https://ntrs.nasa.gov/search.jsp?R=19790010218 2020-03-21T23:21:22+00:00Z

LACIE-00464
JSC-13728

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)

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

79 10122
TM-79986

NASA NOAA USDA

(E79-10122) LARGE AREA CROP INVENTORY EXPERIMENT (LACIE). FIRST INTERIM PHASE 3 EVALUATION REPORT Interim Report (NASA) 129 p HC A07/MF A01 CSCL 02C
N79-18389
Unclas
G3/43 00122

FIRST INTERIM PHASE III

EVALUATION REPORT

NASA
National Aeronautics and
Space Administration

Lyndon B. Johnson Space Center
Houston Texas 77058

JANUARY 1978

LARGE AREA CROP INVENTORY EXPERIMENT
(LACIE)

FIRST INTERIM PHASE III EVALUATION REPORT

Approved By:

R. B. MacDonald, Manager
Large Area Crop Inventory Experiment

Original photography may be purchased from
EROS Data Center

Sioux Falls, SD 57198

January 1978

EXECUTIVE SUMMARY

The overall accuracy of LACIE wheat production estimates for Phases I, II, and III strongly supports the contention that the technology is capable of providing improved early-season and at-harvest production estimates in major wheat-producing regions of the world outside the United States. Results through mid-Phase III of LACIE are particularly encouraging in the winter-wheat regions of the world. The LACIE mid- to late-season estimates of winter wheat were adequate to support the LACIE 90/90 at-harvest goal for production. In Phase II, there was a tendency to underestimate spring wheat production in the United States and Canada, primarily because of spring-wheat acreage underestimates. However, improvements implemented for Phase III are projected to decrease the size of the acreage underestimate.

After 2-1/2 years of LACIE operations, Phases I and II have been concluded on schedule; Phase III activities have begun; and a Transition Year to complete, document, and transfer the LACIE technology to an evolving U.S. Department of Agriculture Application Test System has been approved.

During Phase I, the LACIE system components and technology were developed and successfully exercised. Analysis was primarily limited to the U.S. Great Plains "yardstick" region. Acreage estimation was performed in a quasi-operational mode, whereas yield and production estimates were performed in a feasibility test mode. Wheat acreage classification tests were conducted also on exploratory regions outside the United States. Several improved technology approaches were developed for subsequent implementation in Phases II or III.

In Phase II, quasi-operational wheat acreage, yield, and production estimation was conducted for the U.S. Great Plains "yardstick" region, for Canada, and for indicator regions of the U.S.S.R.

PRECEDING PAGE BLANK NOT FILMED

ORIGINAL PAGE IS
OF POOR QUALITY

The scope of LACIE Phase III has been expanded significantly over Phase II. Operations in the yardstick region are being continued in Phase III, with additional emphasis on evaluations of various technology updates. The U.S.S.R. operations have been expanded from coverage of Phase II indicator regions to include the entire Soviet wheat crop, in order to obtain more reliable independent U.S.S.R. statistics for evaluating LACIE estimates. In cooperation with the Canadian Government, classification technology assessment work has been intensified in Canada with the addition of some 30 blind sites. Two crop years of LACIE operations have resulted in the definition of several key areas for technology improvement. These improvements have been developed and tested in LACIE Research, Test, and Evaluation activity and are being quasi-operationally tested in Phase III.

Since the currently implemented remote sensing technology and approach are in the developmental stage, a significant improvement in crop surveys is expected in the future. As LACIE activity proceeds, the technology is expected to improve greatly. These improvements will be accompanied by a better understanding of factors which affect the accuracy of remote sensing crop surveys.

CONTENTS

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

Section Page

ORIGINAL PAGE IS
OF POOR QUALITY

TABLES

ORIGINAL PAGE IS
OF POOR QUALITY

FIGURES

ABBREVIATIONS

ORIGINAL PAGE IS
OF POOR QUALITY

ACC	adjustable crop calendar
agromet	agricultural-meteorological
ASCS	Agricultural Stabilization and Conservation Service, U.S. Department of Agriculture
ATS	Application Test System
BMTS	Biometeorological Time Scale used in calculating the stages of crop development
CAMS	Classification and Mensuration Subsystem
CAS	Crop Assessment Subsystem
CCEA	Center for Climatological and Environmental Assessment
CMR	CAS Monthly Report
CRD	Crop Reporting District
CRT	cathode-ray tube
CV	coefficient of variation
DEC	Digital Equipment Corporation
ERIPS	Earth Resources Interactive Processing System
ETAC	Environmental Technical Applications Center
FAS	Foreign Agricultural Service, U.S. Department of Agriculture
GIN	Green Index Number
GSFC	Goddard Space Flight Center
I^2S	International Imagery Systems
IMAGE 100	Interactive Multispectral Image Analysis System, model 100
ISOCLS	Iterative Self-Organizing Clustering System
ITS	intensive test site

JSC	Lyndon B. Johnson Space Center
KSU	Kansas State University
LACIE	Large Area Crop Inventory Experiment
MAP	Management and Productivity
MSS	multispectral scanner
NASA	National Aeronautics and Space Administration
90/90 criterion	premise that LACIE at-harvest estimates are to be within 10 percent of the true production at the national level 90 percent of the time
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service, U.S. Department of Commerce
PAYES	Production, Acreage, and Yield Estimation System
pixels	picture elements
Procedure 1	a highly automated cluster-based procedure for LACIE analysis which removes all required analyst functions except interpretation of Landsat and ancillary data products for the purpose of labeling spectral pixels as wheat or nonwheat
RD	relative difference
RFP	request for proposal
RT&E	Research, Test, and Evaluation
SRS	Statistical Reporting Service, U.S. Department of Agriculture
Transition Year	period during which LACIE technology will be completed, documented, and transferred to a U.S. Department of Agriculture test system
USDA	U.S. Department of Agriculture
WMO	World Meteorological Organization

yardstick region	U.S. Great Plains area used in evaluating LACIE technology; includes nine states: Colorado, Kansas, Minnesota, Montana, Nebraska, North Dakota, Oklahoma, South Dakota, and Texas.
YES	Yield Estimation Subsystem

ORIGINAL PAGE IS
OF POOR QUALITY

1. INTRODUCTION

ORIGINAL PAGE IS
OF POOR QUALITY

1.1 PURPOSE AND SCOPE

The purpose of this report is to document the results of the Large Area Crop Inventory Experiment (LACIE) as of August 1977. All accuracy and performance discussions are based on data acquired through near-harvest for winter wheat, inasmuch as analyses were not available for spring wheat at the time this report was compiled. The scope of Phase III has been expanded to include more complex operations than Phase II. The Phase III Landsat operational data volume is almost 200 percent greater than that of Phase II* as a result not only of expanded U.S.S.R. coverage from the Phase II indicator regions to the entire Soviet wheat crop but also of an increase in Landsat sampling density of some 50 percent in the U.S. Great Plains (yardstick)† region. The Phase III scope is expanded further by parallel evaluations of second-generation acreage sampling and yield estimation technology over moderately large regions in the yardstick region and in the U.S.S.R. Additionally, the second-generation Landsat data machine processing technology developed and tested in Phase II is being implemented in a staged system delivery mode over all Phase III regions. Experience through Phase II showed that the estimation of spring-wheat (in comparison to winter-wheat) acreage was somewhat more difficult. Therefore, in Phase III, increased emphasis has been placed on evaluating the second-generation machine processing technology for spring-wheat acreage estimation. In cooperation with the Canadian Government, some 30 Canadian blind sites have been added for this purpose.

---

*The LACIE processed 9277 segments in Phase II and 17 445 in Phase III.

†*U.S. Great Plains* and *yardstick* will be used interchangeably in this report. This region encompasses nine states: Minnesota, Montana, and North and South Dakota (the U.S. northern Great Plains); and Colorado, Kansas, Nebraska, Oklahoma, and Texas (the U.S. southern Great Plains).

This section presents the LACIE background, the project structure, its division into phases, scheduling, and organization. The technical approach is summarized, highlighting the key technical improvements implemented in Phase III. Section 2 discusses the results of LACIE to date, including the accuracy of the winter-wheat acreage, yield, and production estimates, as well as the performance of the improved quasi-operational data analysis system. Results of the LACIE Research, Test, and Evaluation (RT&E) activity are summarized and the key technical issues remaining as of mid-Phase III are reviewed.

