LACIE-00610
JSC-10812

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)

NASA NOAA USDA

PHASE I ACCURACY ASSESSMENT PLAN

National Aeronautics and Space Administration
LYNDON B. JOHNSON SPACE CENTER
Houston, Texas
January, 1976

LACIE Phase 1

Accuracy Assessment Plan

Approved by ______________________

D. E. Pitts, Accuracy Assessment

Approved by ______________________

R. B. MacDonald, LACIE Manager

FOREWORD

The Phase I Accuracy Assessment effort is designed to use the data collected at the intensive test sites and selected LACIE sample segments to determine the components of the LACIE acreage error, and to determine which of these, if any, prevents the LACIE error budget from being met. Some of the major potential error sources which will be investigated are: sampling, crop calendar, registration, wheat proportion estimation, aggregation, and standard statistics algorithms.

TABLE OF CONTENTS

ACCURACY ASSESSMENT TEAM

TEAM MEMBER	
D. Pitts	TF3
R. Bizzell	TF4
A. Feiveson	TF3
J. Carney	TF4
R. Patterson	TF4
J. Ridgely	TF4
R. Chhikara	C09
E. Hsu	C09
C. Liszcz	C09
W. Palmer	C09
F. Solomon	C09
J. Sommers	C09
R. Stuff	C09
G. Baron	C09
W. Crea	TF4
D. Frank	TF4

1.0 OBJECTIVES

The LACIE Phase I accuracy assessment is designed to check the accuracy of products produced by the operations and thereby determine if the procedures used are sufficient to meet the LACIE goals. Thus accuracy assessment is distinct from Quality Assurance which determines if the LACIE procedures are being followed.

The following items are the objectives of the accuracy assessment effort for LACIE Phase I:

1) Estimate the variance, bias, and confidence for the LACIE Phase I output.
2) Assess the components of the LACIE operation to a level of detail that is sufficient to identify the source and magnitude of error contributions.

2.0 SCOPE AND RESOURCE REQUIREMENTS

The following is a list of questions (in priority) which the LACIE Phase I accuracy assessment will try to answer:

1) What are the relative sizes of the error components in wheat area estimation?

   - Sampling - The within county variance will be calculated by photointerpretation. Simulation runs will be made using TRW Error Model.
   - Per Segment Proportion Estimate - Evaluate error sources using ITS by AI/Biostage/Geography.
   - Crop Calendar - Analyses will be run using both nominal and correct biostages. In two areas: one for winter wheat and one for spring wheat.

2) Is the accuracy of the wheat proportion estimate significantly different for:

   - Spring and winter wheat
   - Wheat distribution
   - Strip fallow vs. continuous crop region
   - Single pass and multitemporal
   - Early biostage and at harvest

   The classification results for the ITS and the "blind sites" will be analyzed with respect to the above factors.

3) Are the CAMS wheat proportion estimates significantly biased? Omission and comission errors will be calculated for the ITS and blind sites.

4) Are the CAMS wheat proportion significantly correlated with the actual wheat proportions?

Statistical test will be run comparing the ground truth to the CAMS wheat proportion.

5) Are there significant differences between AI's?

- Three test sites were acquired in all four biostages. All 14 AI's will analyze each of the four biostages for each site to determine difference between AI's relative to biostage.

6) Are the intensive test site representative of the LACIE sample segments?

- The intensive test site ground truth and classification accuracy will be compared to the nearby blind test ground truth and its classification accuracy.
- Discriminant analysis will be performed on the ITS based on soil types, confusion crops, cropping practices, etc.

7) Is the accuracy better for segments judged acceptable by AI-DPA than for those judged unacceptable?

The CAMS results for the "blind sites" will be compared to the ground truthed and analyzed statistically with respect to S, M & U.

8) Does a trend exist between the significance tests at the 1%, 5%, 10% levels and the classification performance?

Standard statistics will be calculated. For the Great Plains at the 10% confidence level, the variance will be used to manually calculate the 1% and 5% confidence levels.

9) Are the CAMS rework estimates of wheat proportion significantly better than the segments processed using the operational procedure for LACIE Phase I?

The rework and the operational products will be compared to the SRS county wheat % and area and where applicable to the blind site wheat area.

Resource Requirements

The resource requirements for these tasks are given in the following Table I.

Computer

If all the biostages were acquired, the total number of ITS computer runs would be approximately 464. Actually 62 acquisitions were accomplished so that 248 individual DPA runs are required to analyze the single pass data. Since most of the runs will be made in the batch mode, the time per run is anticipated to never exceed about eight (8) minutes per run. The estimated total CPU time is 111 hours to process the ITS segments thru ERIPS. An additional 25 hours will be required to process the blind sites.

AI Photo Interpretation

Most of the AI use is required during the tests explained in section 4.1. If all the biostages were acquired for the ITS, then approximately 450 individual photo interpretation Landsat images for training field selection would have to be made. Actually, 62 acquisitions were made so 248 individual photo interpretations will be required. An additional 128 analyses are required to determine the accuracy of the crop calendar effects and the difference between AI's. An additional 75 man days of AI expertise will be required for the blind site analyses.

Statistical Analyses

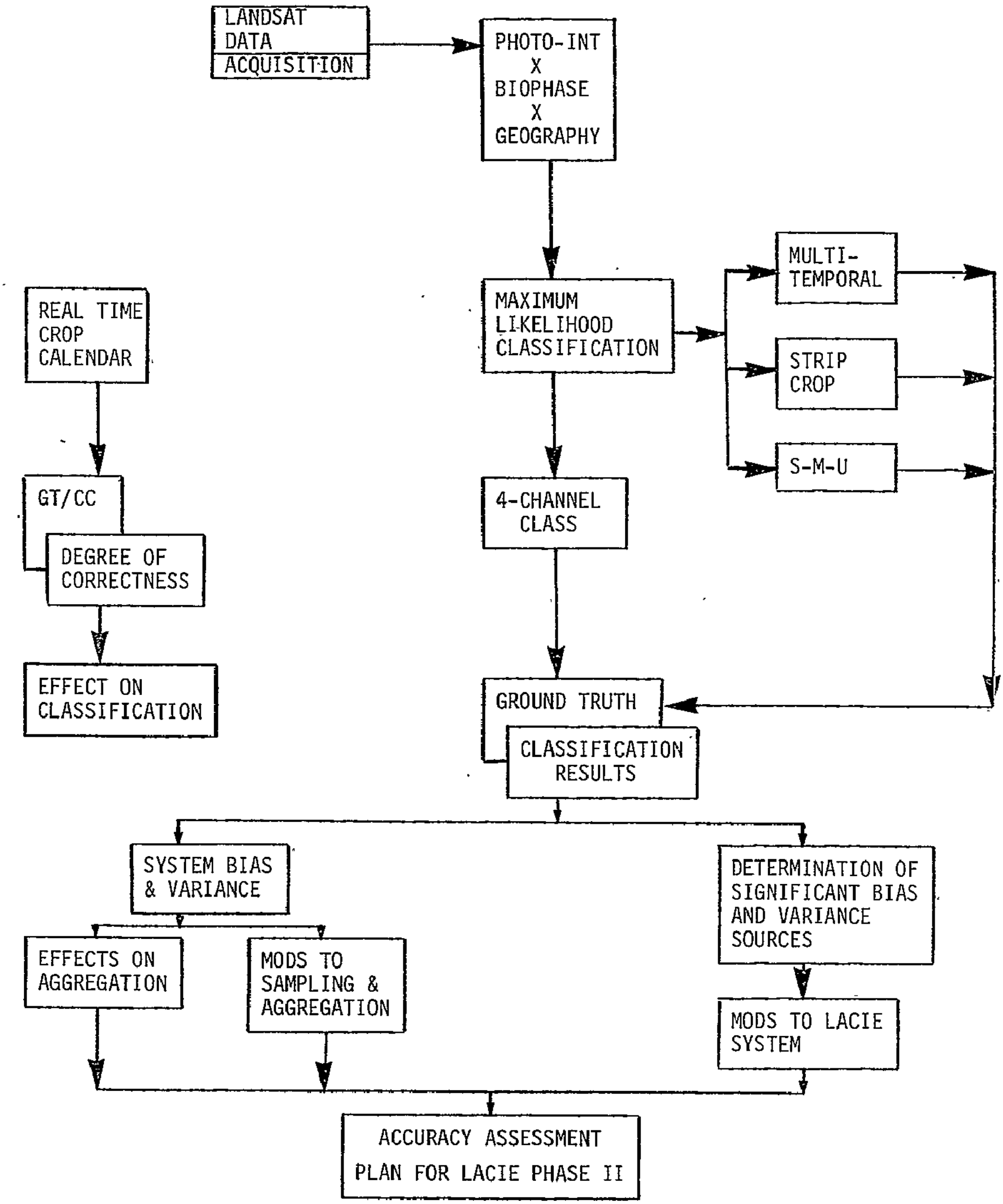
About 2100 hours of statistical analyses will be required to complete the Phase I accuracy assessment analyses.

