8.1-10031

AgRISTARS

FC-L0-00609
JSC-16364

NASA CR-
160777

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

Foreign Commodity Production Forecasting

June 1980

SUPPLEMENTAL U.S./CANADA WHEAT AND BARLEY EXPLORATORY EXPERIMENT IMPLEMENTATION PLAN: EVALUATION OF A PROCEDURE 1A TECHNOLOGY

(E81-10031) SUPPLEMENTAL US/CANADA WHEAT AND BARLEY EXPLORATORY EXPERIMENT IMPLEMENTATION PLAN: EVALUATION OF A PROCEDURE 1A TECHNOLOGY Interim Report (Lockheed Engineering and Management) 44 p HC A03/MF A01

N81-12513

Unclas

G3/43 00031

NASA

Lyndon B. Johnson Space Center
Houston, Texas 77058

1. Report No. JSC-16364	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Supplemental U.S./Canada Wheat and Barley Exploratory Experiment Implementation Plan: Evaluation of a Procedure 1A Technology		5. Report Date June 1980
		6. Performing Organization Code SF-4
7. Author(s) Not applicable		8. Performing Organization Report No. FC-L0-00609
9. Performing Organization Name and Address Crop Applications Branch Earth Observations Division Space and Life Science Directorate National Aeronautics and Space Administration		10. Work Unit No.
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Lyndon B. Johnson Space Center Houston, Texas 77058		13. Type of Report and Period Covered Technical Report
		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This document is a detailed plan for a supplemental experiment to evaluate a new sample allocation technique for selecting picture elements from remotely sensed multispectral imagery for labeling in connection with a new crop proportion estimation technique. The method of evaluating an improved allocation and proportion estimation technique is also provided.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
labeling experiment Procedure 1A crop proportion estimation technique sample allocation	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	44	

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

JSC Form 1424 (Rev Nov 75) NASA — JSC

FC-LO-00609
JSC-16364

SUPPLEMENTAL U.S./CANADA WHEAT AND BARLEY EXPLORATORY EXPERIMENT IMPLEMENTATION PLAN: EVALUATION OF A PROCEDURE 1A TECHNOLOGY

This document describes a supplemental experiment to be conducted in support of the U.S./Canada Wheat and Barley Exploratory Experiment of the Foreign Commodity Production Forecasting Project.

J. L. Dragg
FCPF Project Manager

June 1980

LEMSCO-15042

CONTENTS

Section	Page

TABLES

FIGURES

ACRONYMS

AA	Acuracy Assessment
AI	Analyst-Interpreter
AgRISTARS	Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing
AMOEBA	clustering algorithm
CLASSY	clustering algorithm
CY	crop year
DAHP	Data Acquisition, Handling, and Processing
ERIPS	Earth Resources Interactive Processing System
FCPF	Foreign Commodity Production Forecasting
FY	fiscal year
ISOCLS	Iterative Self-Organizing Clustering System
JTCT	Joint Technical Coordination Team
LACIE	Large Area Crop Inventory Experiment
LARS	Laboratory for Application of Remote Sensing
MSE	mean square error
MYE	man-year equivalent
P1	Procedure 1
PDP	Programmed Data Processor
pixel	picture element
PTRR	Preliminary Technical Review Report
SR	Supporting Research
TY	Transition Year
USGP	U.S. Great Plains

1. INTRODUCTION

1.1 EXPERIMENT BACKGROUND

The Foreign Commodity Production Forecasting (FCPF) Implementation Plan (ref. 1), dated January 15, 1980, provides for a category 3 (test and evaluation) experiment for U.S./Canada wheat and barley scheduled for completion in fiscal year 1980 (FY80). A wheat and barley labeling experiment plan (ref. 2) was developed in late January 1980 to support that exploratory experiment. This document is a detailed plan for a supplemental experiment to evaluate a new sample allocation technique for selecting picture elements (pixels) to be labeled and a new proportion estimation technique. Crop labels provided by the previously planned labeling experiment will be used for this evaluation. The allocation and proportion estimation technique was recommended by the Supporting Research (SR) Project of the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) Program and provides a basically new concept (refs. 3 and 4) that may be used in the replacement of Procedure 1 (P1) [ref. 5].

P1 was tested thoroughly during the 1977 and 1978 crop years (CY). Evaluations of these tests results (refs. 6, 7, 8, and 9) showed that relative count proportion estimate from a stratified scene using supervised clustering and a subsequent maximum likelihood classification did not significantly improve upon the proportion estimate obtained using only a labeled random sample of pixels. Thus, research was initiated by the SR Project to develop an improved allocation and proportion estimation technique to replace P1.

The recommended proportion estimation technique uses the CLASSY clustering algorithm (refs. 10, 11, and 12). The SR Project was responsible for the testing of three clustering algorithms; CLASSY, AMOEBA (ref. 13), and Iterative Self-Organizing Clustering System (ISOCLS) which is defined in reference 14 and is the one used in P1. Results show that all three performed at about the same level for variance reduction over simple random sampling and cluster purity (ref. 3); however, CLASSY was deemed superior for the proposed estimation technique since it produced fewer clusters. On the average, CLASSY produced about one-half as many clusters as AMOEBA, which in turn produced one-half as many clusters as ISOCLS. Theoretically then, with the fewer numbers of clusters, CLASSY would require fewer dots to be labeled in order to obtain proportion estimates with accuracy comparable to that using either AMOEBA or ISOCLS.

A preliminary report (ref. 3) by the SR Project indicated that for a given number of labeled pixels, proportion estimates obtained using the recommended procedure and using CLASSY-defined strata were more accurate than the estimates obtained using P1 and using ISOCLS-defined strata. In order to determine whether the improvement was due to the use of CLASSY or the new recommended proportion estimation technique, the Joint Technical Coordination Team (JTCT) recommended that combinations of allocation and proportion estimation schemes (including those used in P1) also be tested using the CLASSY-defined strata. (See section 1.3.)

This plan presents the second in a series of experiments scheduled for the FCPF Project which addresses small grains, wheat, and barley. The first experiment (ref. 2) in the series is a test and evaluation of labeling techniques which are known as the Reformatted and Integrated Labeling Procedures (refs. 15 and 16, respectively). The experimental design will not be affected by the second experiment; however, crop labels from the labeling test are essential inputs to the experiment described in this plan. In addition to these inputs, ground-truth labels will also be utilized in testing and evaluating the techniques described in this plan. Since the pixel allocations and proportion estimates will be obtained through computer calculations, manpower expenditures will be low. Yield and aggregation activities are not a part of this experiment.

1.2 OBJECTIVES

The general objectives of the supplemental experiment are the following:

a. Evaluate a candidate technology for improved area estimation in follow-on FCPF Project exploratory and pilot experiments. This candidate technology will be referred to in this plan as P1A.

b. Develop recommendations for additional research, development, and tests required to incorporate an improved area estimation technology into follow-on FCPF Project exploratory and pilot experiments.

