OCT 19 1978

JSC-14483

79-10011

CR 151847

LEC

LOCKHEED ELECTRONICS COMPANY

SYSTEMS AND SERVICES DIVISION

16811 EL CAMINO REAL ☐ HOUSTON, TEXAS 77058 ☐ TELEPHONE (AREA CODE 713) 488-0080

Ref: 642-6951
Contract NAS 9-15200
Job Order 73-743-23

TECHNICAL MEMORANDUM

SECONDARY ERROR ANALYSIS: THE EVALUATION OF ANALYST DOT LABELING

By

K. A. Havens

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

(E79-10011) SECONDARY ERROR ANALYSIS: THE EVALUATION OF ANALYST DOT LABELING (Lockheed Electronics Co.) 17 p HC A02/MF A01 CSCL 05B

N79-12527

Unclas
G3/43 00011

September 1978

LEC-12380

A SUBSIDIARY OF LOCKHEED AIRCRAFT CORPORATION

PRECEDING PAGE BLANK NOT FILMED

CONTENTS

TABLES

1. INTRODUCTION

At the close of the Large Area Crop Inventory Experiment (LACIE) Phase III crop year, several investigations were outlined in support of the Classification and Mensuration Subsystem (CAMS). The goal of the secondary error analysis plan was to use Procedure 1 to evaluate as many of the error sources as possible in the small-grains estimate for 5- by 6-nautical-mile segments in the U.S. Great Plains. The component of the plan which is considered herein is the evaluation of analyst labeling errors on type 1 and type 2 dots. The purpose of this study is to determine the effects of the mislabeling experienced on classification and proportion estimation performance. Mislabeling errors are the only secondary errors that introduce both bias and increased variance into the proportion estimates and are thus viewed as the most important source of errors.

2. GENERAL DESCRIPTION OF PROCEDURES

This experiment was conducted by relabeling the Procedure 1 dot grids (ref. 1, pp. 4-8 through 4-10) with ground-truth information and reprocessing the segments. Two types of dot grids were used in the test: the random dot grid used in Phase III and the uniform dot grid that is being used in the Transition Year. Each of these two proportion estimates was then compared to the latest acceptable CAMS estimate for the Phase III crop year.

Relabeling of the Procedure 1 dot grids required an accurate ground-truth map on a pixel-by-pixel basis. To produce the ground-truth map for each segment, the county agent annotated an aerial photograph of the entire segment. The field boundaries were digitized in the photo-cartographic laboratory, and both of these products were then used to produce a tape of all ground-truth information. This tape was, in turn, used in a computer program (ref. 2) to generate a gray-scale map with the crop code represented for each pixel.

The ground-truth maps were registered to the latest acquisition date processed by CAMS for which an acceptable proportion estimate was produced. These acquisitions are referred to as the base acquisition dates for the ground-truth maps. For the reprocessing of a test segment, the same base acquisition date was used in the generation of the ground-truth map, in the Phase III processing of the segment, and in the reprocessing of the segment using ground-truth labeled random and uniform dot grids. Agreement of these base acquisition dates ensured a minimal amount of misregistration between ground-truth maps and images.

The multiple processing of a test segment required that exactly the same multitemporal acquisition dates be used. The same combination of random grid dots was labeled using ground truth as had been labeled in the Phase III processing. Because the processing was kept as nearly the same as possible, the only variable being observed was the dot labeling error.

3. DESCRIPTION OF THE EXPERIMENT

The data set chosen for use in this study consisted of Phase III blind sites that had been worked using Procedure 1. The data included nine states within the U.S. Great Plains, with the number of segments in each state varying from one to five. Initially, a total of 30 test segments was chosen; however, because of incomplete ground-truth coverage and problems in reprocessing, the final total included only 25 segments. These test data, along with information on acquisitions used in processing, are more fully described in table 1. Of the final 25 test segments, 5 segments had additional problems in the computation of a digitized ground-truth estimate (as described in section 2); therefore, a 400-dot-count estimate of the ground-truth proportion was used.