TABLE I

TASK	AI RESOURCES (Photo Interpretations)	COMPUTER TIME (Hrs.)	STATISTICSL ANALYSES (Man Hours)
ITS Single Pass Analysis of Accuracy	248	80	--
ITS Multitemporal Analysis of Accuracy	0	9	--
ITS Strip Crop Accuracy	*	*	160
ITS Stratification	--	--	160
ITS Component Analysis Single Pass (AI-Biostage-Geography)	*	*	120
ITS Difference Between AI's	120	20	160
ITS System Bias and Variance	*	*	160
Sampling and Aggregation	--	5 hrs U-1110	240
Within County Variance	--	20 hrs I-100 time	112
Simulation Studies	--	70 hrs U-1110	500
Crop Calendar	8	2	80 (includes 60 hrs YES)
Blind Sites			
AI Accuracy of Training to Test Fields	29 man days	2 hrs U-1110	--
Determination of True % Wheat Within Segment	15 man days	--	--
Correct Labels	15 man days	22 hrs	--
Reselect Training Fields Using Ground Truth	15 man days	3 hrs	
Training & Test Field Accuracies	--	--	120 hrs
Statistical Analysis	--	--	120 hrs
Phase I Aggregation Accuracy	--	--	160 hrs
Foreign Sites	10 photo interpretations	--	50 hrs

*Included in single pass and multitemporal analysis

2.0 DATA FLOW


LANDSAT DATA ACQUISITION
PHOTO-INT X BIOPHASE X GEOGRAPHY
MAXIMUM LIKELIHOOD CLASSIFICATION
MULTI-TEMPORAL
STRIP CROP
S-M-U
REAL TIME CROP CALENDAR
GT/CC
DEGREE OF CORRECTNESS
EFFECT ON CLASSIFICATION
4-CHANNEL CLASS
GROUND TRUTH
CLASSIFICATION RESULTS
SYSTEM BIAS & VARIANCE
DETERMINATION OF SIGNIFICANT BIAS AND VARIANCE SOURCES
EFFECTS ON AGGREGATION
MODS TO SAMPLING & AGGREGATION
MODS TO LACIE SYSTEM
ACCURACY ASSESSMENT PLAN FOR LACIE PHASE II

3.0 EVALUATION OF INTENSIVE TEST SITES

3.1 Description

The test sites to be used in this accuracy assessment plan are the twenty nine (29) listed in Table II. These sites are located in eight (8) states and one(1) Canadian province, which combine into about four (4) regions - the northwest United States, the northern Great Plains, the southern Great Plains, and the Great Lakes.

3.2 Results Anticipated

In order to accomplish the objectives, it is necessary that the ITS's be representative of the LACIE sample segments, because the results obtained from the test designs using the ITS's will be extrapolated and correlated to the LACIE sample segments. The significance and importance of this statement should not be underestimated. Therefore, statistical techniques will be used to determine if the ITS's are representative at some level of the area in which they are located, be it the CRD, state, or region. Analysis of LACIE accuracy over a group of regular segments (i.e., blind test) which is described in a later section will be compared to the accuracy of the operational procedures over nearby intensive test sites.

3.3 Statistical Approach

Using the general philosophy applied to the partitioning exercise of signature extension and further, to analogous sites in foreign countries - assessing the similarities of geographic parameters such as soil type, confusion crops, cropping practices, climate, etc. - determine which ITS's are statistically similar. This will involve using discriminant analysis techniques or clustering techniques in which the above mentioned parameters and other geographic type parameters will be used. Hopefully, this approach will establish four (4) regions within which the ITS's are statistically the same. And at worst, it will generate partitioning by state, which would give nine statistically similar areas.

3.4 Options

If the analysis of the ITS's for similitude does not yield a result included in the discussion above, then three conditions can occur and two of them are bad.

- All ITS's statistically similar. This would indicate that the ITS's are not representative of their surrounding area and may not even be representative of any other area in the United States except that 5 x 6 n.mi. area they cover.
- All ITS's statistically different. This would indicate huge variations throughout the United States and could be cause for concern as to the possibility of any success for the project.

The following condition is not a bad indication:

- ITS's are statistically the same within regions but are statistically different among regions.

In either case, a possible and very likely alternative would be to discard the present ITS's and select new ITS's from the LACIE sample segments. However, if this were to be done, the same type of analysis would have to be repeated for those segments.

4.0 ITS ACCURACY TEST DESIGN

4.1 Analyst Interpreter/Biostage/Geography

A full factorial type of test design will be used to evaluate the effects of the above mentioned three factors as well as the three pairs of two factor interactions.

4.1.1 Analyst Interpreter

Although the complete analysis of the effects of AI would utilize all AI's, the constraints of time and manpower dictate that something less than that be used. The alternatives were: (1) To use all the AI's in a fractional type of factorial test design, and (2) to employ a representative sample of the AI's and use a full factorial type of test design.

The first method would be practical if one or more of the two factor interactions was known to be insignificant, then the test could be shortened to some reasonable number of photo interpretations. Since this was not the case, in fact, most of the experience to date seems to indicate that just the opposite is true - that the two factor interactions could be very significant. The second method will be used. Additionally, if operational mode is to be maintained, unbiased consultation must be made available. The second method assumes that there is no difference between AI's. Section 9.0 discusses the procedure to identify these differences.

The trade-off point of a representative sample number versus a reasonable number of photo interpretations was judged to be four AI's. These four individuals, labeled AI #1, #2, #3, and #4, will always be the same four AI's.

4.1.1.1 Procedure for the Selection of AA/AI's

The purpose of this procedure is to insure, as much as possible, that an unbiased, objective, representative sample of the AI;s is obtained for purposes of accuracy assessment of the LACIE Operational System.

1. Any AI who has seen the ground truth for the ITS's is not a test candidate.
2. At least four (4) consultant level AI's must be made available for the test AI's on an individual (nonrepeating) basis.
3. Of the remaining AI's, a random selection process will be followed as:
   - list all candidate AI's in any order
   - using a random number list, assign a number to each moded by the number of candidates
   - use the next number in the list as a pointer of how many numbers to skip
   - The next four (4) numbers, moded by the number of candidates, indicate which of the numbered candidates will be used in the accuracy assessment effort.

4.1.2 Biophase

In the current operational mode, it has been decided that the crop development will be quantized into four (4) biostages. Therefore, in this test design, there will also be four biostages.

4.1.3 Geography

The components of this factor are:

(1) Field size

(2) Confusion crops

(3) Crop calendars

(4) Cropping practices

(5) Soil type

(6) Spring/winter wheat

Table II shows the locations of the twenty-nine (29) intensive test sites (ITS's) that will be used to assess the impact of these components of the geography factor.