Finally, in section 3, the outlook for the remainder of Phase III and beyond will be discussed, along with currently envisioned technology modifications required at the end of Phase III. The status of the U.S. Department of Agriculture (USDA) advanced system for transferring the LACIE technology for applications testing and the LACIE follow-on food and fiber program will be discussed. The food and fiber program will focus on adapting the LACIE technology for application to multiple crop inventories.

1.2 LACIE OVERVIEW

1.2.1 OBJECTIVES

The LACIE was initiated in 1974 as a "proof of concept" program. It was designed to assimilate remote sensing technology developed over the previous decade and to apply the resultant experimental system to the task of monitoring a singularly important agricultural commodity (wheat). The experimental approach was to be modified as necessary to demonstrate the technical and cost feasibility of global agricultural monitoring systems.

Timeliness and accuracy goals for LACIE were established in recognition of the essential requirements for global agricultural information. The experiment was designed to establish the

ORIGINAL PAGE IS
OF POOR QUALITY

feasibility of acquiring and analyzing Landsat data within a 14-day interval. Importantly, the at-harvest estimates were to be within 10 percent of the true production at the national level 90 percent of the time (the LACIE 90/90 criterion). An additional performance goal was that of determining how early in the crop year estimates could be produced and with what accuracy and repeatability. Additionally, the estimates were to be made using repeatable and objective procedures with qualitative judgments kept to a minimum.

1.2.2 ELEMENTS AND PARTICIPANTS

Three major elements comprised the LACIE: (1) a quasi-operational element to acquire and analyze Landsat and meteorological data to make experimental estimates of production, (2) an offline element to test and evaluate alternative approaches as required to meet the performance goals of the experiment, and (3) an element to research and develop alternative approaches.

The experiment has been jointly conducted by personnel from the National Aeronautics and Space Administration (NASA), the USDA, and the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce. These government entities represent the many disciplines (including physics, plant pathology, engineering, agronomy, statistics and mathematics, soils sciences, economics, and plant physiology) necessary to meet the objectives of the experiment.

The major components of the quasi-operational element of the experiment include Landsat and its acquisition and preprocessing subsystem; the World Meteorological Organization (WMO) weather reporting system; the NOAA development and operational facilities in the Washington, D.C., and Columbia, Missouri, regions; and the analysis, compilation, and evaluation activities at the NASA Johnson Space Center (JSC) in Houston, Texas. The experiment also draws significantly on the expertise of university and industrial research personnel.

Because of the complexity and importance of LACIE, periodic technical reviews have been held where invited experts have reviewed LACIE results, discussed specific technical issues, and made specific recommendations. This process has made significant contributions to the LACIE.

1.2.3 PHASES AND SCHEDULES

The experiment was scheduled to be conducted in three phases on the timeline shown in figure 1-1, with the following objectives:

a. In Phase I, the technology to estimate the crop proportions of wheat-growing regions would be implemented and tested; and, similarly, the technique to estimate the yield from specific acreages would be developed and tested.

b. In Phase II, the technology modified during Phase I would be tested further over expanded geographic regions and modified as required.

c. In Phase III, the modified technology would be tested and evaluated over an even wider range of geographic conditions.

In addition, a Transition Year extending LACIE through 1978 has been approved. In the Transition Year, the LACIE technology developed in the experiment will be completed, documented, and transferred to a USDA test system.

1.3 LACIE TECHNICAL APPROACH

The LACIE approach utilizes the direct observational capabilities afforded by the Landsat, together with estimates of weather variables to estimate production. This approach requires that each geographic subregion (selected to be relatively homogeneous with regard to wheat acreage and yield) in a country be monitored (1) to forecast the quantity of wheat acres (hectares in LACIE

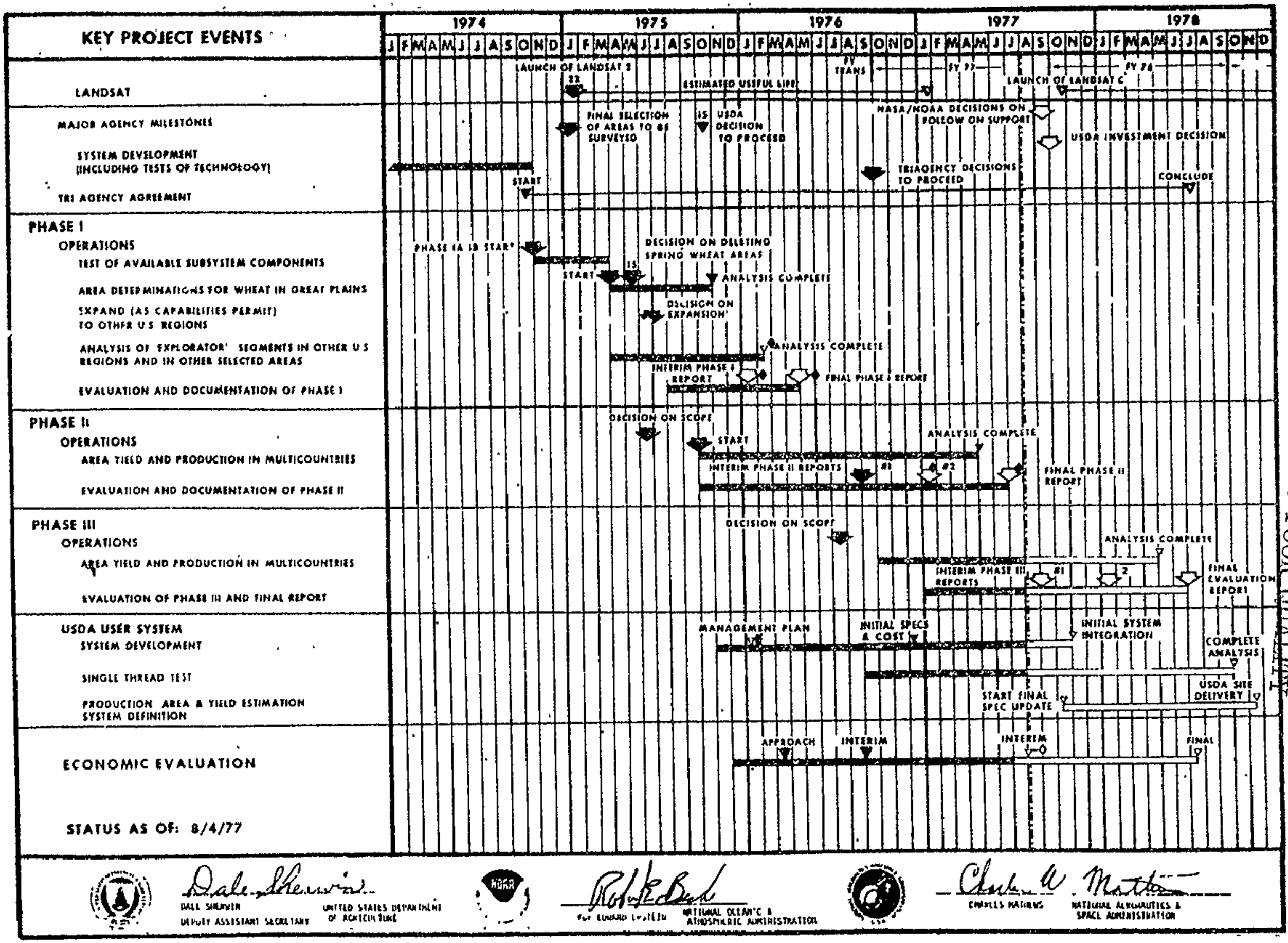


Figure 1-1.- LACIE level 1 schedule as of August 4, 1977.

ORIGINAL PAGE IS OF POOR QUALITY

foreign countries) available for harvest (both winter and spring, individually, in each subregion) and (2) to forecast the expected productivity (yield) for each subregion (based on the acres available for harvest). The total wheat production for each subregion is then forecast by multiplying the available acres for harvest times the average yield per harvested acre. The production estimates for all subregions are then summed to obtain a forecast at the country level. In addition, the subregional forecasts of acres for harvest are summed to obtain a forecast of national acres for harvest. An average yield for all acres harvested nationally is then obtained. It is, by definition, the acreage-weighted average. This acreage-weighted average yield is a desirable estimate to have because, when multiplied by the national acreage, it will produce the national production estimate. The LACIE stratification and sampling approach is similar to the domestic approach utilized by the timely and accurate USDA Statistical Reporting Service (SRS) survey system. The approach has been adopted also by the Canadian and other national governments.