Specific objectives related to (a) in the general objectives follow.

a. To assess the suitability of the recommended allocation approach for improving labeling efficiency and accuracy

b. To determine if the number of dots required to be labeled on each segment can be minimized

c. To determine the effects on proportion estimation error due to the interaction of the labeling procedure with the sampling allocation of pixels within a segment

d. To determine if the recommended proportion estimation techniques can significantly reduce the mean square error (MSE) in proportion estimation with a minimal introduction of bias

e. To assess the reduction in variance over random sampling that can be achieved using the recommended approaches

Recommendations on whether to use one of the tested allocation and proportion estimation techniques in follow-on exploratory and pilot experiments will be made following evaluation of the results of this experiment. Recommendations for further research and development will also be developed.

1.3 APPROACH

The following combination of allocation and proportion estimation techniques will be tested and evaluated utilizing dots labeled in the U.S./Canada Wheat and Barley Labeling Experiment.

a. Random sample allocation using relative count to estimate crop proportions (Recommended by the JTCT)

b. Proportional sample allocation using relative count to estimate crop proportions (Recommended by the JTCT)

c. Proportional sample allocation using a Bayesian approach to estimate crop proportions (Recommended by the JTCT)

d. Sequential Bayesian method for sample allocation using a Bayesian approach to estimate crop proportions (Recommended by the SR Project as a candidate technology for P1A)

Crop labels obtained from the U.S./Canada Wheat and Barley Exploratory Labeling Experiment are limited to acquisitions over the spring wheat and barley sites in the U.S. Northern Great Plains (USGP) obtained during the CY79. Proportion estimates derived from these acquisitions will be made from a fixed number of dot labels (50) and with a variable number determined by the sequential Bayesian allocation method. Some labels may be required by the Bayesian sequential allocation procedure in addition to the 209-grid dots that are usually labeled. In this event, the location of the additional dots will be furnished to the analyst in advance so that no change in the labeling procedures will be noticed.

In working this proportion estimation experiment, it is intended that the sequential Bayesian allocation be applied after the dot labeling has been performed. This procedure (for this experiment only) deviates from the manner in which it would normally be expected to operate. As a consequence, it is possible that the analyst versus allocation procedure interactions, which might occur in an operational mode, will not be present in this experiment.

Stratification of the scene will be accomplished using the unsupervised clustering algorithm CLASSY (refs. 10, 11, and 12). Acquisition dates selected for machine processing with CLASSY will be those dates designated by the Reformatted Labeling Procedure in reference 15. In the event segments are not processable using this Reformatted Labeling Procedure, then acquisitions for those segments will be selected using the Detailed Analysis Procedures given in reference 17.

As a result of evaluating the four proposed techniques using labels from two labeling procedures and ground truth, a minimum of 24 estimates for each segment will be obtained as summarized in table 1-1.

1.4 COMPONENT ROLES

Component elements having major roles in this experiment are Classification, Experiment Design, Accuracy Assessment (AA), and Data Acquisition, Handling, and Processing (DAHP). No aggregations of acreage or yield are planned for this experiment.

Component implementation planning for this supplemental exploratory experiment is described in subsequent sections followed by summaries of (a) data and system requirements, (b) resources, and (c) an integrated experiment schedule.

Component responsibilities are stated in the FCPF Implementation Plan (ref. 1). Overall conformance to the technical aspects of this experiment will be monitored by a representative of the experiment design component.

1.5 PRODUCTS, REPORTS, AND DOCUMENTATION

Each of the technical components will interchange and produce products as specified in the individual component sections. Results from evaluations will be reported on a component basis to be compiled into a Preliminary Technical Review Report (PTRR). Formal documentation of experiment results and recommendations for future pilot experiments and further development will follow the preliminary report.

TABLE 1-1.- PROPORTION ESTIMATES OBTAINED PER SEGMENT

Proportion estimation techniques		Number of dots	Reformatted Labeling Procedure	Integrated Labeling Procedure	Ground truth
Allocation	Estimator				
Random sampling	Relative count	Fixed	--	--	--
		Variable	--	--	--
Proportional allocation	Relative count	Fixed	--	--	--
		Variable	--	--	--
Proportional allocation	Bayesian estimation	Fixed	--	--	--
		Variable	--	--	--
Sequential allocation	Bayesian estimation	Fixed	--	--	--
		Variable	--	--	--

6

2. EXPERIMENT DESIGN

2.1 BACKGROUND AND ISSUES

During the Large Area Crop Inventory Experiment (LACIE) and LACIE Transition Year (TY), several different classification approaches were tested for more accurate and efficient methods of estimating small grain acreage (ref. 5). Over a period of time, improvements were implemented as they were identified.

The approach utilized during Phases I and II of the LACIE Project was a method by which the analyst purposively selected fields as training samples. These samples were assumed to be representative of all the spectral subclasses within the segment. Each spectral subclass was assumed to have a multivariate normal distribution. These samples were used to estimate the means and covariances matrices for each subclass. With these statistics, maximum likelihood classification was performed. The small grain proportion estimates were obtained by the pixel count method.

In a highly automative environment as the LACIE, which required production estimates at a regular interval, the subjective processing approach proved to be unsatisfactory in both accuracy and efficiency. Manual interpretation required approximately 12 hours to complete the fields selection and labeling process. The procedure was more adaptable to areas with large agriculture fields; technical difficulties were encountered when implemented in areas with smaller fields.

Motivated by these and other problems experienced with the LACIE Phases I and II design, a second proportion estimation method, P1 (ref. 5), was constructed from the supporting research to remove or reduce these deficiencies. The use of P1 began in LACIE Phase III and has continued with minor changes through the present time. The main features of the P1 design were that a first set of randomly selected samples (pixels) were labeled and used to start a nearest neighbor clustering algorithm. The cluster statistics produced were the training data for the maximum likelihood classification of the scene into two- or three-class strata. Then stratified proportion estimates were made from a second set of labeled random samples which considered the classes as strata.

The P1 design proved to be a significant improvement in terms of estimation accuracy and efficiency. Some areas of improvement observed were (a) a shift toward the concept of an unbiased estimator, (b) multitemporal analysis and machine processing capability, and (c) the ability to extract and evaluate internal components of the procedure design. With LACIE Phases I and II procedures, it was difficult to distinguish the effects of analyst performance from classification performance on proportion estimation. With P1, several evaluations of classification performance became possible and were conducted. Results from the performance evaluation studies of the maximum likelihood classifier in P1 indicated a definite need for a more advanced approach. Findings indicated (refs. 6, 7, 8, and 9) that the classification portion of P1 did not show an improvement over simply using the labeled random sample of pixels for proportion estimation (even though analyst contact time was significantly reduced). Using an unsupervised clustering routine to perform stratification was shown to produce comparable results (ref. 3). With this in mind, several aspects of an improved procedure was researched. One of the techniques developed involved sequential sampling with a Bayesian estimation of crop proportions (ref. 3 and 4). This method is felt to have very strong potential for revising the current method of estimating small grain proportions. Among the proposed advantages cited are: (a) stratification without the need of the first set of labeled training samples, (b) lower MSE in the proportion estimation, and (c) improved labeling accuracy.