A randomized complete block experimental design was used to analyze the data. The data were blocked by state with segments nested within the states. The treatments consisted of three classifications differentiated by dot labeling techniques: analyst-interpreter (AI) labeling of the random grid system, ground-truth labeling of the random grid system, and ground-truth labeling of the uniform grid system. The model for the three analyses of variance (ANOVA) was as follows:

$$y_{ijk} = \mu + b_i + s_{ij} + t_k + bt_{ik} + e_{ijk}$$

where

μ = the overall mean of the observations

b_i = the block or state effect ($i = 1, 2, \cdots, 9$)

s_{ij} = the segment effect ($j = 1, 2, \cdots, n_i$ as n_i varies with state; $i = 1, 2, \cdots, 9$)

t_k = the treatment or labeling effect ($k = 1, 2, 3$)

bt_{ik} = the block-by-treatment effect

e_{ijk} = the random error for each observation

y_{ijk} = the response variable

In this experimental design, the assumption was made that there is an interaction between treatments and states. This could occur because of varying field sizes among states, a fact which would affect the analyst's labeling capability.

For the first two analyses, y_{ijk} was the true probability of correct classification (PCC) for segment j within state i using labeling procedure k as calculated from type 1 starting dots (PCC1) and type 2 bias correction dots (PCC2), respectively. For these two analyses, only 20 test segments were included ranging over 8 states. This was because of incomplete ground-truth information, which prohibited the calculation of the PCC in the remaining five segments.

The response variable for the third analysis is defined as follows:

$$y_{ijk} = |\hat{p}_{ijk} - p_{ij}|$$

where $\hat{p}_{ijk}$ is the proportion of wheat for segment j within state i using labeling procedure k, and p_{ij} is the ground-truth proportion for segment j within state i.

A Newman-Keuls multiple comparison test was planned in the event that any of these ANOVA tests indicated significant treatment differences.

4. RESULTS

Table 2 presents the results of computing the PCC1 for each of the three labeling procedures (treatments). Table 3 lists the ANOVA test results of the PCC1. The states, the segments within the states, and the treatments were found to be significantly different at the 5-percent level of significance. The Newman-Keuls test was performed, and the results appear in table 4. Means are arranged in ascending order, and differences at the 5-percent level of significance are separated by brackets. There were differences found between the AI labeled procedure and each of the two ground-truth labeled procedures. However, no difference was found between the two ground-truth labeled grid systems: uniform and random.

Table 5 presents the results of computing the PCC2 for each of the three treatments, and table 6 gives the corresponding ANOVA test results. The states and the segments within the states show differences at the 5-percent level of significance. There was no significant difference between treatments, as indicated by the similarity of the treatment means.

Table 7 presents the proportion estimates for the three treatments and the ground-truth proportions. Table 8 presents the differences between the ground truth and the AI labeled proportion (Δ_{AI}); the ground-truth proportion and the ground-truth labeling of the random dot grid proportion (Δ_R); and the ground-truth proportion and the ground-truth labeling of the uniform dot grid proportion (Δ_U). Table 9 is the ANOVA table prepared using these differences. No significant difference was found in any of the factors examined.

5. CONCLUSIONS

From this examination of 25 test segments using AI labeling and ground-truth labeling, the PCC on type 1 dots was found to be significantly better for both types of ground-truth labeled procedures than the PCC obtained using AI labeling. No significant difference in the PCC was found for type 2 dots. However, in all three treatments, the type 2 dots included pixels which fell on boundaries or were mixed pixels (ref. 1, pp. 7-1 and 7-2). This would account for all PCC2 values being equally low.

The proportion estimates achieved in these classifications showed no significant differences between procedures. By reviewing table 8, the large variance among the proportion estimates is obvious; this may account for the nonsignificance found in table 9. However, the ranking of the treatment means places both types of ground-truth labeled procedures above the AI labeled procedure. It was expected that ground-truth labeling would be a significant improvement to the classifier. Using Procedure 1, however, comparable estimates were achieved regardless of the type of labeling procedure.

6. REFERENCES

1. Phase III CAMS Detailed Analysis Procedures. LACIE-00720 (JSC-11693), Aug. 1977.

2. Loe, D. L.; Hayenga, W. B.; and Ahlers, C. W.: "As-Built" Design Specification for PDP 11/45 Accuracy Assessment System Using Disk Data File. Lockheed Electronics Co., Inc., LEC-11881, Feb. 1978.