TABLE II LACIE INTENSIVE STUDY SITES

Segment Number	State	County	Center Coordinates		US/CA	Site Size	Wheat Type	Acquired As
			N. Lat.	W. Long.				
1960	Kansas	Finney	38°04.2'	101°01.7'		5x6 stat.	W	W
1961	Kansas	Morton	37°16.0'	101°54.0'		5x6 mi.	W	W
1962	Kansas	Saline	38°41.8'	97°28.4'		3x3 mi.	W	W
1963	Kansas	Rice	38°17.0'	98°12.7'		3x3	W	W
1964	Kansas	Ellis	38°50.1'	99°13.0'		3x3	W	W
1965	N. Dakota	Burke	48°53.2'	102°10.0'	Y	5x6 mi.	S	S
1966	N. Dakota	Williams	48°19.2'	103°24.7'	Y	5x6 mi.	S	S
1967	N. Dakota	Divide	48°53.6'	103°10.9'	Y	2x10	S	S
1968	Montana	Glacier	48°37.5'	112°33.4'	Y	2x10 mi.	S&W	S
1969	Montana	Toole	48°53.0'	111°46.5'	Y	2x10 mi.	S&W	S
1970	Montana	Liberty	48°44.0'	110°51.0'	Y	2x10	S&W	S
1971	Montana	Hill	48°42.0'	109°55.0'		2x6	S&W	S
1972	Washington	Whitman 1	46°54.6'	117°15.5'		3x3 mi.	S&W	W
1973	Washington	Whitman 2	46°50.4'	117°48.3'		3x3	S&W	W
1974	Washington	Whitman 3	47°08.0'	117°26.3'		3x3	S&W	W
1975	Idaho	Oneida	42°04.5'	112°29.5'		3x3 mi.	S&W	W
1976	Idaho	Franklin	42°08.0'	111°58.0'		3x3	S&W	W
1977	Idaho	Bannock	42°56.5'	112°25.5'		3x3	S&W	W
1978	Texas	Randall	35°09.5'	102°04.4'		3x3 mi.	W	W
1979	Texas	Deaf Smith	34°52.2'	102°22.3'		3x3	W	W
1980	Texas	Oldham	35°15.0'	102°32.0'		3x3	W	W
1981	Indiana	Shelby	39°27.6'	85°47.2'		3x3 mi.	W	W
1982	Indiana	Madison	40°13.5'	85°37.5'		3x3	W	W
1983	Indiana	Boone	40°05.7'	86°33.5'		3x3	W	W
1984	Sask.	Delisle	51°55'	107°28'		2x10 mi.	S	S
1985	Sask.	Swift Current	50°19'	107°53'		2x10	S	S
1687	S. Dakota	Hand 1	44°35.0'	98°58.0'		5x6 stat.	S&W	S
1986	S. Dakota	Hand 2	44°21.0'	98°45.1'		5x6 mi.	S&W	S
1987	Minnesota	Westpolk	47°49.0'	96°41.0'		5x6 mi.	S	S

TABLE II (CONTINUED)

List of Intensive Test Sites and Biostages Acquired

Segment #	Biostages I	II	III	IV
1687	133		205	
1960	291		150	
1961	291			169
1962	324	131		
1963	289	131		
1964	290			
1965	155	191		
1966				
1967	137	191		227
1968	143	180	216	
1969	161	179	215	233
1970	142	179		233
1971	142			
1972	268			218
1973	268		201	218
1974	268		182	218
1975	159*	178*	195*	213*
1976	299	177	195	213
1977	299		196	214
1978	291		133	
1979	291		133	
1980	291		133	
1981	105*			176*
1982	299	140		
1983	281	141		
1984		195		
1985				
1986	150	169	187	
1987				

*Segments were moved to be coincident with ground truth. Because they were moved, they had to be reordered.

4.1.4 Test Design

Test design for a 4 x 4 x 29 = 464 individual photo interpretations and computer classification runs.

Biophase	1				2				3				4			
AI	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
ITS																
1																
2																
3																
4																
5																
.																
.																
.																
27																
28																
29																

One of the response measurements that will be used in this experiment is the difference of the percent wheat in an ITS segment* as generated by the LACIE system, and that obtained from the ground truth. Other possible responses are: for each class proportion of training fields correctly identified and proportion of test fields correctly classified, number of training fields for each subclass of wheat and non-wheat, proportions of each subclass field and number of subclasses. Also, evaluate the significance level or some other means of separation of classes in a segment.

* ITS segment means the intensive test site area enclosed in a segment.

4.1.5 Missing Data Contingency Plan

4.1.5.1 An attempt will be made to statistically group the various ITS's based on similarities between and within their components; (e.g., soil types, climate, confusion crops). This will be tried at both the regional and state level. If successful at either level, randomly missing segments will not affect the data analysis, though the level of confidence in statistical inference will probably be reduced.

4.1.5.2 If case 4.1.5.1 fails, there are statistical methods for estimating the responses of missing data that will be used. Their disadvantages lie in the fact that they only tend to support the trend of the data that is available if there is one. If no trend exists, it might generate a trend due to the nature of the available data.

4.2 Strip Crop Classification

This will be a simple comparative test to determine if there is a significant difference between the ability to accurately classify an ITS that is predominantly strip cropped versus non-strip cropped, given that the conditions in both are the same.

4.3 ITS Registration

In case of multitemporal processing of a segment errors in registration of one pass to another is expected to cause additional errors in classification. Since multitemporal classification runs will be made by CAMS and compared with single pass classification for accuracy assessment, the knowledge of registration errors will be helpful in making comparative analysis. If the registration error component is relatively significant, the effect of registration errors on the multitemporal classification performance will be assessed.

MPAD has developed an off-line registration program that will be used to evaluate the registration errors for ITS segments. The accuracy assessment team in their comparative analysis will evaluate the multitemporal classification performance in the light of these errors.

4.4 Multitemporal Classification Accuracy

The acquisitions of the intensive test sites are given in Table II. Multitemporal classification runs will be made on all possible combinations of these biostages and these results will be compared to the ground truth to determine the accuracy of each combination. This will require 64 classification runs.

Action: CAMS and Accuracy Assessment Team

Start November 1 -- Complete December 1

4.5 Foreign Area Study

The foreign exploratory segments will be stratified and compared to the ITS according to climate, soil type, crop mix, field size, etc. Due to the high degree of specificity, only a small number (i.e., 10) of exploratory segments and ITS should pair up "exactly." The AI accuracy in the ITS will then be assumed to be similar in the foreign segments. Intensive AI analysis will then be compared with the operational analysis of the foreign segments.

5.0 ITS DATA REQUIREMENTS FOR ACCURACY ASSESSMENT

5.1 Test Field Selection

For selecting the test fields from the ITS segment, the following ITS field data output products are required:

- 1:24,000 scale field boundary overlay
- Periodic observation data form
- Field identification data form

There will be approximately ten wheat test fields and approximately ten non-wheat test fields randomly selected from the "field boundary overlay" for each segment. The "periodic observation data form" and "field identification data form" will be utilized for identifying the crop classes of those test fields. These ground truth information will not be accessible to the AI's nor to the ADP analysts. The crop classes for test fields will be updated according to all available ground truth information of each phase. The final winter wheat ground truth will be used to verify fall planted ground truth. All discrepancies will be noted and appropriate changes made to insure correct results.

The ITS segment will be sent to each of the four AI's for determining the field coordinates for the first phase. Then, the coordinates will be updated for each successive phase. The method for choosing the coordinates should be consistent with standard operational procedures.

5.2 AI Photo Interpretation Report

An AI output should include the following information:

- Segment number
- Acquisition data and biophase

- Analyst interpreter
- Transmittal sheet, including
    - Crop classes
    - Test field numbers by classes
    - Training field numbers by classes
    - Field coordinates

The above results are then sent to the accuracy assessment team. The AI should also prepare the necessary information to transmit to the DPCA for entering into the field data base for data processing. Any variation from standard methods of field determination shall be noted in all reporting of results.