The Bayesian sequential allocation approach gives consideration to a prior distribution of cluster purities (appropriately modeled). In a small preliminary development test, it produced a lower MSE with a sample selection that increased labeling accuracy.

This supplemental experiment is intended to study the three basic subcomponents of the Bayesian sequential allocation procedure that make it different from P1. These subcomponents are: (a) the CLASSY clustering algorithm (refs. 10, 11, and 12) as a means of providing an unsupervised stratification of the segment [P1 used a supervised clustering approach, ISOCLS], (b) the Bayesian proportion estimator as an improvement over the relative count proportion estimator of P1, and (c) the allocation of data sequentially rather than proportionally.

2.2 OBJECTIVES

The general objectives of this supplemental experiment are:

a. To evaluate proportion estimation techniques as a candidate technology for Procedure 1A in preparation for its use in the area estimation of small grains for future FCPF Project exploratory and pilot experiments

b. To develop recommendations for additional research, development, and tests required to incorporate into follow-on FCPF Project exploratory and pilot experiments

Based on the results, it is anticipated that the current technology will be enhanced for small grains applications. The main technical issue to be studied is how much improvement in segment proportion estimation for small grains can result from using the techniques of stratified random sampling, Bayesian estimation, and sequential allocation.

Specific objectives have been established to support the general objectives listed above. They are:

a. To assess the suitability of using Bayesian sequential allocation to improve labeling efficiency and accuracy

b. To determine if the internal MSE threshold of the Bayesian sequential allocation method of estimating can be used to minimize the labeling process in each segment

c. To examine the effects of crop proportion estimation error due to the interaction of the labeling procedure with the sampling allocation of pixels

d. To determine if Bayesian estimation can reduce the MSE in proportion estimation with a minimal introduction of bias for each labeling procedure

9

e. To assess and determine if a reduction in variance over random sampling can be achieved using proportional allocation with CLASSY-defined strata

Findings from the supplemental experiment should aid in determining what technology should be used in the AgRISTARS/FCPF Project for the 1981 U.S./Canada Wheat and Barley Pilot Experiment.

2.3 TECHNICAL APPROACH

To assess the feasibility and advantages of the proposed improvements to the small grains proportion estimation technology, four combinations of sample allocation and proportion estimation techniques will be applied utilizing the pixel labels from the U.S./Canada Wheat and Barley Exploratory Labeling Experiment (ref. 2) Labels will be available from three sources: the Reformatted Labeling Procedure, the Integrated Labeling Procedure, and labels from digitized ground truth.

Three components of the procedure recommended by the SR Project are sequential allocation of pixels to be labeled, stratification using CLASSY, and Bayesian proportion estimation. The effects of each component will be tested and evaluated separately. Proportion estimates based on a labeled random sample of pixels will be used as the standard for comparison. To evaluate the effects of stratification, relative count estimates will be made with sample allocation proportion to the size of the CLASSY clusters and compared with the random sample estimates. To evaluate the effects of Bayesian proportion estimation, Bayesian proportion estimates will be made from the proportionally allocated pixel sample and compared with the stratified relative count estimates. To evaluate the added effects of Bayesian sequential allocation, Bayesian proportion estimates will be produced from labeled pixels sequentially sampled from the CLASSY clusters and compared with the Bayesian proportion estimates obtained from the proportionally allocated pixel sample.

With the Bayesian sequential allocation, the number of labeled pixels required can be determined in two ways: either by a predetermined fixed number or by a set maximum threshold value for the internal MSE estimate. Using the internal

MSE, this would potentially provide a uniform accuracy across segments by sampling more pixels from the more "difficult" segments and less pixels from the less "difficult" segments. To verify the expected performance of the MSE threshold, the sample-segment proportion estimates will be made with both a MSE threshold value to determine the number of dots, as well as a fixed number (50). The MSE threshold value will be provided by developmental testing prior to this exploratory. The fixed value of 50 was recommended by the SR Project.

From this experiment, it is intended that sequential allocation be applied after the dot labeling has been performed. This deviates from the manner in which it would normally be expected to operate. Based on knowledge of the allocation approach and use of the Integrated Labeling Procedure, the possibility exists that the analyst versus allocation procedure interactions may result in practice that will not occur in this test. That interaction would be defined as a tendency for the analyst, based on prior knowledge of allocation, to become biased in relabeling. This should not be a factor to consider with the Reformatted Labeling Procedure.

During the developmental testing, it was discovered that pixels selected for labeling by the Bayesian sequential allocation procedure were more accurately labeled than pixels randomly selected. Since there will be labels from two analyst labeling procedures and from ground truth, this contention can be assessed. If verified, this would represent an important finding.

2.4 METHODOLOGY

The number of analysts and segments used in this test will be the same as those specified in the U.S./Canada Wheat and Barley Exploratory Labeling Experiment Implementation Plan (ref. 2).

Acquisition selection will be determined by guidelines specified by the Reformatted Labeling Procedure. Depending on the availability, two to four acquisitions will be used to run CLASSY. For a single time period, an acquisition may be substituted for a missing biowindow. Otherwise, selection will be made from procedures in reference 17.

//

Using the selected acquisitions, the entire scene is clustered with an unsupervised adaptive maximum likelihood algorithm (CLASSY). Initially, 75 dots are proportionally allocated to clusters according to the cluster size. This set is then unioned with the set of 209-grid intersection dots to form the selection set. The analysts are provided with spectral aid products for the selected dots generated from a modified version of the CLASSY software. The backup support for these aids will be available on the Programmed Data Processor (PDP).

After labeling for small grains and if barley separation is feasible, spectral aids for barley separation are generated for final labeling. The labels are input into the Bayesian sequential allocation estimator with 50 dots and the MSE threshold number of dots for each of the three sets of labels to form the allocated dots. If the allocation is satisfied from the dot label set, then the remaining three proportion estimation techniques with 50 dots and the MSE threshold number will be run. However, if the number of dots required for a cluster from the Bayesian sequential allocation technique is not obtained, then an allocation of 150 dots, which includes the original 75, will be made and the process will continue until the need is satisfied. The analyst will repeat the labeling process on these extra dots, and the four estimation techniques will be applied with both the 50 and the variable number of dots determined by the MSE threshold. All the estimates, as well as intermediate products, will be on the Laboratory for Applications of Remote Sensing (LARS) files for use by the Accuracy Assessment (AA) Component in evaluation.

2.5 PRODUCTS, REPORTS, AND DOCUMENTATION

The technical aspects of this experiment will be monitored by a representative of the Experiment Design Component. The representative will assist in the training sessions before the test to ensure conformity with experiment design. Pertinent observations made by the experiment design representative during the course of the experiment will be included in the PTRR. Inputs on any changes to the initial design which occur during the experiment will be included in the final report. Recommendations for future exploratory and pilot experiments will also be included in the report.