TABLE 1.— TEST DATA

Segment*	Location (county, state)	Acquisitions used	Robertson biostage†
1005 (W)	Cheyenne, Colo.	7177, 7159, 6326, 6254	6.0, 4.8, 2.5, 1.0
1032 (W)	Wichita, Kans.	7194, 7086, 6326, 6254	6.0, 2.7, 2.5, 1.0
1033 (W)	Clark, Kans.	7156, 6288	5.1, 2.2
1853 (W)	Ness, Kans.	7193, 7067, 6253	6.0, 2.0, 1.0
1861 (W)	Kearny, Kans.	7194, 7158, 7104, 6326	6.0, 4.9, 3.0, 2.5
1512 (S)	Clay, Minn.	7193, 7156	5.2, 3.5
1520 (S)	Big Stone, Minn.	7174, 7156, 7120	4.4, 3.6, 2.0
1544 (S)	Sheridan, Mont.	7198, 6294	5.8, 0.0
1739 (M)	Teton, Mont.	7222, 7168, 7132, 6263	6.0/6.0, 4.2/3.8, 3.1/2.5, 1.8/1.0
1582 (W)	Hayes, Nebr.	7194, 7158, 7086, 6254	6.0, 4.6, 2.6, 1.0
1604 (S)	Renville, N. Dak.	7143, 7125	2.7, 1.9
1606 (S)	Ward, N. Dak.	7197, 7125	5.3, 1.9
1648 (S)	Bowman, N. Dak.	7179, 7125	4.5, 2.0
1661 (S)	McIntosh, N. Dak.	7159, 7123	3.5, 1.8
1902 (S)	McKenzie, N. Dak.	7197, 7125	5.2, 2.0
1231 (W)	Jackson, Okla.	7156, 7066, 6288	5.4, 3.0, 2.2
1242 (W)	Canadian, Okla.	7173, 7155, 7101, 6287	6.0, 5.6, 3.4, 2.0
1367 (W)	Major, Okla.	7155, 7101, 6287	5.3, 3.4, 2.0
1677 (S)	Spink, S. Dak.	7211, 7193, 7176, 7140	6.0, 5.7, 4.5, 2.9
1690 (S)	Kingsbury, S. Dak.	7211, 7193, 6325	6.0, 6.0, 0.0
1803 (W)	Shannon, S. Dak.	7178, 7159, 7123, 6255	6.0, 4.7, 3.2, 2.0
1805 (M)	Gregory, S. Dak.	7211, 7158, 6307, 6290	6.0/6.0, 4.7/4.0, 2.5/0.0, 2.4/0.0
1056 (W)	Moore, Tex.	7158, 7049, 6290	5.2, 2.7, 2.4
1059 (W)	Ochiltree, Tex.	7157, 7121, 6325, 6307	5.1, 3.8, 2.5, 2.4
1060 (W)	Sherman, Tex.	7158, 7068	5.2, 2.8

*W = winter wheat; S = spring wheat; M = mixed wheat.

†Robertson scale adjustable crop calendar.

TABLE 2.— PCC1 RESULTS

Segment - state	Labeling procedure		
	Random	Uniform	AI
1005 - Colo.	87.2	86.0	89.7
1032 - Kans.	89.7	91.8	84.6
1033 - Kans.	88.0	92.0	87.8
1853 - Kans.	96.8	90.9	87.1
1055 - Tex.	95.0	93.7	89.7
1059 - Tex.	87.8	92.0	83.3
1060 - Tex.	97.8	93.7	87.0
1231 - Okla.	100.0	97.7	97.2
1242 - Okla.	100.0	100.0	96.8
1544 - Mont.	91.2	71.7	76.5
1582 - Nebr.	97.8	97.9	97.8
1604 - N. Dak.	90.2	81.0	61.5
1606 - N. Dak.	82.6	80.6	75.6
1648 - N. Dak.	83.7	92.9	71.4
1661 - N. Dak.	87.5	91.7	75.0
1902 - N. Dak.	88.0	100.0	96.0
1677 - S. Dak.	90.5	83.3	87.5
1690 - S. Dak.	100.0	91.7	94.9
1803 - S. Dak.	100.0	100.0	98.0
1805 - S. Dak.	92.0	90.6	80.0
Total	1845.8	1819.2	1717.4
Average	92.3	91.0	85.9

TABLE 3.— PCC1 ANOVA TEST RESULTS

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-value	Level of significance, %
State	7	1600	229	9.2	5
Segment within state	12	927	77	3.1	5
Treatment	2	459	230	9.2	5
State by treatment	14	438	31	1.2	NS
Error	24	590	25		
Total	59	4014			