Within each of the described subregions, Landsat multispectral scanner (MSS) data are collected every 18 days from 9- by 11-kilometer (5- by 6-nautical-mile) segments drawn at random from each stratum. Wheat is distinguished from nonwheat within each segment by monitoring the temporal development of the crops from planting through harvest. The areal percentage of wheat in each segment in the stratum is then estimated; and, using this information, an average percentage for the stratum is determined. The average areal percentage of wheat can then be multiplied by the total agricultural acreage in the stratum* to estimate total wheat acres for the stratum.

---

*Stratum agriculture is delineated on full-frame Landsat imagery and planimetered to determine total agricultural acreage within a stratum. Agriculture is defined to be any area of the image for which field patterns are evident.

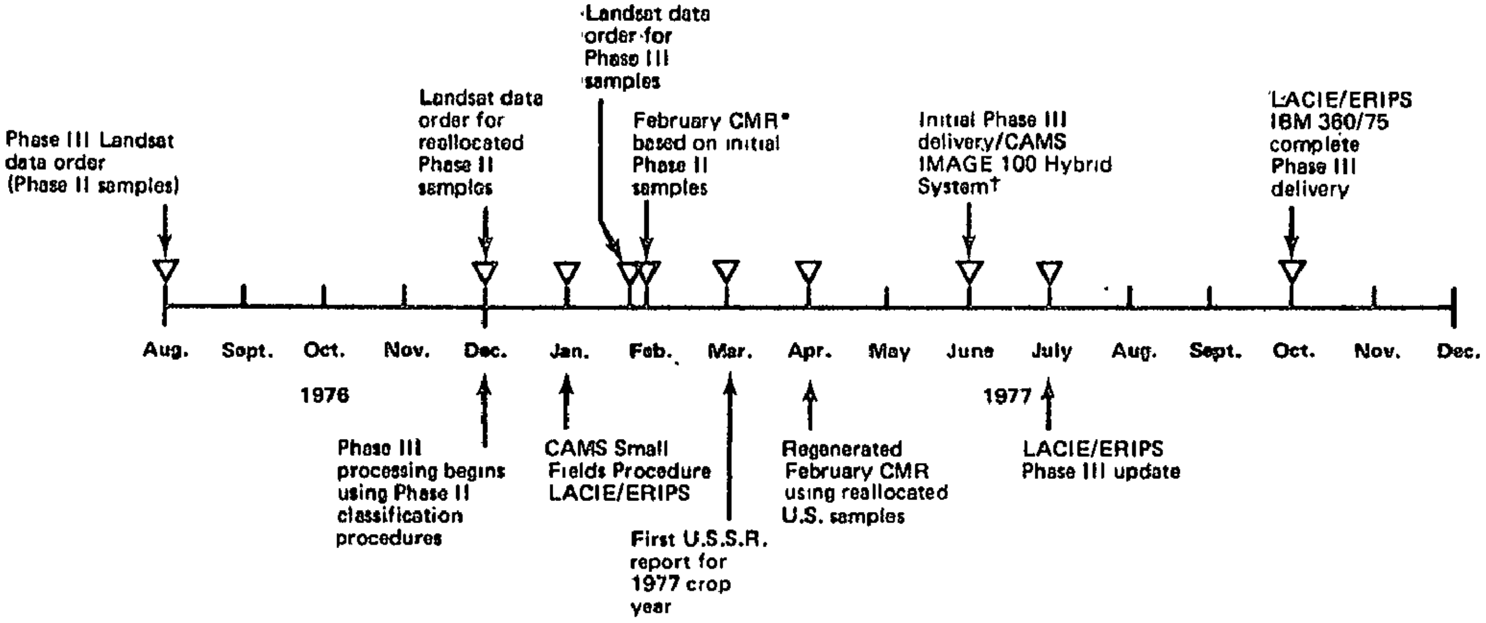
The yield for harvested acres is forecast in LACIE through the use of regression models which utilize weather-related variables obtained from the ground-based stations of the WMO network. These models are referred to as agricultural/meteorological (agromet) models. The first-generation models currently used in LACIE are developed around monthly averages of temperature and precipitation. In the yardstick region, both winter and spring wheat models cover 15 subregions. The yield and climatic data base used to derive the yardstick models is approximately 45 years in length. The yield data are obtained by aggregating the SRS estimates of harvested acreage and production to obtain yield in bushels per harvested acres for both winter and spring wheat individually in each of the 15 subregions. The climatic data consist of monthly climatic division averages of precipitation and temperature. These averages are weighted using acres harvested to obtain the monthly average temperature and total precipitation for a given region. A piecewise linear trend is used to model the technology trend.

A more detailed illustration of the LACIE technical approach was presented to the Eleventh International Symposium on Remote Sensing of Environment in April 1977 (ref. 1).

1.4 LACIE TECHNICAL REVIEWS — TECHNOLOGY MODIFICATIONS

Recognizing the value of periodic technical reviews, LACIE personnel schedule formal, in-depth technical reviews by selected technical personnel inside and outside LACIE having expertise relevant to the LACIE technology. Reviews have been held at approximate 6-month intervals, and recommendations are tracked to logical final disposition. The most significant recent changes resulting from these reviews have been a second-generation sampling strategy, improved Landsat classification procedures, and improved yield models. The schedule for these various system modifications and/or deliveries is given in figure 1-2.

ORIGINAL PAGE IS
OF POOR QUALITY



*Crop Assessment Subsystem (CAS) Monthly Report (CMR).

†Classification and Mensuration Subsystem (CAMS) Interactive Multispectral Image Analysis System, model 100 (IMAGE 100), Hybrid System. This system consists of batch processing on the LACIE/Earth Resources Interactive Processing System (LACIE/ERIPS) utilizing the IBM 360/75 and interactive processing on the General Electric IMAGE 100.

Figure 1-2.— Schedule of LACIE Phase III system modifications and/or deliveries.

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

1.5 PHASE III FIRST-GENERATION SAMPLING STRATEGY CHANGES

The first-generation sampling strategy utilized in Phase II was designed to achieve a 2-percent sampling error at the U.S. country level. The sampling strategy was modified in Phase III to achieve a 5-percent coefficient of variation (CV) in the LACIE yardstick area production estimate. This CV permits the 90/90 criterion to to be met even with a reasonable degree of bias in the production estimate. This modification necessitated an increase in samples from 431 to 601 in the yardstick area. These samples were reallocated using improved estimates of the distribution of wheat based on small grains identified from Landsat imagery. Their reallocation was based also on interpretations of full-frame Landsat data for agricultural areas and empirical estimates of classification and yield estimation error.

The modified allocation and location of segments was not completed prior to the Phase III data order submission in August 1976 for the 1977 crop year, at which time the initial Phase III Landsat acquisitions were ordered for the Phase II sample segment locations. The initial LACIE Phase III crop report in February 1977 (ref. 2) was based on these sample segments acquired through December 1976. The Phase III sample locations were completed and data were ordered retrospectively on January 31, 1977. The new segments acquired through December 1976 were processed, and a 601-segment allocation report, replacing the earlier 431-segment allocation, was regenerated on April 6, 1977.

The first-generation sampling strategy also was improved in the U.S.S.R. for Phase III to accommodate the updated agricultural/nonagricultural delineation using Landsat full-frame imagery. Originally, the agricultural area had been defined without the benefit of Landsat data. As a result of using the imagery, approximately 700 segments erroneously located in nonagricultural areas were relocated to agricultural areas. In addition, based on improved estimates of U.S.S.R. stratum agricultural area variance, 160 segments were reallocated for more efficient sampling of the agricultural area. The U.S.S.R. sample segment locations for Phase II were ordered for Phase III in August 1976, and the new Landsat sample segment data were ordered retrospectively on December 1, 1976. As a result, the publishing of the first U.S.S.R. report for crop year 1977 was delayed from January 1977 until March 1977.

