available.

If problems with the acquisition choice or with the field definition exist, the analyst may choose to plot the field and acquisition set using IMAPLT before entering the classification.

c. Classification:

Classifidation will be done in the batch-processing mode on the Interactive Multispectral Image Analysis System, Model 100 (Image 100), PDP 11/45 image processor. The analyst will input the field coordinates and acquisition set via cards. The required computer time is 20 to 60 minutes.

d. Evaluation of results:

Products of classification will be:

- A summary sheet of input values and calculated parameters based on the training field data with a numerical summary of pixel classification results (fig. 6-1)
- A full-scene classification map generated as a film product for analyst use (fig. 6-2)

The film product classification map should have evident field patterns. Fields should be well filled out with a minimum of blank spaces (pixels rejected as spring wheat) in the field interiors. Blank areas also should be clear; i.e., be reasonably free of scattered pixels classified as spring wheat. Scattered pixels or blank areas may indicate an overclassification or an underclassification. The classification map should be overlaid on the PFC film products to check agreement of the classification with the analyst's identification.

On the summary sheet, calculated parameters should be checked; the estimated planting date of the training field as generated on each channel should be the same within the estimated planting date error, and the chi-square fit should be less than 10 in each channel.[5] A scene

[5]More objective criteria for evaluating the statistical output need to be determined. Currently, visual examination of the channel graphs is used to assess the compactness of the training field data and the adequacy of the curve approximation of the data. With this, a chi-square fit of less than 10 is meaningful. The chi-square fit value is dependent upon the data variation; an evaluation value which combines these parameters should be defined for operational use.

proportion estimate can be computed from the pixels classified as spring wheat, and this estimate can be compared with available statistics. As an addition to the summary sheet of figure 6-1, the numerical proportion of pixels correctly classified as spring wheat should be supplied for each of the analyst-defined test fields; the test field accuracy should be 70 percent or better.

If the classification is unsatisfactory, rework will consist of selecting an alternate training field or acquisition set. This choice should be graphed, using the IMAPLT program, before use as the basis for segment reclassification.

ACCURACY ASSESSMENT QUALITY ASSURANCE CLASSIFICATION

PRØCESSING DATE - 12-MAR-80 AT 02:58:04

SEGMENT NUMBER - 1653   CRØP OF INTEREST - SWHE

IMAGE FILES USED IN CLASSIFICATION - DB2:[111,3]165378136.IM2
DB2:[111,3]165378154.IM2
DB2:[111,3]165378191.IM2
DB2:[111,3]165378191.JM2
DB2:[111,3]165378208.IM2

TRAINING FIELD - LINE NØ.   SAMP. NØ.

LINE NØ.	SAMP. NØ.
73.0	112.0
70.0	129.0
73.0	130.0
75.0	112.0

MEANS AND STD. DEV. FØR TRAINING FIELD BASED ØN 40 PIXELS -

CHANNEL NUMBER		----------ACQUISITION DATES---------- 78136	78154	78191	78191	78208
1	MEAN	30.11	23.03	21.17	21.17	23.21
	STD. DEV.	2.17	1.51	0.93	17.76	0.99
2	MEAN	33.46	17.33	18.25	18.25	24.71
	STD. DEV.	2.32	1.73	1.97	37.74	1.63
3	MEAN	37.36	40.83	45.82	45.82	40.61
	STD. DEV.	2.01	2.18	1.74	35.88	2.26
4	MEAN	31.63	36.80	44.35	44.35	36.88
	STD. DEV.	1.63	1.09	2.98	61.04	1.91