TABLE 4.— NEWMAN-KEULS TEST OF PCC1

Treatment	Mean
AI	85.9
Uniform	91.0
Random	92.3

TABLE 5.— PCC2 RESULTS

Segment - state	Labeling procedure		
	Random	Uniform	AI
1005 - Colo.	71.7	78.3	68.3
1032 - Kans.	76.3	76.7	79.7
1033 - Kans.	89.5	83.3	86.2
1853 - Kans.	85.0	73.3	78.3
1056 - Tex.	65.0	76.7	81.7
1059 - Tex.	84.5	85.0	77.2
1060 - Tex.	94.9	85.0	83.1
1231 - Okla.	89.8	96.6	91.5
1242 - Okla.	86.8	86.8	83.6
1544 - Mont.	55.0	61.7	64.7
1582 - Nebr.	96.7	93.2	96.7
1604 - N. Dak.	76.7	60.0	63.3
1606 - N. Dak.	76.6	62.5	72.3
1648 - N. Dak.	66.7	79.3	61.7
1661 - N. Dak.	81.1	76.0	71.7
1902 - N. Dak.	83.3	86.4	90.0
1677 - S. Dak.	72.5	76.9	84.3
1690 - S. Dak.	88.3	90.0	88.3
1803 - S. Dak.	98.3	98.3	98.3
1805 - S. Dak.	92.9	92.5	89.0
Total	1631.6	1618.5	1609.9
Average	81.6	80.9	80.5

TABLE 6.— PCC2 ANOVA TEST RESULTS

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-value	Level of significance, %
State	7	4068	581	17.1	5
Segment within state	12	1873	156	4.6	5
Treatment	2	12	6	.2	NS
State by treatment	14	253	18	.5	NS
Error	24	811	34		
Total	59	7017			

Treatment	Mean
AI	80.5
Uniform	80.9
Random	81.6

TABLE 7.— PROPORTION ESTIMATES

Segment - state	Labeling procedure			Ground truth, %
	Random	Uniform	AI	
1005 - Colo.	38	48	20	37
1032 - Kans.	37	40	23	39
1033 - Kans.	9	13	2	*9
1853 - Kans.	35	35	26	31
1861 - Kans.	6	25	31	*35
1056 - Tex.	17	26	32	30
1059 - Tex.	38	43	38	46
1060 - Tex.	20	22	17	31
1231 - Okla.	72	74	76	74
1242 - Okla.	50	50	51	47
1367 - Okla.	62	58	36	54
1512 - Minn.	16	28	31	*33
1520 - Minn.	22	22	21	*31
1544 - Mont.	60	40	43	38
1739 - Mont.	33	37	20	*24
1582 - Nebr.	16	14	18	19
1604 - N. Dak.	53	54	35	52
1606 - N. Dak.	25	33	19	32
1648 - N. Dak.	33	26	36	38
1661 - N. Dak.	37	35	33	41
1902 - N. Dak.	11	8	7	9
1677 - S. Dak.	28	40	24	34
1690 - S. Dak.	18	26	9	21
1803 - S. Dak.	2	3	2	1
1805 - S. Dak.	16	19	6	16

*Indicates that the 400-dot-count estimate was used.

TABLE 8.— DIFFERENCES BETWEEN PROPORTION ESTIMATES

Segment - state	Labeling procedure		
	Δ_R	Δ_U	Δ_{AI}
1005 - Colo.	1	11	17
1032 - Kans.	2	1	16
1033 - Kans.	0	4	7
1853 - Kans.	5	5	5
1861 - Kans.	29	10	4
1056 - Tex.	13	4	2
1059 - Tex.	8	3	8
1060 - Tex.	11	9	14
1231 - Okla.	2	0	2
1242 - Okla.	3	3	4
1367 - Okla.	8	4	18
1512 - Minn.	17	5	2
1520 - Minn.	9	9	10
1544 - Mont.	22	2	5
1739 - Mont.	9	13	4
1582 - Nebr.	3	5	1
1604 - N. Dak.	1	2	17
1606 - N. Dak.	7	1	13
1648 - N. Dak.	5	12	2
1661 - N. Dak.	4	6	8
1902 - N. Dak.	2	1	2
1677 - S. Dak.	6	6	10
1690 - S. Dak.	3	5	12
1803 - S. Dak.	1	2	1
1805 - S. Dak.	0	3	10

TABLE 9.— PROPORTION ESTIMATE ANOVA

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-value	Level of significance, %
State	8	234	29	0.88	NS
Segment within state	16	508	32	.97	NS
Treatment	2	96	48	1.45	NS
State by treatment	16	490	31	.94	NS
Error	32	1067	33		
Total	74	2395			

Treatment	Means
Uniform	5.0
Random	6.8
AI	7.8