1.6 SECOND-GENERATION SAMPLING STRATEGY EVALUATION

Phase III includes an evaluation of a second-generation sampling strategy. The first-generation sampling strategy is a stratified random method where the strata and sample allocations are primarily based on historical data augmented by Landsat imagery to delineate agricultural land and to estimate within-stratum sample variance. Because historical acreage and yield data are used, these strata are confined necessarily to the political reporting boundaries for which these data have been historically generated. The second-generation approach utilizes Landsat full-frame imagery along with climatological and soil information, to develop strata along naturally occurring boundaries and to determine the optimal segment allocations to each stratum. Such an approach was known from the outset of LACIE to be an improvement over the use of historical data, particularly in countries having limited historical information. However, it was not possible to implement this approach until late in Phase II because Landsat imagery for foreign countries was not available and techniques for discerning the small grains on the imagery were not fully developed. A year and one-half of data collection by Landsat and a similar period of image analysis experience in the LACIE have made implementation of such techniques possible.

ORIGINAL PAGE IS
OF POOR QUALITY

1.7 PHASE III CLASSIFICATION PROCEDURE CHANGES

Because a number of needed improvements were discovered during the LACIE Phase I and II evaluation of the first-generation Landsat data processing procedures, a second-generation classification procedure development effort was initiated at the end of Phase II. This effort was successfully completed and resulted in a set of design requirements for a procedure referred to as Procedure 1. These requirements were implemented at various stages throughout Phase III. The first stage of system delivery occurred in January of 1977. It provided the analyst the capability of selecting four-picture-element (pixel) line fields as training data (hence it was called the "Small Fields Procedure") and incorporated many improvements to machine clustering. The Small Fields Procedure was a cluster-based machine procedure. In Phase I and early in Phase II, the existing clustering procedures had been found faulty and were not used during a majority of Phases I and II. This prohibited any significant amount of multitemporal machine processing during the first 2 years of LACIE. In June 1977, two new systems were delivered with an implemented software system capable of supporting analysis with the second-generation method, Procedure 1. Procedure 1 is a highly automated, cluster-based procedure which has removed all required analyst functions except interpretation of the Landsat and ancillary data products for the purpose of labeling the spectral pixels or data as wheat or nonwheat.

1.8 WHEAT AND SMALL-GRAINS PROCEDURES

In Phases I and II, wheat could not be reliably differentiated from certain small grains. Thus, the LACIE determined only the percentages of small grains from Landsat data and then applied historical wheat/small-grains ratios to derive a wheat percentage for a LACIE sample segment. During Phase III, a procedure was developed for separating spring wheat from total small grains using Landsat data. The procedure, which is being tested in

North Dakota, is based on the crop calendar and general spectral characteristics of each category of small grains. While the spectral reflectance patterns of spring wheat and other small grains are similar, subtle but detectable temporal and spectral differences have been noted in the greenness and brightness of the different grains. Using the knowledge of these differences, quantitative spectral displays of spectral band combinations which relate to crop greenness are used by the analyst to separate the different classes of small grains (ref. 3).

1.9 CLIMATOLOGICAL YIELD MODELS

During Phase III, the yield models of the Center for Climatological and Environmental Assessment (CCEA) of the NOAA, somewhat modified from the Phase II models, continued to provide operational yield estimates used for aggregating production in the United States and the U.S.S.R. These models were applied in a somewhat different manner than in Phase II, however.

- The models were reconfigured (1) for removing overlap in coverage between modeled regions, (2) for achieving greater homogeneity within modeled regions, and (3) for extending coverage to include previously unmodeled areas. The areas included in each model for the United States and the U.S.S.R. are shown in figures 1-3, 1-4, 1-5, and 1-6.
- The models were operated for pseudozones, which are aggregates of Crop Reporting Districts (CRD's), rather than for CRD's as was done in Phase II. Tests indicate this did not cause a measurable degradation in yield estimates but did significantly improve the confidence in the estimate of the variance.

1.10 SECOND-GENERATION YIELD MODEL

The Feyerherm model developed at Kansas State University (KSU, ref. 4) was the only second-generation model available for

ORIGINAL PAGE IS
OF POOR QUALITY

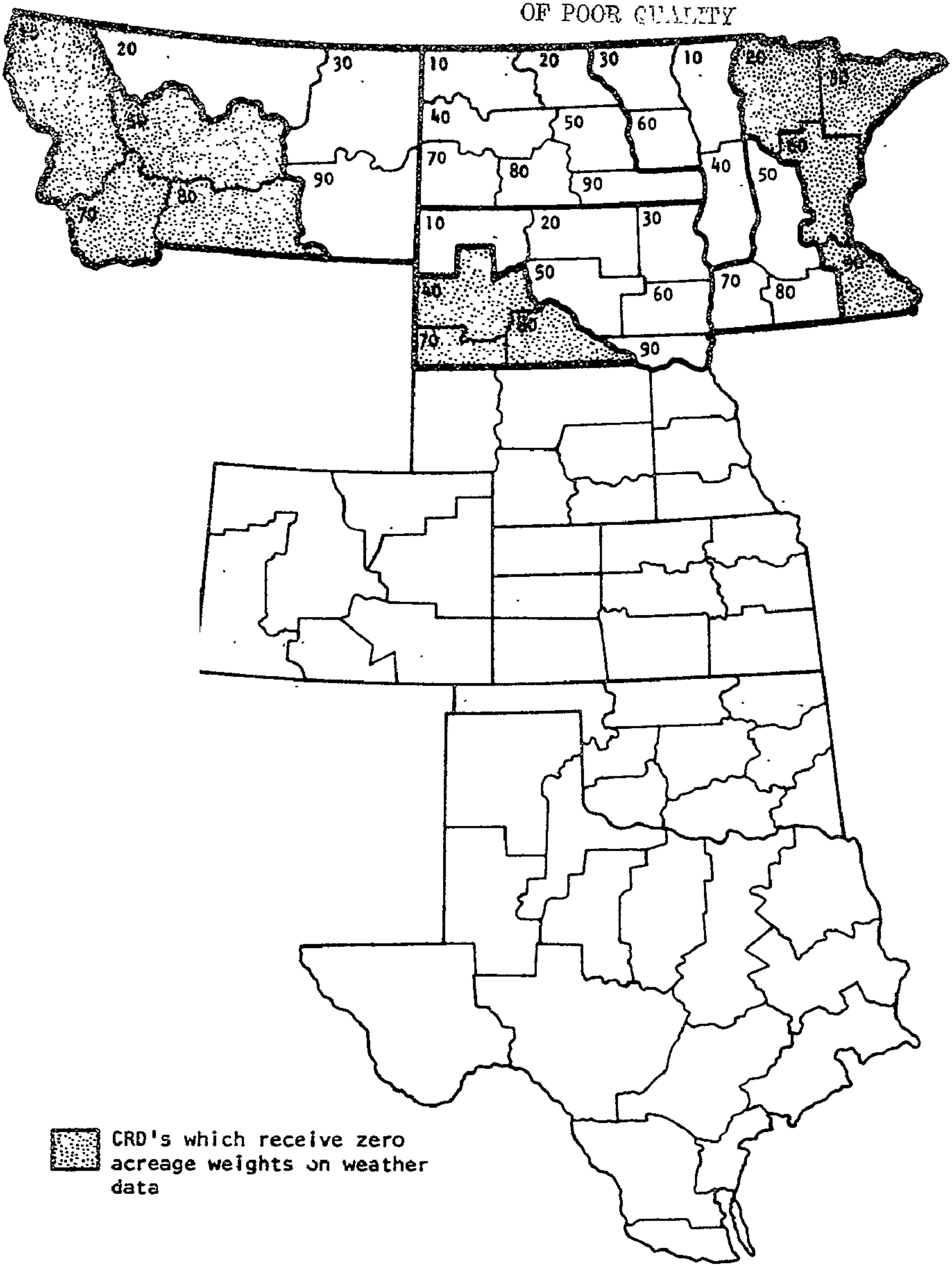


Figure 1-3.— U.S. Great Plains CCEA spring-wheat model boundaries.