CØNSTANTS FØR MØDEL -

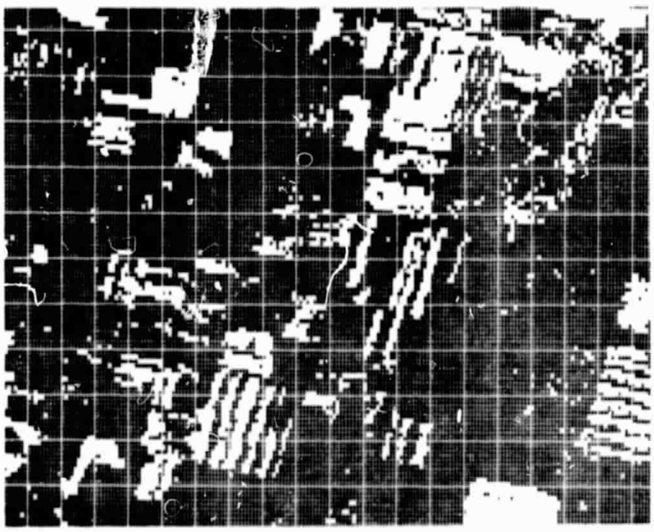
CHANNEL NUMBER		A	ALPHA	BETA	TØ	CHISQ
1	INITIAL	3.80	-5.37	-0.82	1.20	
	FINAL	3.77 +- 5.75	-5.51 +- 1.65	-0.84 +- 0.28	1.19 +- 2.19	0.11
2	INITIAL	5.00	-15.19	-2.49	1.20	
	FINAL	4.69 +- 8.94	-15.16 +- 2.64	-2.49 +- 0.45	1.15 +- 1.21	0.97
3	INITIAL	3.22	3.91	0.62	1.20	
	FINAL	3.31 +- 5.40	3.47 +- 1.52	0.54 +- 0.26	1.15 +- 3.05	1.24
4	INITIAL	3.00	5.23	0.81	1.20	
	FINAL	3.09 +- 0.00	4.17 +- 0.00	0.63 +- 0.00	1.16 +-0.00	2.22

CHISQ THRESHØLD - CHANNEL	1	2	3	4
THRESHØLD	7.04	7.04	7.04	8.17

CLASSIFICATIØN RESULTS -

PIXELS CLASSIFIED SWHE - 9689   CUT BY CH2 - 4456
PIXELS SCREENED - 0   CUT BY CH3 - 3521
PIXELS CLASSIFIED NØN-SWHE - 13243   CUT BY CH4 - 5266

Figure 6-1.- Example of a statistics summary sheet.



ORIGINAL PAGE IS
OF POOR QUALITY

Figure 6-2.- Film product classification map of the full scene.

APPENDIX A

SOFTWARE PROGRAMS USED FOR CLASSIFICATION OF SPRING WHEAT USING BADHWAR PROFILE SIMILARITY TECHNIQUE

APPENDIX A

SOFTWARE PROGRAMS USED FOR CLASSIFICATION OF SPRING WHEAT USING BADHWAR PROFILE SIMILARITY TECHNIQUE

A.1 IMUNLD2A

IMUNLD2A takes an image unload tape generated on the Earth Resources Interactive Processing System (ERIPS), edits it using SCREEN (ref. 6), 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 on a PDP 11/45 disk.

a. Input: ERIPS image unload tape

b. Output: screened images with adjusted Landsat-3 acquisitions on a PDP 11/45 disk

A.2 IMAPLT

IMAPLT (ref. 8)[6] 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 (two for each Landsat 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. The segment number, the acquisitions used, the coordinates of the field, the 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 planting date of the field (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 are presented on the second plot.

---

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

a. Input: field coordinates in order; acquisition set of four or five acquisitions

b. Output: ght graphs as described above

A.3 CLASFYT

CLASFYT (ref. 9) computes the constants for the curves from the training field data, compares the values for each pixel in the segment with these curves (the crop profile in each channel),[7] and rejects those pixels which are not within a specified chi-square measure of the profile. The technique for rejection is to compare the pixel channel values with the profiles in channel 2, channel 3, channel 4, then channel 1 in succession and reject the pixel if the comparison in any single channel is unsatisfactory. Variability of the time of planting and/or emergence is allowed for in the comparison of individual pixels with the crop profile (refs. 2 and 3). Accepted pixels are labeled as spring wheat; rejected pixels, nonspring wheat.

a. Input: four or five image files, coordinates of one crop-of-interest field to establish crop profiles, and initial values for the function constants as computed in IMAPLT (to aid convergence of the approximating curve)

b. Output: classification file on disk which has a designation of spring wheat or nonspring wheat for each pixel in the segment; line-printer sheet summarizing the following:

- Acquisitions used
- Training field coordinates and the number of pixels in the field
- Mean and standard deviation for each channel and each acquisition (field averages)
- The input and the final constants (with error) for the model

---

[7]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. Pixels in the training field failing to pass this edit step were excluded from processing, hence from affecting the crop profiles. However, screened pixels were restored before classification of the segment so all 22 932 pixels are designated "wheat" or "nonwheat" (ref. 10).