3. ACCURACY ASSESSMENT

3.1 OBJECTIVES

The purpose of the AA Component in this supplemental experiment is the evaluation of proportion estimators for small grains. The evaluation will use segment proportion estimates resulting from four sampling and estimation techniques using labels from the Integrated and Reformatted Labeling Procedures, and ground-truth in an effort to detect differences in effects. If improvements (over P1) in the accuracy or the efficiency of crop proportion estimators are found, the technique responsible for the greatest improvement will be recommended for P1A.

3.2 APPROACH

Tests and evaluations will be made for two different sample sizes: a fixed sample size of 50 dots per segment and a variable segment sample size determined by requirements set on the segment MSE under sequential allocation. Deviations of proportion estimates under the two analyst-interpretor (AI) labeling procedures from proportion estimates based on ground-truth labels will be modeled and an analysis of variance run.

The statistical model is as follows:

$$\delta_{ijk\ell} = \mu + \tau_i + \lambda_j + (\tau\lambda)_j + b_k + \varepsilon_{ijk\ell}$$

where

$\delta_{ijk\ell}$ = $(\hat{P}_{AI} - \hat{P}_{GT})$ is the deviation of the proportion estimate under one of two labeling procedures from the proportion estimate based on the same dots with ground-truth labels.

μ = the overall mean deviation

τ_i = the effect of the i^{th} sampling and estimation technique $\sum_i \tau_i = 0$

λ_j = the effect of the j^{th} labeling procedure $\sum_j \lambda_j = 0$

$(\tau\lambda)_j$ = the effect of the interaction between the i^{th} technique and the j^{th} labeling procedure

b_k = the effect of the k^{th} block where blocking is done on the segments $b_k \sim N(0, \sigma_b^2)$

$\varepsilon_{ijk\ell}$ = the random error due to the ℓ^{th} segment in the i^{th} technique, j^{th} labeling procedure, and k^{th} block $\varepsilon_{ijk\ell} \sim N(0, \sigma^2)$

This model and analysis will be used to test (a) the effect of CLASSY clustering in producing small grain proportion estimates when sampling is proportional to cluster size as opposed to random, (b) the effect of Bayesian estimation in producing proportion estimates under proportional allocation as opposed to a relative frequency estimator, (c) the effect of sequential allocation with a Bayesian estimator in producing proportion estimates as opposed to proportional allocation with a Bayesian estimator, and (d) the effect of the Reformatted Labeling Procedure as opposed to that of the Integrated Labeling Procedure. If there is an interaction between techniques and labeling procedures, within-level evaluations will be made. For further evaluation purposes, nonparametric tests will be made on the error, $|\hat{P}_{AI} - \hat{P}_{GT}|$.

Squared deviations of proportion estimates from those under ground-truth labels will also be modeled with the same form as above, and analysis of variance procedures performed to test the effect of the sampling and estimation techniques on the MSE.

To test the effect of sequential allocation on labeling accuracy, a comparison will be made between the labeling accuracy of small grain proportion estimates for the samples under sequential allocation and the overall labeling accuracy for all dots labeled for each segment. The expected value of this difference under random sampling and proportional sampling is zero.

3.3 PRODUCTS, REPORTS, AND DOCUMENTS

A report will be prepared containing the results and conclusions of these evaluations. The report will also include any appropriate recommendations.

The AA Component will receive digitized ground-truth tapes, CLASSY cluster map file, and 418-dot overlays from the DAHP Component. The Classification Component will furnish the required proportion estimates and dot labels under the Reformatted and Integrated Labeling Procedures. The AA Component will prepare ground-truth labels.

4. SAMPLING AND AGGREGATION

The CY79 segments from the U.S./Canada Wheat and Barley Exploratory Labeling Experiment will be used in this test. No aggregation of data generated during this experiment is planned.

5. WEATHER ANALYSIS AND INTERPRETATION

No special requirements are placed on the Weather Analysis and Interpretation Component other than those specified in the U.S./Canada Wheat and Barley Exploratory Labeling Experiment (ref. 2).

6. CLASSIFICATION

6.1 OBJECTIVES

The objective of the Classification Component is to conduct segment processing for the exploratory tests by integrating the classification and labeling efforts as required by the experiment design. The results will form the basis for consideration of alternate spring grain classification procedures. The intent of this supplemental exploratory experiment will be satisfied with the completion of the following in FY80:

a. The documentation of a recommended procedure for use in the 1981 U.S./ Canada Wheat and Barley Pilot Experiment

b. The recommendation for further research and development of crop proportion estimation techniques based on the results from this supplemental exploratory experiment

6.2 APPROACH

The Classification Component approach will be to implement the candidate crop proportion estimation techniques using a selected set of data for spring small grains and barley, which are chosen from the labeling procedures in test 2 of the U.S./Canada Wheat and Barley Exploratory Labeling Experiment (ref. 2). Dots which must be labeled to satisfy this test will be determined prior to analyst labeling; thus, the crop proportion estimation experiment should be transparent to the labeling experiment.

Off-grid dots which must be labeled are selected by the use of proportional allocation of dots to clusters that were developed using the CLASSY algorithm. After labeling of the (209+) dots has been performed, crop proportions will be determined using the various techniques. If insufficient dots have been labeled, a new allocation will be determined and the analysts will label the additional dots. Proportion estimates and the CLASSY partitions will be transmitted to the Accuracy Assessment Component. Coordination with the DAHP Component will be provided by the Classification Component.

18

6.2.1 ACQUISITION SELECTION

A team of analysts will select two to four acquisitions to initialize processing using CLASSY on LARS and a maximum of six acquisitions for spectral aid production. For segments worked by both the Reformatted and Integrated Labeling Procedures, acquisition will be as specified in the Reformatted Labeling Procedure. For those segments that are worked only by the Integrated Labeling Procedure, acquisitions will be selected according to the procedures specified in reference 16.

6.2.2 DOT SELECTION

A dot selection and listing is needed for the initial labeling by the analysts and for the generation of the spectral aid package. This selection should be compatible with the Reformatted and Integrated Labeling Procedures and the various crop proportion estimation techniques. The minimum number of dots selected will be 209. This listing is acquired only when a segment has two or more usable acquisitions defined by the labeling procedure. These acquisitions are submitted to the DAHP Component to prepare for production of the dot listing. The dot listing will be provided back to the analysts by the DAHP Component.

6.2.3 CROP LABELS

After labeling is completed and corrections are made on the appropriate forms, all processing results that are the product of a completed analysis will be placed in an envelope labeled with the segment number, procedure used, test name, analyst name, and date. The imagery packet is forwarded to the DAHP Component, and the separate envelope is forwarded to the Classification Component for storage. Dot labels will be keypunched and forwarded to the AA Component.

A second copy of all forms will be submitted to the DAHP Component so that final dot labels can be extracted and inputted into the LARS computer system for application of the various proportion estimation techniques.