5.3 ADP Classification Report

When the training and test fields of a segment (for a certain phase) have been loaded into field data base, the data will be scheduled for classification under standard operational procedures for LACIE Phase I. For each classification, the following information will be required:

- Segment number
- Biophase(s)
- Statistics report
- Clustering reports (if applicable)
- Training segment number (if different from the recognition segment)
- Classification results for
    - Each training and test field
    - ITS area, and
    - Entire segment

The results of each completed classification run with the above information will then be passed to the accuracy assessment team for final assessment.

5.4 (Segment) Accuracy Assessment Report

A report consisting of the following parametric evaluations is to be prepared by the accuracy assessment team.

- Difference between LACIE % wheat and ground truth % wheat
- Proportion training fields correctly identified for each class.
- Proportion test fields correctly identified for each class.

6.0. ITS EXPERIMENTAL CONTROLS

Adequate experimental controls are essential for attainment of the performance assessment objectives. Three major requirements are:

1) Restricted Access to Ground Truth Data. The AI personnel assigned to accuracy assessment must be precluded from assessing the ground truth data. This restriction is vital to insure that knowledge of actual ground conditions and is not directly or indirectly communicated between AI's, analysts, or others involved in making the classification. On the other hand the accuracy assessment evaluation personnel will have access to ground truth data.

2) Replication Independence. Procedures to insure independence of replications must be established. When a segment is to be replicated, neither the analyst nor the AI can have any previous knowledge of the data for the segments. Moreover, in performing the classification, they must avoid consultation with anyone who does have previous knowledge of data for the segment. Normal consultation with other individuals at their location is not only acceptable, but is recommended.

3) Realism. Full four pass classifications will be performed according to the same restrictions that would apply in an operational system. If data quality is poor, or some passes are missing, the segment should be processed as well as possible.

4) Segments will be worked according to the acquisition sequence. Acquisitions which are over and above the normal segment acquisition of one per biostage will not be used by the AI in choosing training or test fields.

(5) The analyst interpreters will be provided classification results and classification maps from each interpretation before continuing with the next.

7.0 ITS DATA ANALYSIS TECHNIQUES

7.1 Factorial Test Designs

These test designs, used for tests described in sections 4.1, 4.2, and 4.4 are specifically ordered to enable standard analyses of variance (ANOVA) techniques to be employed for data analysis. Computer programs, which are currently operational on the Univac-1110 onsite and incorporate these techniques, will be utilized.

Preliminary tests on the data, where possible, to determine the goodness of fit, homogeneity of sample variances, etc., will also be performed.

7.2 Statistical Analysis

Statistical techniques of regression, analysis of variance, test of significance, and confidence interval estimation will be utilized for data analysis and making statistical inference.

8.0 REPORTING MILESTONES

The first report will be issued on December 15, 1975, for the Intensive Test Site evaluation. Interim reports for statusing and tracking purposes will be issued on a monthly basis. The final report will be issued on March 15, 1976.

9.0 ITS ASSESSMENT OF DIFFERENCE BETWEEN AI'S

Three intensive test sites were acquired for each of the 4 biostages: 1975 (Oneida, Idaho) and 1976 (Franklin, Idaho) and 1969 (Toole, Montana). The remaining 10 AI's will interpret these segments (requiring 120 additional interpretations and 120 additional DPA runs).

Start: December 1, 1975   Complete: January 15, 1975

10.0 ITS SYSTEM BIAS AND VARIANCE

The test described in section 4.1, AI vs. Biophase vs. Geography processed through the LACIE system should determine the bias and repeatability (variance) of the operational segment oriented portion of the system. The determination of bias may enable the introduction of the proper correction factor at some TBD level (to the normal LACIE sample segments) prior to including the data in the aggregation model. The determination of the variance for the ITS, and then the extrapolation to the LACIE sample segments will establish the degree of error that can be associated with the mechanics of the system, and its contribution to the overall average estimating procedure, as opposed to sampling and aggregation variance.

The accuracy of the LACIE system will be verified in mixed spring and winter wheat areas for both: 1) winter wheat area acquired using spring wheat biowindows (Montana and South Dakota ITS) and 2) spring wheat area acquired using winter wheat biowindows (Washington, Idaho ITS).

Start December 1, 1975 Complete January 15, 1976

11.0 SAMPLING AND AGGREGATION

After the degree of accuracy of the system's classification of the ITS has been established, extrapolation upwards to estimate the accuracy associated with the entire segment may be performed. Then, based upon the study described in section 4.0, the ITS accuracy parameters may be transferred to the LACIE sample segments that are located in the same area. The aggregation model will be evaluated with respect to item K-1 of the aggregation issue defined by the RID analysis team (AES Acreage Review, August 11-15, 1975).

For the purpose of testing the adequacy of the sampling, the 1974 SRS county wheat areas will be used as an input to the aggregation model in place of the CAMS output. Aggregations will be made in this manner and standard statistics will be calculated for each of the Great Plains states.

12.0 WITHIN COUNTY VARIANCE DETERMINATION

To determine the sampling error, the within county variance will be determined for one MSS frame in Kansas, one MSS frame in Nebraska, one in North Dakota, one area in Saskatchewan, one area in the Ukraine, and two other TBD areas in the USSR. The areas chosen as highest priority shall be crop reporting districts with all sample groups represented (i.e., groups I, II, and III). They are:

Kansas crop reporting districts 2, 6, 7

North Dakota crop reporting districts 4, 9

Nebraska crop reporting district 1

This within county variance information will be utilized to assess the level at which the county variances can be assumed homogenous and to determine the effectiveness of the regression technique in the variance estimation.

In each case imagery will be chosen that currently exists in-house that was acquired when confusion crops are minimized. A grid of 330 5x6 n.mi segments will be overlayed on the 9x9 in. color transparency and every 5x6 n.mi will be interpreted as to percent wheat to the nearest 1.0% within each agricultural segment in the pseudo county. The grid will be aligned according to the grid used in the original sample selection and only those segments denoted as "agricultural" in the original sample selection shall be interpreted. Only counties that are completely within the chosen frame shall be interpreted. It is estimated that this will take 14 man-days to complete and will be done by AI's assigned to RTEB. When this test is complete, it will be used to develop techniques for doing similar tests on a broader scale for Phase II of LACIE.

13.0 SIMULATION STUDIES

13.1 Acquisition

Study, through Monte Carlo simulation techniques, the effects of various levels of missing data on the aggregation model. For example, if the satellite acquires only 60% of the segments of the country, but acquire 90% of the segments in the major areas.

13.2 Estimation Error

Based upon the results obtained from the test described in 4.1, simulate, using the LACIE IOC Error model, the effects of the bias and variance on the various levels of reports generated by the aggregation model. Both random and consistent bias will be investigated. Determine the variance in the state estimates that will still allow the 90-90 LACIE goal to be achieved at the national level.

13.3 Combined Error

Using the results of both 13.1 and 13.2, simulate the expected error of the aggregation estimate for all cases of biostage, level, region, country, monthly report, etc.

14.0 CROP CALENDAR VERIFICATION AND THE EFFECT OF CROP CALENDAR ERRORS ON CLASSIFICATION

Since the evaluation of a crop calendar does not yield a specific right or wrong situation, but rather a degree of correctness and a spread of data, its effect on correct classification of wheat/non-wheat is not assumed to be a simple comparative method of evaluation. However, one would expect a correlative type of relationship and, therefore, initial verification techniques will be based on regression/correlation methods.

There are three important areas to be tested in relation to the ACC (adjustable crop calendars):

1) Within an ITS or blind test within the Great Plains segment how do the ACC estimates compare with the actual (ground truth) wheat growth development stage (mean). If it does not match closely the following causes will be investigated: a) The site lies on the perimeter of the CRD. In this case, the ACC output for the adjacent CRD will be investigated; b) the site is at a different elevation than the mean elevation for the CRD; c) the agricultural practices are dissimilar from the rest of the CRD; d) model error; e) observer error.

   Action: YES

   Start November 1, 1975 End December 1, 1975

2) Determine the crop calendar error (in days) for each biostage and resulting CAMS performance degregation for one ITS in North Dakota and one ITS in Kansas (see attached Biostage vs. Julian Days in figures 1 and 2).