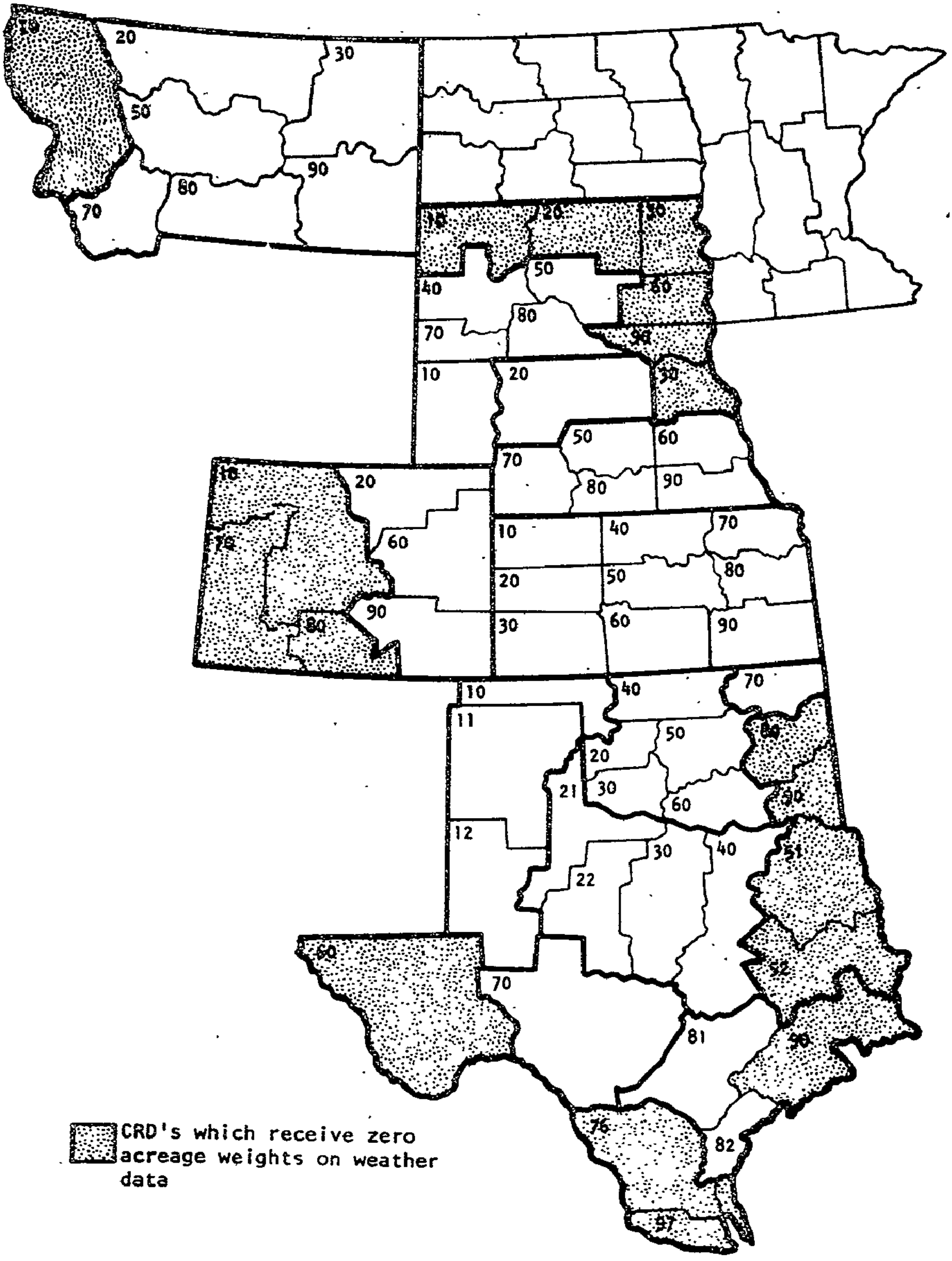


Figure 1-4.— U.S. Great Plains CCEA winter-wheat model boundaries.

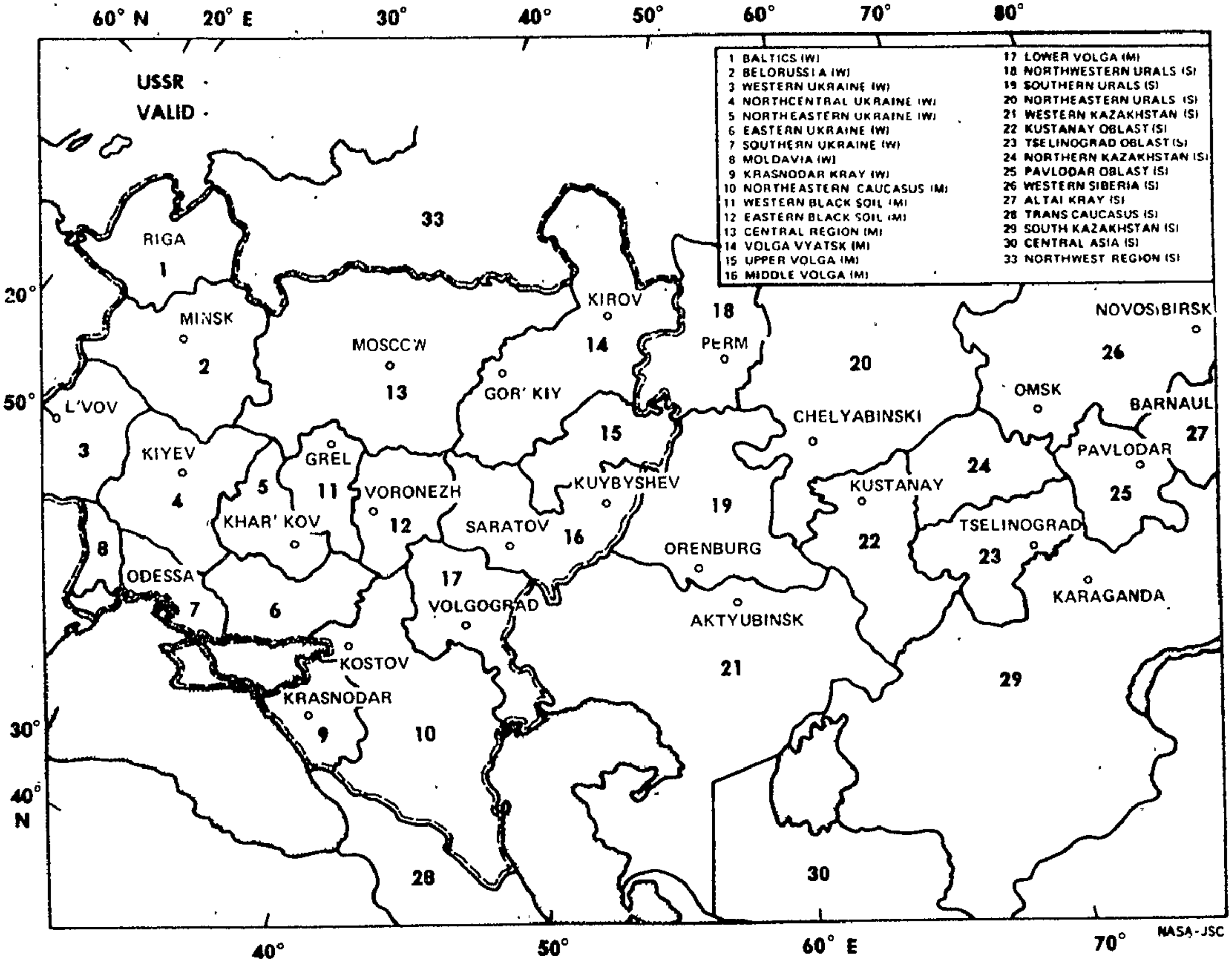


Figure 1-5.— U.S.S.R. CCEA winter-wheat yield model boundaries.