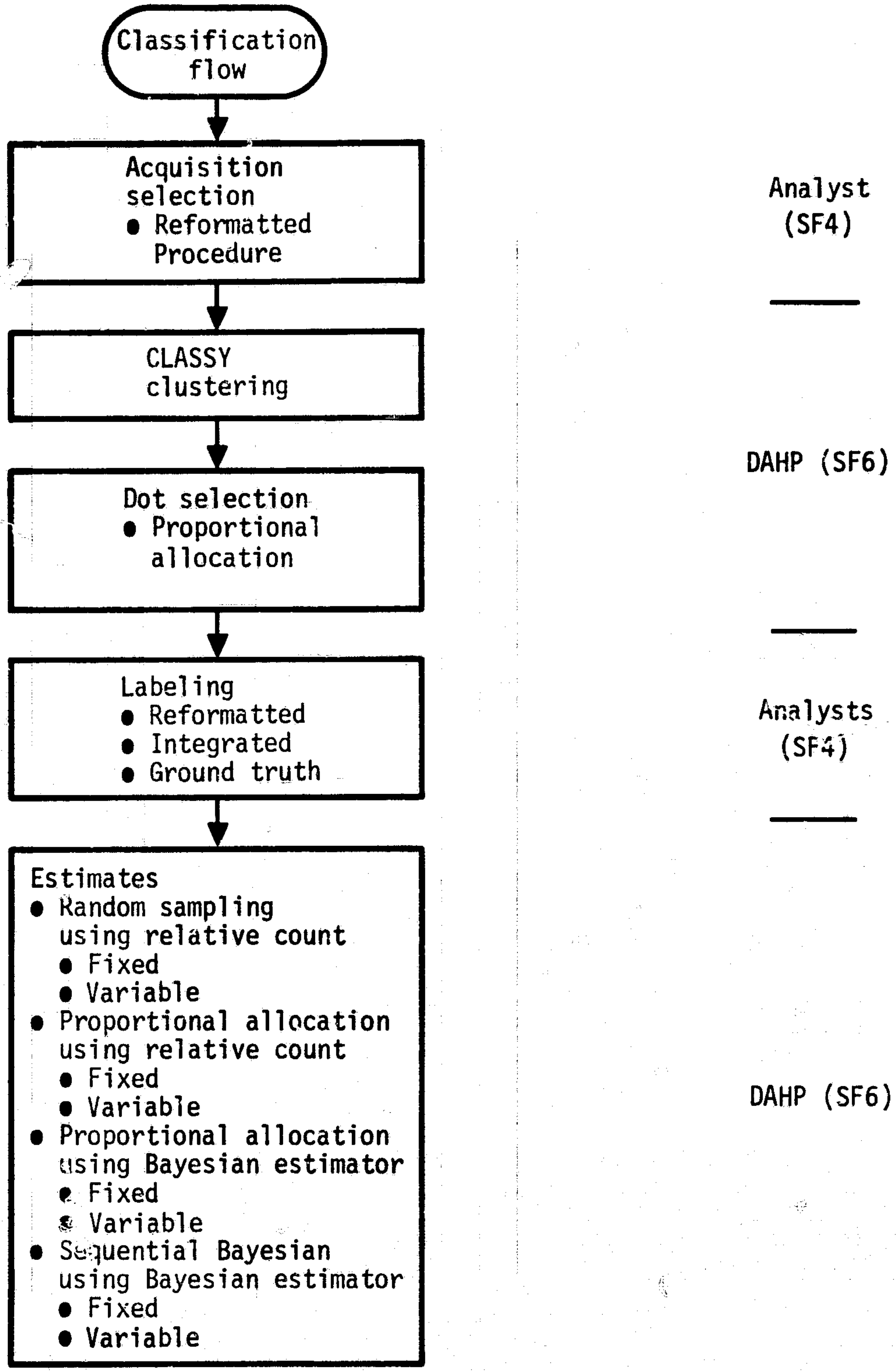

Figure 6-1.- The classification process flowchart.

6.2.4 CROP PROPORTION ESTIMATES

Once an analyst has selected acquisitions for clustering, a dot selection will be made proportionally according to the size of the clusters, and cluster information on each dot selected will be stored. These dots will then be independently labeled according to each labeling procedure and submitted to the DAHP Component for input into the various crop proportion estimation techniques in the LARS computer system. The crop proportion estimates will be developed and transmitted to the AA Component for evaluation. The classification process flow is illustrated in figure 6-1.

6.2.5 STORAGE OF PROCESSING RESULTS

The separate envelope containing the interpretation products, labels, and proportion estimates will be maintained by the Classification Component for evaluation purposes. The analyst packet will be provided to the DAHP Component and will either be returned to storage or reassigned for the second or later analysis as directed by the Classification Component. The envelope containing the analysis results will be provided to the Classification Component for storage where it will be available for evaluation purposes.

6.2.6 STATUSING AND TRACKING

Tracking will be done by the Classification Component. For each segment, the state, strata, procedure, and AI are recorded.

6.2.7 PROCESSING COORDINATION

The Classification Component will interface with the DAHP Component for crop proportion estimation generation and for the receipt and dissemination of products. Activities associated with data base status and other interaction with the data bases will be through the DAHP Component. Responsibility for ensuring the coordination of processing activities, for computer scheduling, for data flow, for accuracy assessment of data requirements, and for other evaluation of data requirements will be with the Classification Component.

6.2.8 TEST CONTROL

The Classification Component will provide consultation on procedural questions, provide feedback to all functional components, and provide an interface between labeling teams and the DAHP Component. The Classification Component will also coordinate meetings with the Accuracy Assessment Component and be responsible for resolving procedural deficiencies. The Classification Component will monitor packet flow between analysts. Packets will be checked for the required listing of dots that are needed for initial interpretation.

Finally, the results from each processing (i.e., the spectral aids, processing forms, and crop proportion estimates) will be stored by the Classification Component in a manner to be easily tracked and made available for accuracy assessment evaluation or any subsequent evaluation.

6.3 PRODUCTS, REPORTS, AND DOCUMENTATION

The Classification Component will provide the required acquisitions and labels to the DAHP Component and will receive dot lists and proportion estimates from the DAHP Component. The AA Component will be provided with the proportion estimates. Results on the implementation of these procedures will be documented for subsequent use in exploratory and pilot experiments. Preparation of the PTRR will be supported. Recommendations for further crop proportion estimation procedure improvements and recommendations for future exploratories and pilots will be submitted for inclusion in the final U.S./Canada Wheat and Barley Exploratory Experiment report.

7. CROP CALENDARS

No special requirements are placed on the Crop Calendar Component other than those specified in the U.S./Canada wheat and Barley Exploratory Labeling Experiment (ref. 2).

8. DATA ACQUISITION, HANDLING, AND PROCESSING

8.1 OBJECTIVES

The objective of the DAHP Component is to provide an approach to the orderly acquisition, storage, retrieval, and dissemination of the products necessary to support this supplemental exploratory experiment. Although the data sets that support this plan are historical rather than realtime, the management and the status and tracking approach is essentially the same as it would be for realtime data processing.