- Final chi-square values for each channel (training field data)
- Estimated planting date of the training field (with error) as derived for each channel
- Chi-square thresholds in each channel applied as cutoff values in classification
- The number of pixels cut for exceeding the chi-square threshold, hence removed from consideration as spring wheat, in each channel
- The final numerical results: the number of pixels classified as spring wheat, the number of pixels screened (always zero in this study), and the number of pixels rejected as spring wheat

A.4 A2SGMAP

A2SGMAP provides a full-scene classification map (22 932 pixels) of the results obtained using CLASFYT. The scale is the same as that used for the AA digitized ground-truth maps. Pixels classified as spring wheat are designated "C" (crop of interest), and those rejected as spring wheat are left as blank spaces on the map.

a. Input: classification file from CLASFYT

b. Output: line-printer map of the full-scene classification

A.5 TAPEOUT

TAPEOUT (ref. 11) reads the data files produced by CLASFYT and creates Universal-formatted 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.

a. Input: classification file from CLASFYT

b. Output: black-and-white classification map of the full scene on film

A.6 MISMAP

MISMAP (ref. 9) compares the classification file produced by CLASFYT with the AA digitized ground-truth inventory map for the segment. A line-printer map with the following codes is generated:

a. Ground-truth spring wheat classified as spring wheat appears as S.

b. Ground-truth nonspring wheat rejected as spring wheat is left blank.

c. Ground-truth nonspring wheat classified as spring wheat appears as +.

d. Ground-truth spring wheat rejected as spring wheat appears as -.

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

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

A numerical scene summary is given in confusion matrix form.

MISMAP line-printer maps can be generated for all pixels or for AA pure pixels only.

a. Input: classification file from CLASFYT and ground-truth inventory map file

b. Output: full-scene line-printer map comparing the classification file produced by CLASFYT with the AA digitized ground-truth inventory map and a confusion matrix numerical summary of results

APPENDIX B

REFERENCES

APPENDIX B

REFERENCES

1. Austin, Willa W.: Implementation of Badhwar Classification of Corn/ Soybean Segments. Lockheed Technical Memorandum LEC-14064, Oct. 1979.

2. Badhwar, G.: A Semi-Automatic Technique for Multitemporal Classification of a Given Crop. AgRISTARS SR-J0-00481 (JSC-16381), July 31, 1979.

3. Badhwar, G.: Crop Emergence Date Determination From Spectral Data. Photogram. Eng. Remote Sensing, vol. 46, no. 3, Mar. 1980, pp. 369-377.

4. Holmes, Q. A.; Horvath, R.; Cicone, R. C.; Kauth, R. J.; and Malila, W. A.: Development of Landsat-Based Technology for Crop Inventories: Appendices. AgRISTARS SR-E9-00404$_2$, Contract NAS 9-15476, Dec. 1979.

5. Heydorn, R. P.; Bizzell, R. M.; Quirein, J. A.; Abotteen, K. M.; and Sumner, C. A.: Classification and Mensuration of LACIE Segments. Proc. Technical Sessions, Vol. I, The LACIE Symposium, October 1978. JSC-16015, July 1979, pp. 73-86.

6. Lambeck, P. F.; and Nalepka, R. F.: Signature Extension Preprocessing for Landsat MSS Data. ERIM Final Report 122700-32-F, Nov. 1977.

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. SF3/6541, Feb. 5, 1979.

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

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

11. Hocutt, W. T.: A.A. Application System Improvements. Letter to A. G. Houston (SF4), ref. 644-1618, Apr. 1980.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1980 - 771-014/ 26 98