ORIGINAL PAGE IS
OF POOR QUALITY

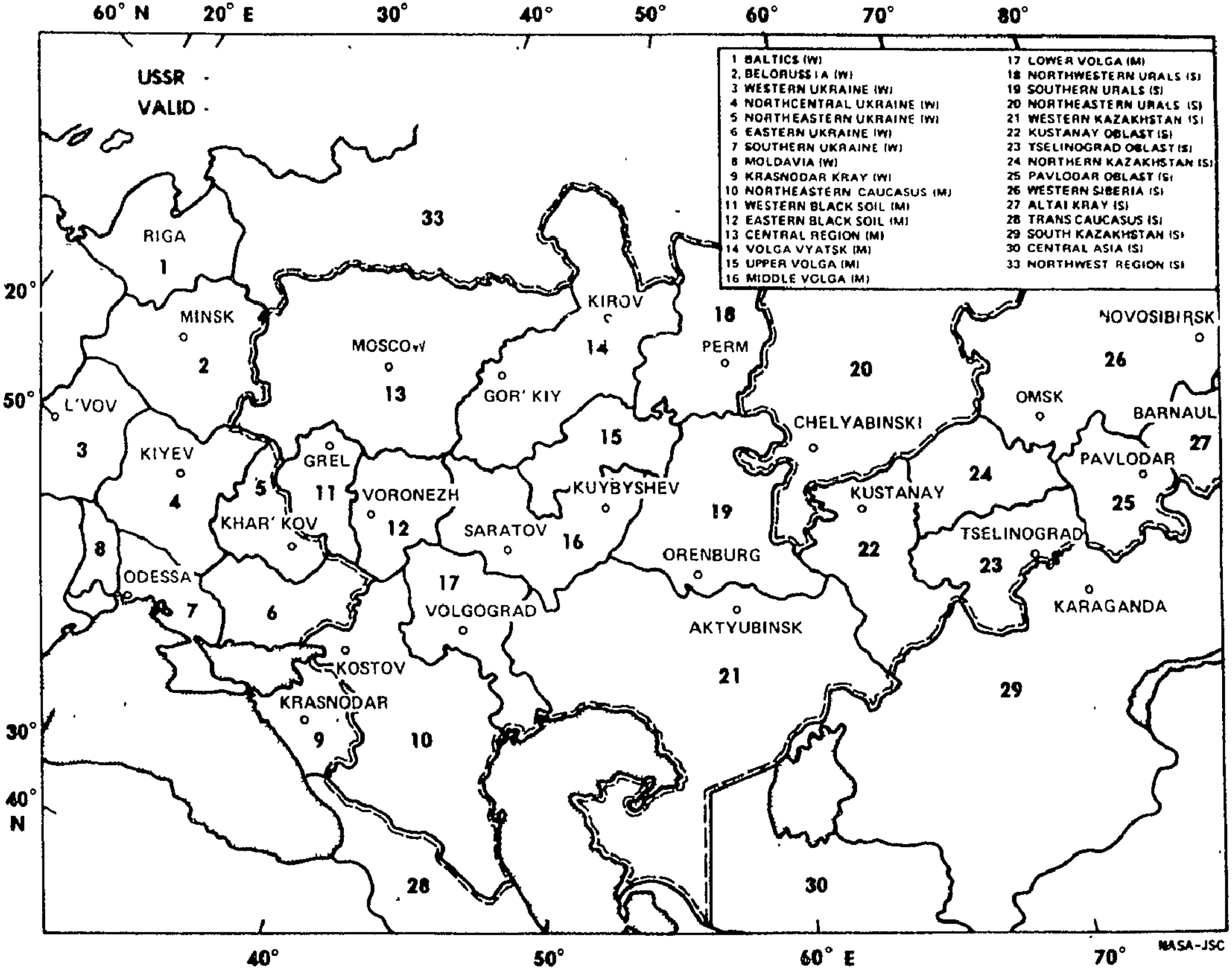


Figure 1-6.– U.S.S.R. CCEA spring-wheat yield model boundaries.

Phase III. Because this model was still in a stage of development where improvements or changes were being suggested frequently, its application to Phase III was in a limited pseudo-operational capacity. This pseudo-operation was performed as an *initial* evaluation of model prediction accuracy. (The term initial is emphasized because 1 year of operation is not considered a sufficient test.) The Feyerherm model was used also as a means of determining if the data system and the computer system could support the use of daily data. The winter-wheat model has been operated for the State of Kansas and the Khmel-Nitsky Oblast, U.S.S.R. This operation has shown that, while the input meteorological data appear adequate to operate the model, the model computer program needs revision in order to be utilized operationally on the Suitland, Maryland, computer system. The input/output structure of the current program is not compatible with the Suitland priority system and will not execute in a timely fashion. An investigation is underway to determine the feasibility of reprogramming the model for more timely operation.

The operation of the second-generation spring-wheat yield models has been deferred pending a test of a modified version of this model for the State of North Dakota and the Kurgan and Tselinograd Oblasts.

The Feyerherm model has been applied to partitions in Kansas, and an aggregation was performed using an area from the new sampling strategy. A procedure has been devised for applying the models to partitions in places where no historical yield data exist below the zone level, provided some information is known about soil yield potential in the partitioned area. Tests of this procedure are underway.

2. RESULTS OF EXPERIMENT

2.1 SUMMARY

The LACIE results at mid-Phase III are discussed in detail in terms of the agromet conditions which existed, the accuracy of estimates, and systems performance (sections 2.2, 2.3, and 2.4, respectively). These results are summarized briefly in this section.

The agromet conditions which existed in the U.S. Great Plains during the 1976-77 crop year were quite different from those of either Phase I or II in several respects. After starting the Phase III winter-wheat season with subsoil and topsoil moisture shortages, September rains provided the needed moisture for plant emergence and establishment. Abnormally cold weather and sparse precipitation in October caused plants to enter dormancy with thin stands and little vegetative cover. The overall moisture deficit, coupled with lack of snow cover and poor conditions of the stands, left many areas open to wind damage. The early cold and dry conditions were manifested in the Landsat data as very weak wheat signatures through February. Wheat signatures became more visible as the temperatures warmed and timely rains persisted through the spring.

LACIE acreage estimates were in close agreement with SRS estimates, and an operational system with a 14-day Landsat data turnaround could have produced an accurate acreage estimate (one which satisfied the 90/90 criterion) 1-1/2 to 2 months before harvest. Low yield estimates resulting from agromet conditions not taken into account in the yield models caused production estimates to be correspondingly low. However, both yield and production estimates satisfied the LACIE 90/90 criterion for winter wheat in the yardstick region.

The implementation of Procedure 1 resulted in more efficient, multitemporal processing of Phase III wheat segment data. By August 1, the number of acquisitions and analyses doubled that of the entire Phase II; per-segment analyst time was reduced significantly; and interactive reworking of segments was reduced to less than 1 percent, allowing more computer time for batch operations.

Yield models for the yardstick region were revised to eliminate data overlap areas, and additional models were developed for five regions in the U.S.S.R. LACIE early season results for hectarage (acreage) in the U.S.S.R. winter-wheat region are encouraging and are projected to be in reasonable agreement with the end-of-season U.S.S.R. estimates. The LACIE yield estimates for the U.S.S.R. are somewhat below the USDA Foreign Agricultural Service (FAS) estimates; however, it is too early in the season to obtain conclusive comparisons. The LACIE winter-wheat production estimates for the U.S.S.R. are within 5 percent of those of the FAS; however, the FAS estimates do not provide as reliable a gauge for within-season comparisons as do the U.S.S.R. estimates (which will be available 5 months after harvest).

Improvements in crop calendar estimates were made by providing for analyst feedback to the adjustable crop calendar (ACC). LACIE personnel prepared and published weekly meteorological summaries for use by CAMS analysts. Data for the summaries were furnished by the National Weather Service (NWS), the Environmental Technical Applications Center (ETAC), the CCEA, and foreign newspaper reports.

Because of the various technology modifications, the average turnaround time observed in Phase III cannot be used to project turnaround time for an operational system.

The RT&E program is responsive to the technical issues identified in Phase II. Two major tasks are being conducted in Phase III: the test and evaluation of the modified first-generation yield models and the test and evaluation of Procedure 1 performance.

2.2 AGRICULTURAL AND METEOROLOGICAL CONDITIONS

A developing wheat crop can appear a variety of ways to the analyst. Throughout the year and from one region to the next, the spectral properties of wheat are quite variable. The major characteristics of crop appearance to Landsat are its condition, the color of the background soil, the growth stage, and environmental factors such as Sun angle (time of year and latitude) and atmospheric haze. These factors are, in turn, strong indicators of the meteorology throughout the year. The analyst must recognize wheat as a product of diverse agromet conditions during the year, all of which must be considered as influential in attaining wheat identification accuracy. It is important, therefore, to review the significant conditions which existed during Phases II and III in order to better understand the performance of the LACIE estimation system and to compare results from one year to the next.