8.2 DATA

All data requirements are the same as specified for the U.S./Canada Wheat and Barley Exploratory Labeling Experiment (ref. 2).

8.3 PACKET PREPARATION AND COORDINATION

Packets will be provided to the analysts by the DAHP Component based on analyst requirements. Coordination of processing activities will also be accomplished by the DAHP Component. Processing forms will be submitted by the Classification Component to the DAHP Component. Analyst inputs will be translated by the DAHP Component into computer run decks and/or typed on terminals into files to be executed on the required development software systems.

8.4 STATUSING AND TRACKING

Statusing, tracking, and reporting by the DAHP Component will not exist during this experimental test. Any proportion estimates produced by the experiment will not be stored in the DAHP Component data base.

8.5 ELECTRONIC DATA MANAGEMENT

The required data bases on the LARS computer will be provided based upon the requirements of the experiment design. Specifically, the CY79 Landsat data will be made available on a tape data base at LARS.

8.5.1 LARS CLASSY PROCESSOR 1

Cluster files will be developed in preparation for the crop proportion estimation experiment. Required spectral plots containing the regular 209-grid dots and the necessary off-grid dots will then be prepared using the LARS CLASSY Processor 1. The Classification Component will submit the necessary data processing forms to the DAHP Component. Listings of the dots to be labeled and spectral plots will be transmitted to the Classification Component.

8.5.2 LARS CLASSY PROCESSOR 2

Proportion estimates will be computed using the LARS CLASSY Processor 2. Inputs will be provided by the Classification Component. The proportion estimate outputs will be transmitted to the Classification Component.

9. YIELD

Actual or estimated yield values are not required for this experiment.

10. DATA AND SYSTEMS REQUIREMENT SUMMARY

The data and systems requirements necessary to support the components of the experiment are summarized in this section and in the U.S./Canada Wheat and Barley Exploratory Labeling Experiment Implementation Plan (ref. 2).

10.1 EXPERIMENT DESIGN

In order to make recommendations for future exploratory and pilot experiments, the Experiment Design Component requires results from the final accuracy assessment analysis.

Data requirements for this test are the same as specified in reference 2 plus the following:

a. All clusters derived using the clustering algorithm CLASSY

b. A list of off-grid dots to be labeled

c. Labels from the U.S./Canada Wheat and Barley Exploratory Labeling Experiment

d. Ground-truth labels corresponding to the labeled dots

e. Wall-to-wall ground truth for a subset of the segments

f. Proportion estimates from the required combinations of the allocation and proportion estimation techniques

The LARS computer system will be used to produce cluster files, develop lists of off-grid dots, and calculate the final proportion estimates. The PDP 11/45 and the LARS computer system will be used by the AA Component in the evaluation.

10.2 ACCURACY ASSESSMENT

In order for the AA Component to carry out its evaluation activities, all data specified in reference 2 and in section 10.1 are required. The PDP 11/45 and the LARS computer systems will be used in evaluating results.

10.3 CLASSIFICATION

Data requirements for the Classification Component are specified in reference 2. Because the Earth Resources Interactive Processing System (ERIPS) is no longer available, the LARS computer system or the PDP 11/45 will be used by the DAHP Component to produce cluster partitions, lists of off-grid dots, and lists of proportion estimates.

10.4 DAHP

The LARS computer system or PDP 11/45 and an image data base (FY79 wheat/barley blind sites) extracted from ERIPS are required for the plan. Developmental computer software for producing clusters, spectral plots, and the required proportion estimates must be available to support the plan.

11. EXPERIMENT SCHEDULE AND RESOURCE SUMMARY

Pursuant to the FCPF Implementation Plan (ref. 1), an exploratory experiment for wheat and barley in the United States and Canada has been scheduled for FY80. Supplemental to this experiment, a small grain/wheat and barley exploratory proportion estimation experiment has been planned. The testing and evaluation of candidate procedures for follow-on exploratory and pilot experiments will be reported in a PTRR followed by formal documentation early in FY81. The PTRR presentation is currently scheduled, consistent with data and critical resource availability, to allow the earliest possible incorporation of findings into FY81 experiment planning.

11.1 EXPERIMENT SCHEDULES

In order for an experiment to arrive at a successful conclusion, an integrated schedule must be developed portraying each component's relationship to the other. The experiment schedule shows this relationship in summary form (fig. 11-1).

11.2 RESOURCE SUMMARY

Resources necessary to conduct this exploratory experiment will be drawn from those provided for the project as defined by task sheet in the FCPF Project Implementation Plan. Resource scoping across organizational elements is generally consistent with individual task sheet estimates. Both civil service and contractor personnel are involved to varying degrees in all facets of the experiment. The following sections present the projected requirements for civil service and contractor personnel.

11.2.1 CIVIL SERVICE RESOURCE SUMMARY

Civil service personnel are engaged in numerous tasks related to the Project Management/Support and the DAHP areas which are involved in directing and monitoring the Experiment Design and Accuracy Assessment contractor effort and in performing analysis work during certain classification tests. Total civil service manpower involved in this experiment is 0.5 man-year equivalents (MYE's).