   Action: YES and AI and Accuracy Assessment Team

   Start December 1, 1975 End January 15, 1976

3) Determine the number of wheat subclasses that the AI chooses as a function of site and wheat/confusion crop calendars.

Action: AI

Start: January 15, 1976 Complete: January 20, 1976

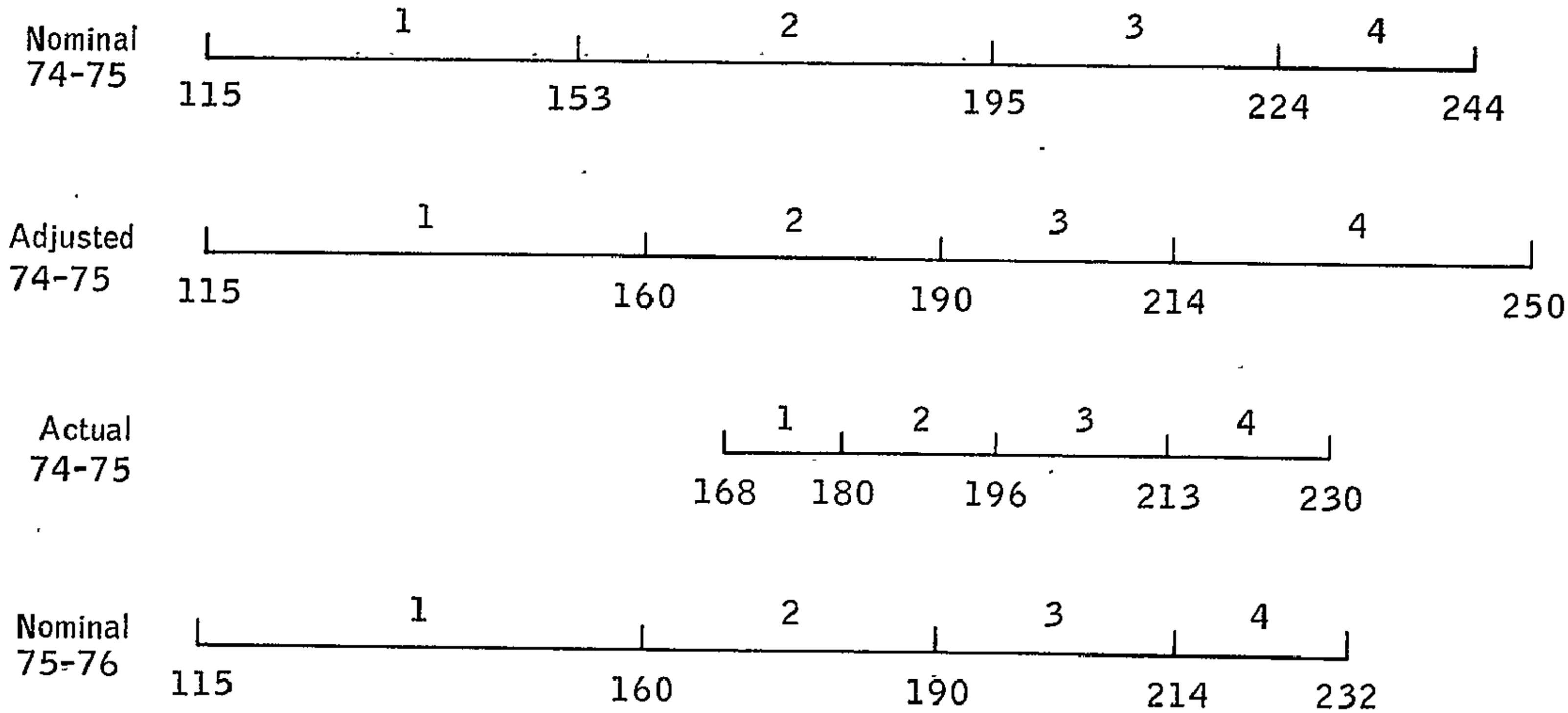


Figure 1.- Biostages versus Julian days - North Dakota CRD 5 spring wheat.

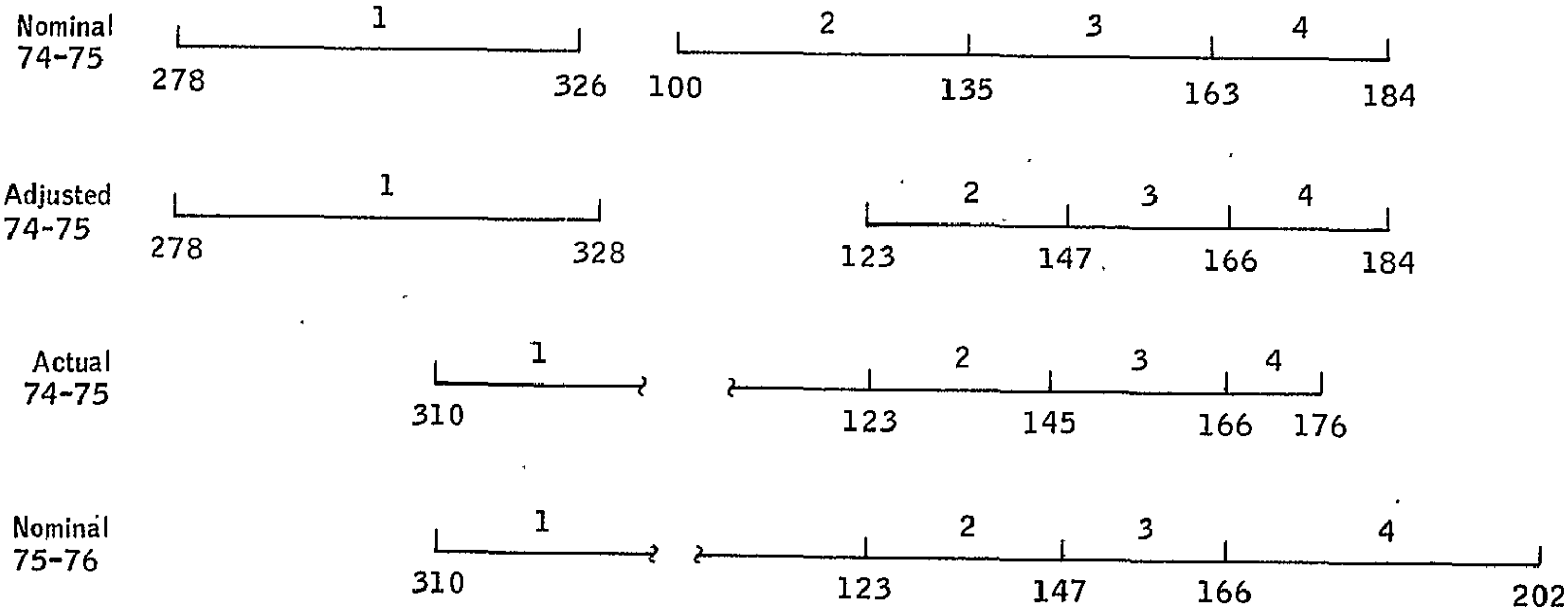


Figure 2.- Biostages versus Julian days - Kansas CRD 5.

15. BLIND TEST

15.0 INTRODUCTION

In order to determine the accuracy of the LACIE system over a set of regular LACIE segments, 30 segments in Montana and North Dakota with one or more Landsat 2 MSS acquisition were chosen to be ground truthed. The segments were also chosen so that all the operational AI were represented so that an even balanced factorial design could be used. The segments chosen, the biostages acquired, and how the ground truth was collected are shown in Table III. Table IV shows the crop key used by the ground truth teams. Color IR Aircraft photography at 1 to 24,000 was obtained in mid-August 1975 and was used in the following week as a base map for annotation by the ground truth teams. Three teams of two persons each were deployed for ten days in mid-August to gather ground truth using light aircraft and limited backup work on the ground.