The conditions which existed during the 1976-77 U.S. Great Plains crop year (Phase III) were quite different from those of either Phase I or II. Before discussing the detailed results of Phase III, the Phase II agromet conditions in the yardstick area and their effects on Phase II results will be summarized. The presence of these conditions in Phase III will also be discussed.

2.2.1 PHASE II WINTER WHEAT

In the fall of 1975, the beginning of Phase II, inadequate moisture in much of the yardstick area caused uneven stands of wheat. Greater-than-normal moisture to these areas in November was too late in the season to establish stands and growth before the onslaught of winter winds. Lack of top growth, snow cover, and winter moisture allowed substantial crop damage because of topsoil removal by winds.

Although the winter-wheat crop broke dormancy early in most areas because of the dry winter, the persisting dry conditions caused growth progress to be slow in the U.S. yardstick area during early spring. In the southern states — Colorado, Kansas, Oklahoma, and Texas — heavy rains during middle to late April alleviated the drought, and crop growth was ahead of normal. Freezing temperatures in early May and subsequent cooler weather in eastern Kansas caught the crop at a critical stage of development, thus lowering yield potential. However, rains during May and June improved crop prospects in all areas of the U.S. southern Great Plains.

Harvesting in Texas began about mid-May. June rains in Kansas, Oklahoma, and Texas slowed the harvest effort; and, by July 4, only 44 percent of the Nation's acreage had been cut, as compared with the national average of 51 percent. Good weather during July permitted timely completion of harvest, except for the northwestern states where the crop normally matures later. As discussed in the LACIE Phase II Evaluation Report (ref. 5), the results of the early dry season, the delayed spring greening up, and late spring moisture produced early-season sparse wheat signatures which were misidentified as bare soil; in addition, many late greening crops were mistaken as spring crops. The anomalous condition created a significant acreage underestimate in Oklahoma and tended to bias the central plains winter-wheat estimate downward, although not significantly, in terms of the accuracy required to support the 90/90 criterion.

ORIGINAL PAGE IS
OF POOR QUALITY

2.2.2 PHASE II SPRING WHEAT

Regarding spring wheat in the yardstick region, 1976 crop seeding in major growing areas was completed much earlier than normal because of dry weather conditions in early spring. Seeding was virtually complete in Minnesota and South Dakota by mid-May and in North Dakota and Montana by May 25. Growth and development occurred ahead of normal in the Dakotas because of early seeding and dry conditions. Severe drought in parts of Minnesota and South Dakota caused a sharp reduction in yield in those areas, particularly in South Dakota where the SRS estimated a yield of 11 bushels per acre. Meanwhile, the LACIE South Dakota yield models were estimating 17 bushels and would have estimated 13 bushels per acre even if zero precipitation had been entered into the model. This model behavior in episodic situations, such as the South Dakota drought, tends to cause the prediction of yields that vary to a large extent from the average.

2.2.3 PHASE III WINTER WHEAT

The Phase III crop year was quite different from either the Phase I or the Phase II crop year. The 1976-77 winter-wheat crop started its growing season with both topsoil and subsoil moisture shortages over a large portion of the yardstick area. Precipitation amounts during the previous August were sparse, and drought conditions encompassed the major portion of the U.S. Great Plains. A series of rain-producing systems passed through the yardstick area in September, replenishing topsoil moisture and giving wheatfields the needed moisture for plant emergence and establishment. Rainfall amounts were generally above the normal in all areas except the Dakotas. Producers' reports were generally optimistic as plants responded to the generous September rains.

The month of October began with moderate temperatures; however, by the end of the first week, cold weather ensued and it became one of the coldest Octobers on record. Precipitation amounts were negligible, restricting the U.S. northern Great Plains to less than 25 percent of the normal precipitation amount. Late-seeded fields, faced with abnormally cold temperatures and sparse precipitation, showed very little additional growth after the initial cold wave; henceforth, they went into dormancy with thin stands and little vegetative cover.

November through January persisted in a pattern that brought cold, dry, arctic air spilling across the plains. The snow cover was variable over the U.S. northern Great Plains and very often nonexistent over the U.S. central and southern Great Plains, leaving plants vulnerable to winterkill. The overall moisture deficit, coupled with the lack of snow cover and poor conditions of the stands, left many areas open to wind damage.

The wheat production outlook improved considerably from March through winter-wheat harvest, as above normal temperatures and timely precipitation persisted through the spring. These conditions permitted recovery of the U.S. southern Great Plains winter-wheat crop to near-normal yields despite the cold winter and soil moisture shortages.

The early cold and dry conditions were manifested in the Landsat data as very weak wheat signatures through February. The warming temperatures and timely rains were generally in evidence as the wheat signatures became more visible. The LACIE Phase III acreage estimates increased considerably over those of Phase II from February to May, as shown in the CMR's (refs. 2, 6, 7, 8, 9) and in the CAS unscheduled reports (refs. 10, 11).

2.2.4 CROP CONDITION ASSESSMENT ACTIVITIES

A Crop Condition Assessment Team was formed as an *ad hoc* group composed of persons with agronomic expertise within various

ORIGINAL PAGE IS
OF POOR QUALITY

elements of the project. The team was charged specifically with assessing the influence of weather and other external factors on wheat conditions in LACIE countries. The assessments were made primarily from the meteorological and spectral data which would be routinely available to an operational crop assessment system.

The crop condition assessment reports were a routine part of the CMR's for the United States and the U.S.S.R. The reports qualitatively evaluated wheat responses and isolated other potential problems.

In the U.S.S.R., the team evaluated the potential for winterkill by closely watching the interaction between snow cover and cold temperatures. Temperature data were routinely available to provide extreme values, whereas snow cover was assessed from daily meteorological satellite imagery. At the end of the winter, the team was able to infer that winterkill had affected only a very small portion of the growing region and that losses of fall-sown grains to the cold would be less than normal.

While growing conditions in the U.S.S.R. winter-wheat areas were very good during the spring, abundant rainfall persisted through the ripening period and into the normal harvest time. During June, the team assessed that the U.S.S.R. wheat would have extensive disease problems, especially stalk rot, and a considerable amount of lodging could be expected. Weeds were also expected to develop in the moist soil as the wheat matured and stands became thinner. A visiting U.S. team, which had toured the winter-wheat region of the U.S.S.R. during June, returned with reports of extensive lodging and evidence of wheat rust and commented that harvesting would be difficult in some areas because of heavy weed infestation.

Because of a lack of resources in CAMS to evaluate Landsat imagery for indicators of crop condition, only minimal information could be elicited from the spectral data. The Green Index Number (GIN) has been developed as a tool to identify drought stress (ref. 12); and, as a result of this effort, data were provided to the team for use in delineating stressed spring wheat in both the U.S.S.R. and the U.S. Great Plains. The analysis allowed inferences to be made about the amount of the wheat yield which was likely to be below normal and provided a greater resolution than was possible with only the yield estimates for the crop regions.

In addition to the assessments prepared for the CMR, the team identified potential problem areas to CAMS and indicated where atypical wheat signatures might be expected. The team's attention was directed to a particular segment in the spring-wheat region of the U.S.S.R. which showed growing moisture stress as the season progressed, thus confirming suspected dryness. A comparison of sequential acquisitions of that particular segment is shown in figure 2-1.

The Crop Condition Assessment Team is a project resource to investigate any area having an agronomic problem.