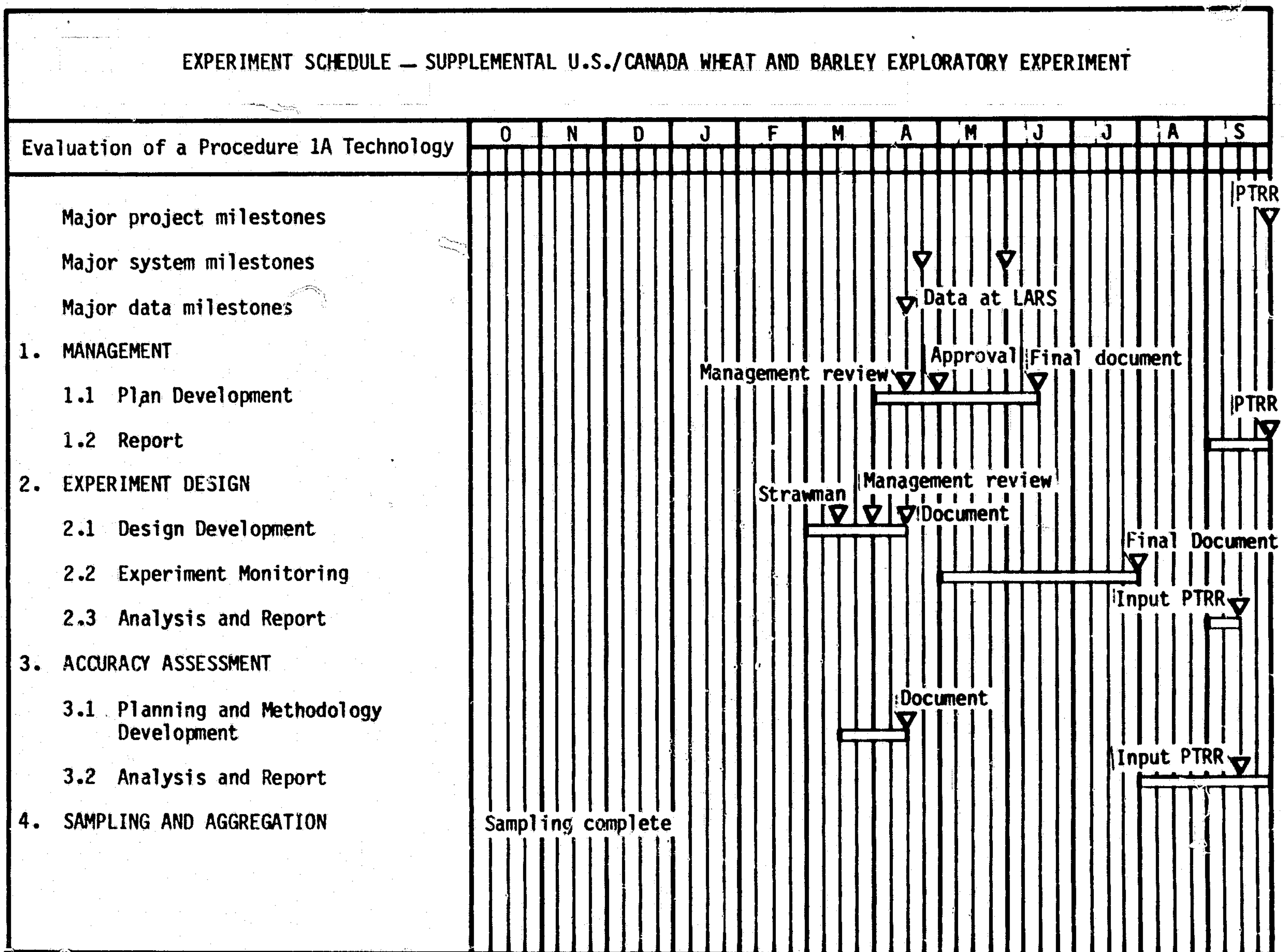

Figure 11-1.- An integrated schedule of the experiment's components.

30

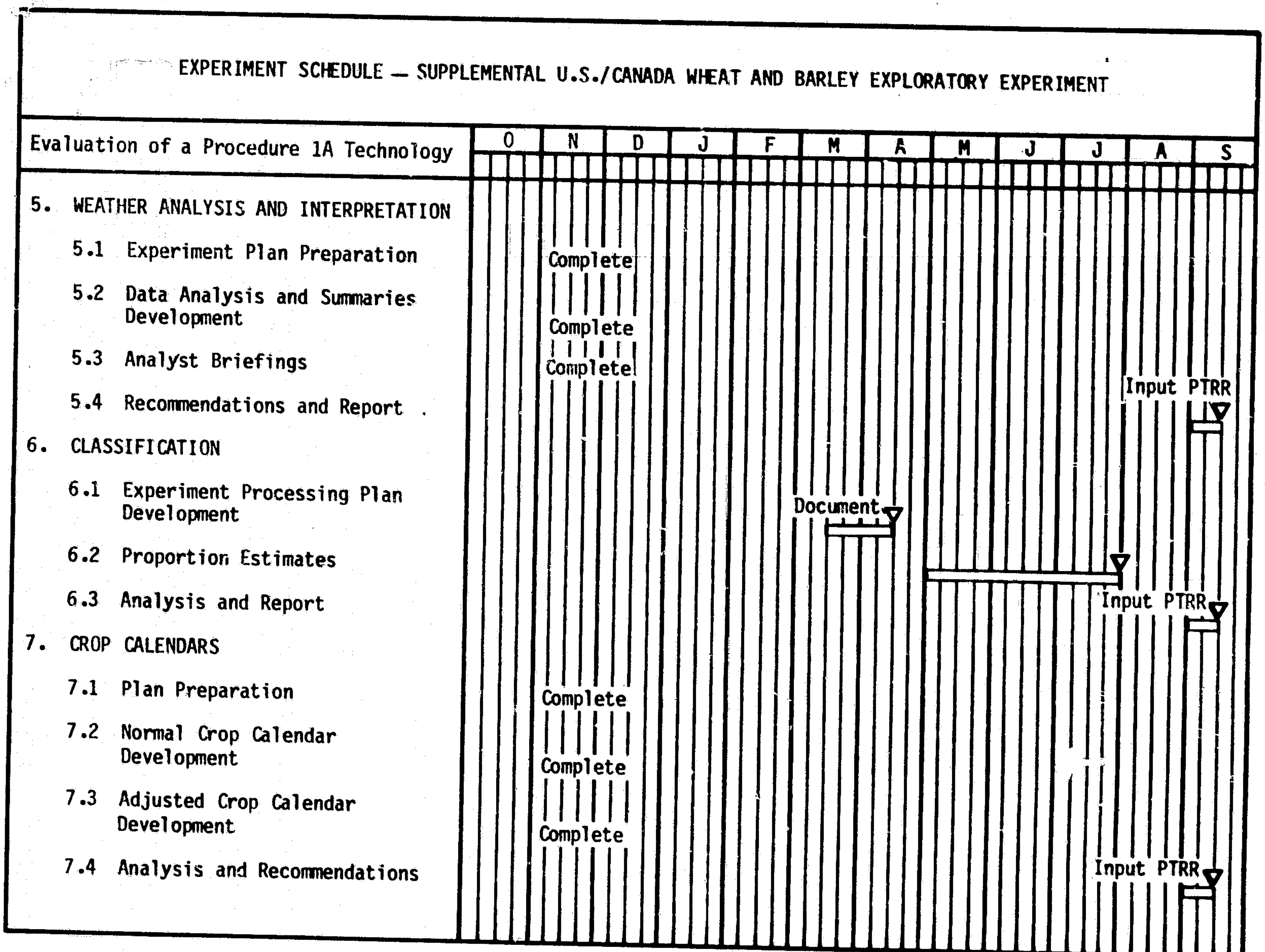

Figure 11-1.- Continued.