15.1 Ground Truth Data Collection

The ground truth were collected from the ground by having each of the three teams deploy first to an intensive test site and use existing photo coverage and ground truth to calibrate with. Next each team drove to a regular LACIE test site and investigated each field that had an adjacent road. The next day each team flew a light aircraft over their regular LACIE site and the intensive test site and annotated the field identification on the week-old aircraft imagery which was obtained using the Zeiss 6 in. at about 20K ft. The tasks would have been futile if it were not for this excellent fresh imagery. On following days the teams would alternate using aircraft at 1500 ft. and doing backup field work

TABLE III- BLIND SITES

COUNTY/STATE	SAMPLE SEGMENT	BIOSTAGE ACQUIRED	GROUND TRUTH ACQUIRED	AIRCRAFT COVERAGE
McHenry/ND	1613	1,2,3	A	100%
McHenry/ND	1612	1,2,3	A	South 90%
Adams/ND	1646	1,2,3	A	100%
Rolette/ND	1615	1,2	A	East 95%
Ramsey/ND	1622	1,2,3,4	A	100%
Ward/ND	1606	1,2,4	A&G	100%
Ward/ND	1605	1,2,4	G&A	100%
Williams/ND	1608	1,2,4	G&A	100%
McKenzie/ND	1627	1,2,3	G&A	100%
Morton/ND	1656	1,2,3,4	A	100%
Richland/ND	1663	1,2	A	100%
Kidder/ND	1634	1,2,3	A	100%
McKenzie/ND	1626	1,2,3	G	100%
Fallon/MT	1555	1,2	G&A	100%
Richland/MT	1540	1,2	G	100%
Carter/MT	1553	1,2,3	A	South 70%
Sheridan/ND	1635	1,2	G&A	100%
McCone/MT	1538	1,2,3	G&A	100%
Sargent/ND	1664	1,2	A	100%
McClean/ND	1629	1,2,3	A&G(part G)	100%
Hettinger/ND	1650	1,2,3	A	100%
Mercer/ND	1630	1,2,4	A	100%
Oliver/ND	1631	1,2,3	A	East 70%
Dawson/MT	1534	1,2,4	A&G	100%
Yellowstone/MT	1552	1,2	A	100%
Sheridan/MT	1544	1,2	G&A	South 90%
Burleigh/ND	1653	1,2,3	A	South 85%
Sheridan/MT	1543	1,2,3	A&G	South 85%
Bottineau/ND	1610	1,2,3,4	A&G	100%

A = Ground truth acquired by air
G = Ground truth acquired by ground

TABLE IV.- CROP KEY

Montana and North Dakota Ground Truth

of Regular LACIE Segments for Accuracy Assessment

Key	Crop
W	Wheat with awns
WA	Wheat awnless
WW	Wheat windrowed
WH	Wheat harvested
SMG	Small grains
F	Fallow
G	Grass (not cut for hay and no fence)
H	Hay (any visable signs of hay activities)
A	Alfalfa
P	Pasture
C	Corn
SF	Safflower
SU	Sunflower
SG	Sudan grass
SR	Sorghum
SY	Soybeans
SB	Sugar beets
FX	Flax
M	Mustard
T	Trees

TABLE IV.- Concluded

Key	Crop
R	Rye
B	Barley
X	Homestead - nonagricultural
BN	Beans
O/W	Oats/wheat mix
W/O	Wheat/oats mix
[ ]	Interpretation used around the code and not the fields
O	Oats

1. Use standard key for all identification.

2. Use Mylar for all coding in ink.

3. Label each photo on back side and Mylar in upper right and left corners for appropriate east and west half of site.

TABLE IV.- EVALUATION FORM

SEG. NO.________ ACQ. DATE___________ BIO PHASE____________
LOCATION____________________

A1 CODE	GROUND/AR TRUTH	#INCORR PIXELS	TOTAL PIXELS	% CORRECT	CORRECTED LABEL
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					
☐☐☐ ☐☐☐					

on the ground. Each team averaged one site per day. Hand-held photographs were used for reference at times.

15.2 Scope

The blind test will be used to test the following areas (in priority):

1) What are the error components of the CAMS per-segment wheat proportion estimate caused by AI training field selection, labeling, and DPA processing?

2) What are the relative sizes of the LACIE error components caused by sampling and the per-segment wheat proportion estimate?

3) It will be determined if there exists a significant difference in the accuracies as a function of:
   a) Multitemporal vs. single pass.
   b) Early biostage vs. early and late biostages.
   c) SM and U.
   d) Strip-fallow vs. continuous crop.

4) Are there significant differences between the accuracy of analysis of different AI's on regulat LACIE segments? The blind sites were chosen so as to have all the AI's represented.

5) Are the intensive test sites representative of the regular LACIE segments?

15.3 Analyst Interpreter/Ground Truth Comparisons

North Dakota/Montana

Action: AI specialist

Time: 1 man-day per segment

- Select segment.
- Pull AI packet.

- Extract temporal crop interpretation form.
- Extract product 1 imagery.
- Provide temporary storage.

- Pull ground truth packet, 1:24,000 G.T. packet photos, 1:48,000 photos, hand-held photography.
- Identify ground classifications for each training fields selected (all biophases).
- Document ground truth descriptions for each training field, test field, and designated other area, on evaluation form.
- Record number of pixels that are not equivalent to the AI descriptor codes.

  e.g. - 100 pixel wheat field (AI interpretation) when compared against ground truth shows that field is actually 80 pixels of wheat and 20 pixels of barley. Twenty pixels would be recorded in "number of incorrect pixels" column of evaluation form.
- Submit evaluation form to accuracy assessment team.

15.4 Determination of Total Proportion of Wheat (or Small Grains) Within a Segment

Action: The analyst interpreter who did the original interpretation.

Time: 4 man-hours per segment

Equipment: H. Dell Foster Digitizer

- Plot the ERTS scene product 1 boundary on the 1:48,000 photography.
- Using the area mode feature (zero scale factor) of the H. Dell Foster, measure the segment area in one thousandths of a square inch on the 1:48,000 scale photography.

- Next, measure the area of each wheat field (or small grain field in the case of aerial observations) on the 1:48,000 scale photography.
- Divide the sum of the individual wheat/small grain fields by the total area of the segment (thousands of an inch). The result is the percent of wheat or small grain within the segment.
- Submit result to AI specialist.

15.5 Provide "Correct" Labels for the Biophase 4 (or Other Biophase as Designated by the Accuracy Assessment Team) AI Selected Fields

Action: Analyst interpreter

Time: 4 hours per segment (single biophase)

Equipment: H. Dell Foster Digitizer and Purdue Lars Terminal

- Generate a "corrected" fields deaf deck on the H. Dell Foster.
- Perform the fields data conversion and edit on the Lars Purdue terminal.
- Prepare an ADP transmittal packet containing the "corrected" flap printout, Polaroid, and a xerox of the (corrected) temporal crop interpretation form and the evaluation form.
- Submit transmittal packet to the AI specialist.

15.6 Reselect Training Fields (Single Biophase) Using All Available Ground Truth

Action: Analyst interpreter

Time: 4 hours per segment (single biophase)

Equipment: H. Dell Foster and Purdue Lars Terminal

- Select optimum training fields using ground/aerial truth and all available acquisition imagery.
- Prepare fields overlay and temporal crop interpretation form.
- Review optimum training field selections with DPA representative.
- Generata a "optimum" training fields deaf deck on the H. Dell Foster.
- Perform the fields data conversion and edit on the Lars Purdue Terminal.
- Prepare an ADP transmittal packet containing the "optimum fields" flap printout, a Polaroid, and a xerox of the "optimum fields" temporal crop interpretation form.
- Submit transmittal packet to the AI specialist.

15.7 Status and Tracking

Action: AI specialist

Time: 1 hour per segment transmittal

- Perform edit of all transmitted data.
- Ensure errors are corrected.
- Transmit materials to accuracy assessment team (or as directed).
- Maintain "real time" status on a daily basis.