31

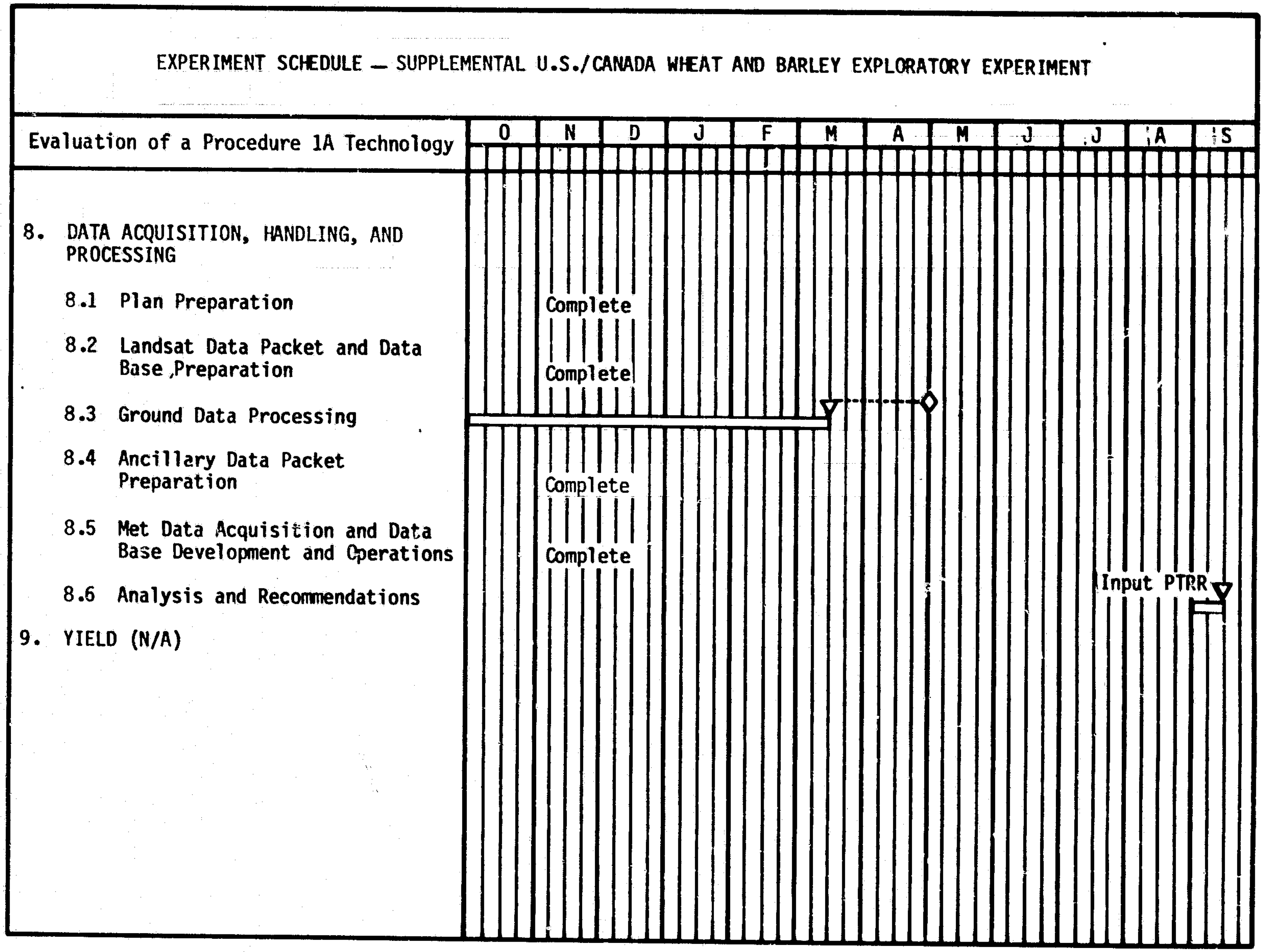

Figure 11-1.- Concluded.

11.2.2 CONTRACTOR RESOURCE SUMMARY

Contractor personnel participate in the implementation of this experiment through all components. The level of man-power involvement on a per component basis is shown in table 11-1.

TABLE 11-1.- CONTRACTOR MAN-YEAR EQUIVALENTS (MYE) PER COMPONENT

Section	Component	Support contractor MYE
1.	Experiment Plan Development	0.2
2.	Experiment Design	0.15
3.	Accuracy Assessment	0.5
[a]4.	Sampling and Aggregation	--
[a]5.	Weather Analysis and Interpretation	--
6.	Classification	0.5
[a]7.	Crop Calendars	--
8.	Data Acquisition, Handling, Processing Status, and Data Bases	--
9.	Yield	--

[a]Resources defined in U.S./Canada Wheat and Barley Exploratory Labeling Experiment Implementation Plan (ref. 2).

12. REFERENCES

1. AgRISTARS, Foreign Commodity Production Forecasting Implementation Plan. AgRISTARS/FCPF Report, FC-J0-C0604, JSC-16344, Jan. 1980.

2. FCPF, U.S./Canada Wheat and Barley Exploratory Labeling Experiment Implementation Plan. AgRISTARS/FCPF Report, FC-J0-00600, JSC-16336, Jan. 1980.

3. Lennington, R. K.; and Johnson, J. K.: Clustering Algorithm Evaluation and the Development of a Replacement for Procedure 1. LEC-13945, JSC-16232, November 1979.

4. Pore, M. D.: Bayesian Techniques in Stratified Proportion Estimation. Proceedings of the 1979 Annual Meeting of the American Statistical Association, LEC-13490, August 1979.

5. Heydorn, R. P., et al.: Classification and Mensuration of LACIE Segments. Proceedings of Technical Sessions - the LACIE Symposium, vol. 1, JSC-16015, October 1978.

6. Carnes, J. G.: Detailed Analysis of CAMS Procedures for Phase III Using Ground-Truth Inventories, LEC-13343, JSC-13343, April 1979.

7. Havens, K. A.: Secondary Error Analysis: The Evaluation of Analyst Dot Labeling. LEC-12380, JSC-14493, September 1978.

8. Havens, K. A.: Further Evaluation of Procedure 1, Secondary Error Analysis. LEC-13180, JSC-14746, May 1979.

9. Baird, J. E.: LACIE Transition Project Second Interim Accuracy Assessment Report. LACIE-00476, JSC-13753, LEC-13075, February 1979.

10. Lennington, R. K.; and Malek, H.: The CLASSY Clustering Algorithm, Description, Evaluation and Comparison with Iterative Self-Organizing Clustering System (ISOCLS), LEC-11289, March 1978.

11. Lennington, R. K.: and Rassbach, M. E.: CLASSY - An Adoptive Maximum Likelihood Clustering Algorithm. LEC-12145, May 1978 [Presented at the Ninth Annual Meeting of the Classification Society, North American Branch, Clemson University (Clemson, South Carolina), May 21-23, 1978].

12. Lennington, R. K.; and Rassbach, M. E.: Mathematical Description and Program Documentation for CLASSY, an Adoptive Maximum Likelihood Clustering Method. LEC-12177, JSC-12177, April 1979.

13. Bryant, J.: On the Clustering of Multidemensional Pictoral Data. Pattern Recognition, vol. 11, 1979, pp. 115-125.

14. Kan, E. P.: The JSC Clustering Program ISOCLS and Its Applications, LEC-0483, July 1973.

15. Palmer, W. F.: The Reformatted Labeling Procedure for Spring Small Grains and Barley. AgRISTARS/FCPF Report (to be published; preliminary documentation in transmittal letter LEC 644-1472, December 20, 1979).

16. Payne, R. W.: The Integrated Analysis Procedure for Identification of Spring Small Grains and Barley. AgRISTARS/FCPF Report FC-LO-00451, LEC-14385, May 1980.

17. Daily, C. L.: Detailed Analysis Procedures for Transition Project (FY79). LACIE-00724, JSC-13756, LEC-13222, May 1979.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1980—671-099/17 33

36