15.8 Training and Test Field Accuracies

Action: Accuracy assessment team

Time: 4 hours per segment

- Using ADP classification summary reports, record the total number of pixels for each training/test field on the evaluation form (Table IV).
- Compute the correct percent value for each AI selected field and record on the evaluation form.
- Record the correct label for each field in the corrected label column

of the evaluation form.

- Return the evaluation form(s) to the assigned AI specialist.

15.9 DPA Procedures

1) The first method of classifying the 30 "blind" segments is that the fields chosen for production will be relabeled by the AI according to the ground truth and the segment will run interactively on ERIPS after the normal operational classification is complete for all of LACIE Phase I.

The DPA will run the segment according to normal procedures with the addition of redefining the subclasses of the fields that were incorrectly labeled. That is:

a) Make a merged image using one pass (the pass that the fields were defined on).
b) Retrieve fields from the data base.
c) Redefine the fields according to AI's instructions.
d) Compute the training statistics.
e) Classify using four channels.
f) Make a class summary and a class map.

This will require about 45 minutes of interactive time plus 20 minutes necessary for doing the calculations by hand (for each segment). It would facilitate the DPA to have the AI update the data base via cards so that segment could be done in batch and the calculations could be done on the Univac 1110.

2) The second method for classifying the segments requires that the AI define new fields according to the ground truth. These will be input to the system via cards. The DPA will run a normal production batch job using subclass statistics of the new fields. The calculations can be done on the Univac 1110.

This will require about 5 minutes of DPA time per segment.

3). Each classification will be evaluated according to the evaluation procedures described in CAMS Detailed Analysis Procedures for the LACIE Operations (Section 2.3.6).

The results from the three classifications (production, method I, and Method II) will be compared using these same criterion.

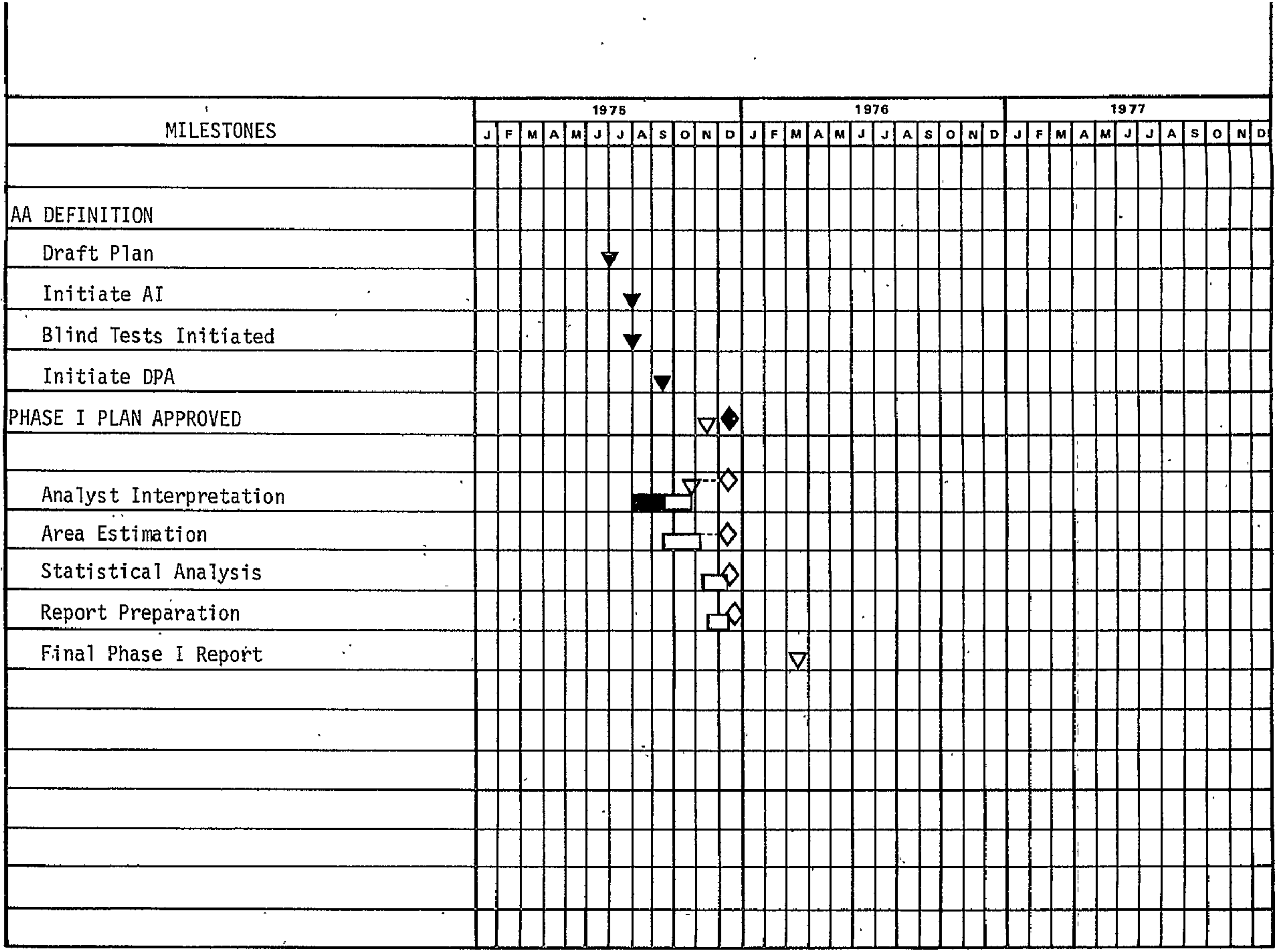
Action: DPA analysts

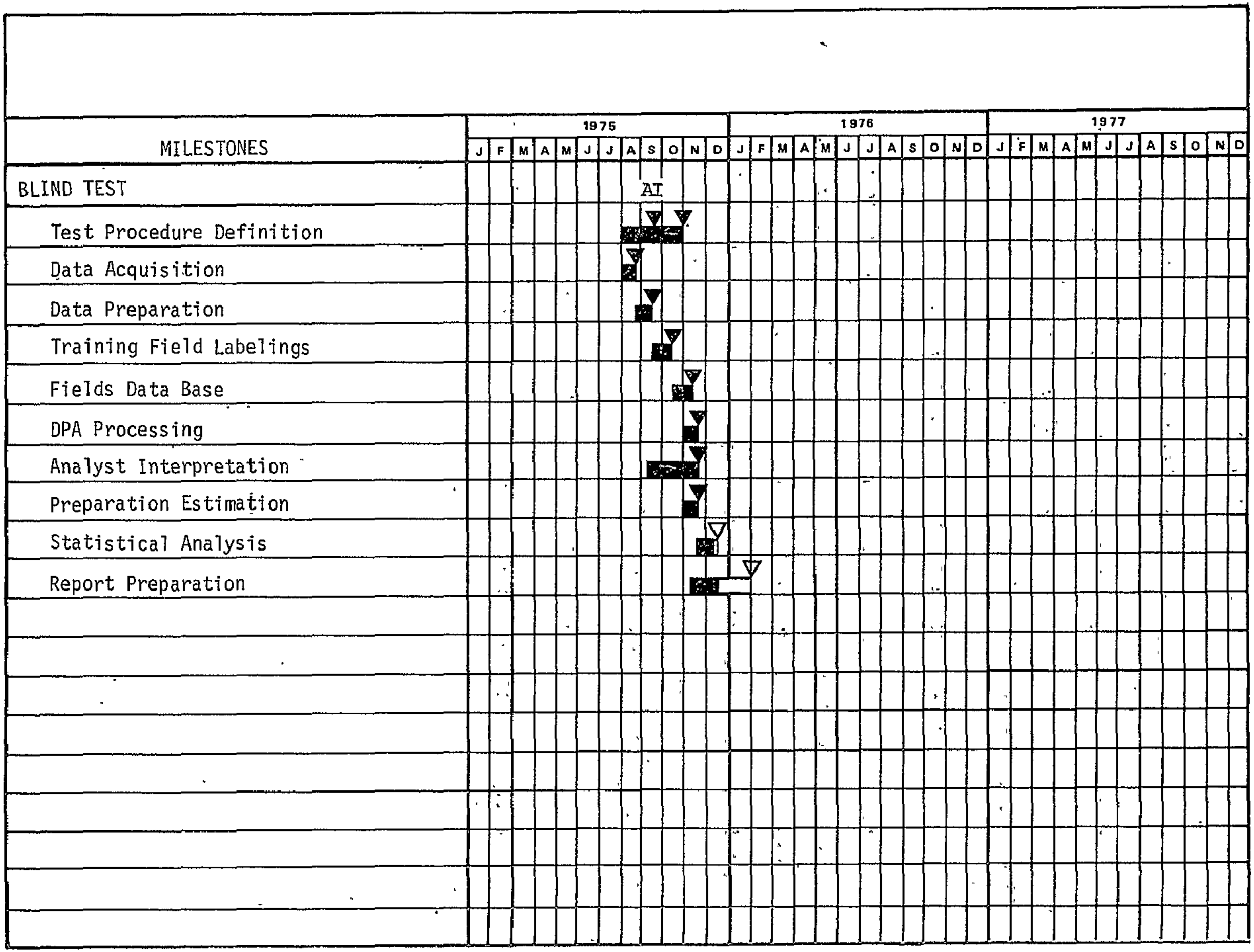
Computer time: 10 minutes/segment

DPA time: 1 hour/segment

15.10 North Dakota Aggregations

Any aggregations performed in operations in North Dakota will be double checked using the North Dakota blind site data.


MILESTONES
1975
1976
1977
J F M A M J J A S O N D
AA DEFINITION
Draft Plan
Initiate AI
Blind Tests Initiated
Initiate DPA
PHASE I PLAN APPROVED
Analyst Interpretation
Area Estimation
Statistical Analysis
Report Preparation
Final Phase I Report


MILESTONES
1975
1976
1977
J F M A M J J A S O N D
BLIND TEST
AT
Test Procedure Definition
Data Acquisition
Data Preparation
Training Field Labelings
Fields Data Base
DPA Processing
Analyst Interpretation
Preparation Estimation
Statistical Analysis
Report Preparation

16.0 AI IMPACTS

16.1 Operations Throughput

Priorities must be assigned to the accuracy assessment effort required ITS photo interpretation. Production/Operations segments must not be allowed to constantly receive higher priorities.

16.2 Constrained Consultant Population

Normally, if an AI were to have a problem selecting training fields, he would be able to consult with any other AI. But for the purpose of this experiment, he will only be able to consult with the other AI's in his immediate vicinity according to Phase I operational procedures. This is required to insure that against an AI who has knowledge of the ground truth for an ITS not being used as a consultant.

16.3 Changed Procedures

No changes are to be implemented without the approval of the AA/CCT personnel.

PRECEDING PAGE BLANK NOT FILMED

17.0 Method for Calculating Accuracy of CAS Aggregations

17.1 Standard Statistics[1]

The standard statistics for area estimation and aggregation are composed of standard deviation (σ), coefficient of variation (CV), 90% confidence limits (CL), and probability of achieving 10% error or less (Ω). Each statistic is a function of variance (V) as follows:

a) $\sigma_s = \sqrt{V_s}$ where s = any area element (3.1.2-1)

b) $CV_s = \frac{\sigma_s}{A_s}$ where A_s = area estimate of the sTH element (3.1.2-2)

c) $CL_s^{(+)} = A_s + (1.645)\ \sigma_s$ (3.1.2-3a)

$CL_s^{(-)} = A_s - (1.645)\ \sigma_s$ (3.1.2-3b)

d) $\Omega_s = \text{erf}\left(\frac{1}{\sqrt{2}}\ \frac{1}{10CV_s}\right)$ (3.1.2-4)

where $\text{erf}\left(\frac{\theta}{\sqrt{2}}\right)$ represents the error function associated with normal probability curve and θ is the standardized random variable.

---

[1] Each area aggregation to the zone (state), region, or country level will have the standard statistics calculated.

17.2 Variance Estimation

To calculate/estimate within substrata variance, regress CAMS wheat proportion output for segments in a state to their corresponding 1969 historical county wheat proportions. Assuming that within county variance is same for all substrata in a zone, consider the residual mean square for an estimate of the within substrata variance. Once this is estimated use the variance formula (see CAS' Requirement Document) to obtain variance estimates for group I, group II, and group III at the strata level.

PROCEDURE

Suppose

$$E[y_{ij}] = a + bx_i$$

where

y_{ij} = CAMS/LACIE wheat proportion for j^{th} segment of i^{th} substrata

x_i = 1969 historical wheat proportion for i^{th} substrata

$E[y_{ij}]$ = Expected value of y_{ij}.

For a set of observed data, y_{ij}, $j = 1, 2, \ldots, n_i$ and $i = 1, 2, \ldots m$, in a zone, using simple regression technique, obtain the best fit

$$\hat{y}_i = \hat{a} + \hat{b}\, x_i \tag{1}$$

and determine the residual mean square

$$\mathcal{S}^2 = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \hat{y}_i)}{\left(\sum_{i=1}^{m} n_i - 2\right)} \tag{2}$$

Let σ_i^2 be the within substrata variance for the i^{th} substrata in a zone. For its estimate, let

$$\hat{\sigma}_i^2 = \mathcal{S}^2, \quad i = 1, 2, \ldots, m.$$

Now variance estimates at CRD and state levels are obtained using the following formulas:

CRD (strata)

Group I:

$$V = \delta^2 \sum_{j=1}^{M} \frac{N_i - n_i}{n_i} N_j R_j^2 \tag{3}$$

where

M = number of counties for which at least one segment is acquired.

N_j = number of segments in j^{th} pseudo county for which at least one sample segment is available.

n_j = number of sample segments in j^{th} psuedo county

$$R_j = \frac{\text{Area in } j^{th} \text{ county}}{\text{Number of segments in } j^{th} \text{ gross pseudo county}}$$

Group II:

The interim procedure is to treat group II counties as group III. Then the variance estimate of CRD/strata acreage estimate is

$$\hat{\sigma}^2 = \left(1 + \frac{w_2}{w_1}\right)^2 V \tag{4}$$

where

w_1 = 1969 wheat acreage for group I counties for which at least one segment is available.

w_2 = 1969 wheat acreage for group II, III and group I counties for which no sample segment is available.

CRD for which no segment processed

$$\hat{\sigma}^2 = \left(\frac{W}{\sum_{j=1}^{k} W_j}\right)^2 \sum_{j=1}^{k} \sigma_j^2 \tag{5}$$

where

W = 1969 wheat acreage for the CRD for which no segment processed.

W_j = 1969 wheat acreage for the j^{th} CRD for which segments processed and aggregation made.

σ_j^2 = Variance estimate given by (4) for the j^{th} CRD for which segments processed and aggregation made.

State (Zone)

$$\text{Variance} = \left(1 + \frac{\sum_{i=k+1}^{L} W_i}{\sum_{j=1}^{k} W_j}\right)^2 \sum_{j=1} \hat{\sigma}_j^2$$

where

W_{k+1}, W_{k+2}, - - - , W_L are 1969 wheat acreages for the (k+1)th, (k+2)th, - - - - , Lth CRD, respectively, for which no segment processed.

18.0 PHASE I AGGREGATION ACCURACY

Task 1: Define and conduct an effort to define the variances in SRS estimates at state and national levels, determine the SRS methods for determining "planted acres" and "acres for harvest," and define where the LACIE estimates and SRS estimates are observing different quantities. Using standard statistics, estimate the expected LACIE variances at state and national levels, and estimate the expected discrepancy between the LACIE estimates and SRS estimates of "planted acres" and "acres for harvest" at state and national levels.