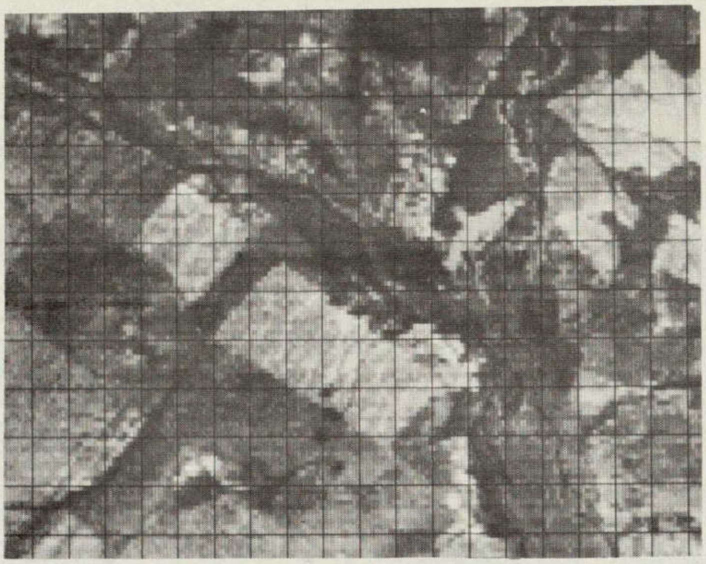
2.3 ACCURACY OF ESTIMATES

Determining the accuracy of LACIE Phase III estimates necessarily encompasses an examination of the many factors that affect wheat production estimates; that is,

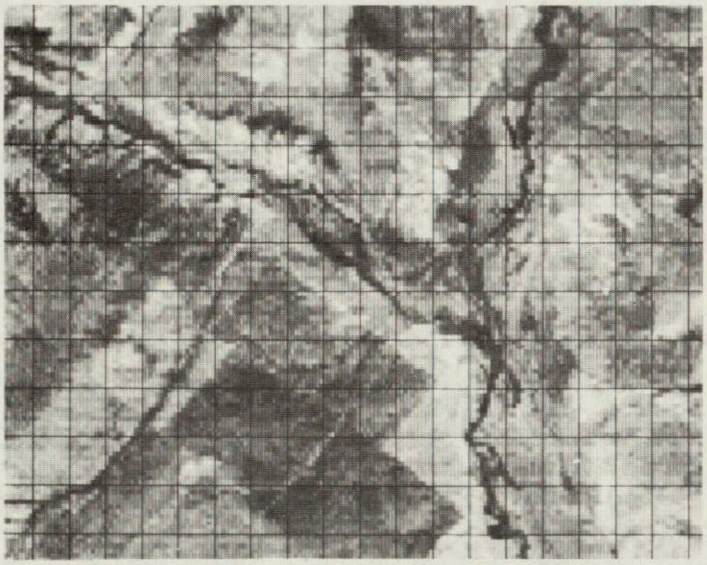
- The acreage and yield estimates, which are factors of the production estimate
- The accuracy of improved classification procedures utilizing Procedure 1

ORIGINAL PAGE IS
OF POOR QUALITY

NASA-S-77-11521



May 16, 1977



June 20, 1977

Kustanay, U.S.S.R.

Figure 2-1.– Example of deteriorating moisture conditions in U.S.S.R. spring wheat. Notice the decrease of the dark signature in the natural drainage ways and light signatures in the wheatfields.

- The accuracy of the blind-site estimates
- The accuracy of the crop calendar

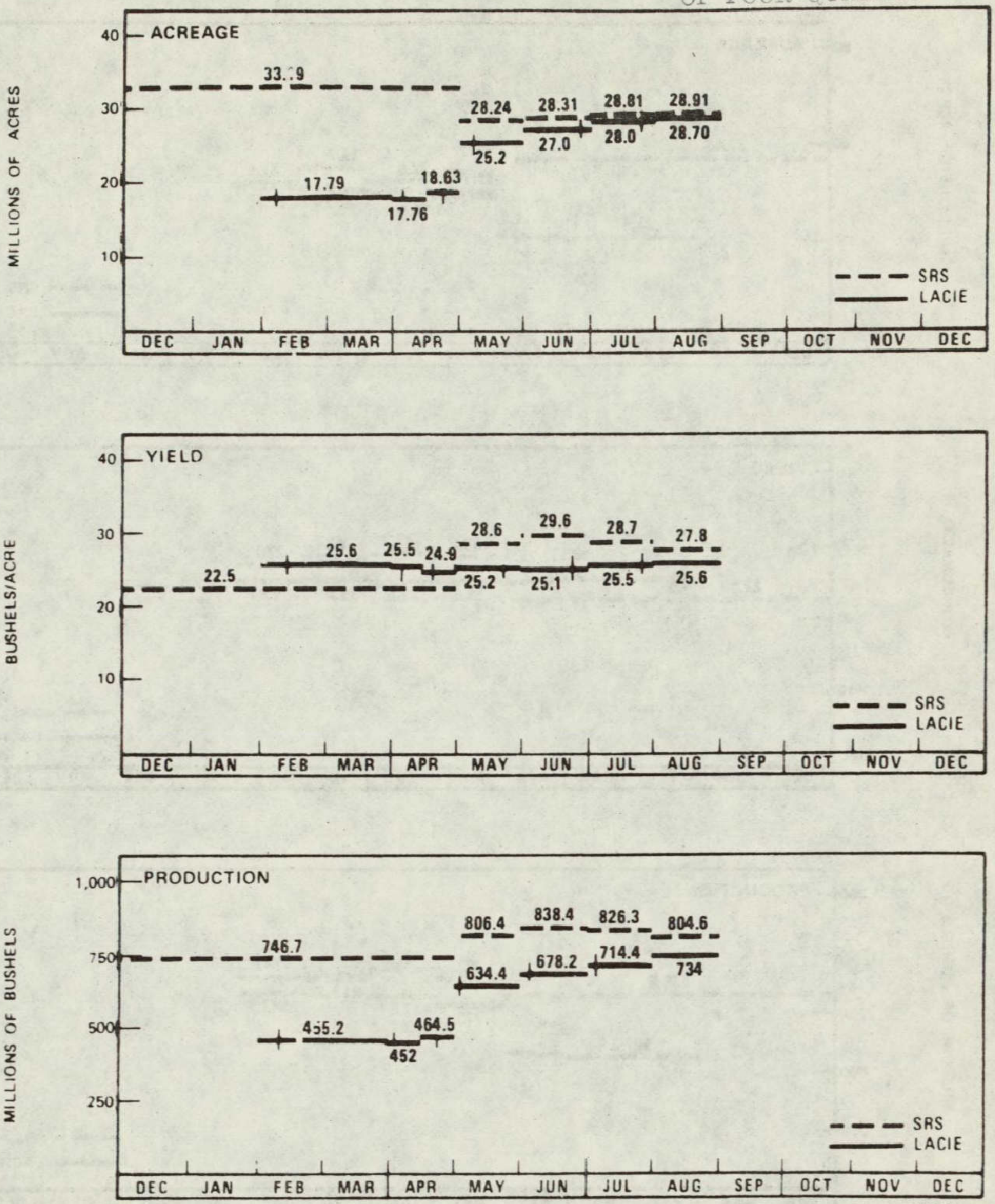
All of these and special studies, such as thresholding and sampling methods and the differentiation of wheat from small grains, form an integral part of the LACIE procedures which culminate in acreage, yield, and production estimates.

The Phase III results for the U.S. Great Plains, as reported in the August CMR (ref. 9), indicate that the LACIE production estimates for winter wheat at the seven-state level supported the 90/90 criterion for the yardstick area. The Phase III winter-wheat acreage estimates are significantly improved over those of Phase II and were supportive of the 90/90 criterion for production as early as June 1977 (ref. 7). The June CMR was based on Landsat data acquired through April 1977. It is projected that an operational system with a Landsat data turnaround of 14 days could have produced an acreage estimate to satisfy the 90/90 criterion not later than mid-May, some 1-1/2 to 2 months prior to harvest. For the first time in the three operational phases, a moderate but not statistically significant difference existed between the LACIE and the SRS yield estimates. The LACIE yield estimate was lower than that of the SRS; however, as can be seen from figure 2-2, this difference is not significant and decreased in July and August as a result of decreases in the SRS forecast. Because of the difference in yield estimates, a corresponding difference in production is apparent; however, the production estimate at the seven-state level supported the 90/90 criterion.

As shown in figure 2-3, early-season results for hectarage (acreage) in the U.S.S.R. winter-wheat region are very promising and are projected to be in reasonable agreement with the end-of-season U.S.S.R. estimates. Figure 2-3 shows that the LACIE yield estimates were lower but not significantly lower than the FAS

ORIGINAL PAGE IS
OF POOR QUALITY

PHASE III



LACIE – REAL TIME REPORTING

5 STATE – WINTER WHEAT

COLORADO, KANSAS, NEBRASKA, OKLAHOMA AND TEXAS

T329

8/29/77

Figure 2-2.– Comparison of LACIE and SRS acreage, yield, and production estimates of winter wheat in the yardstick region.

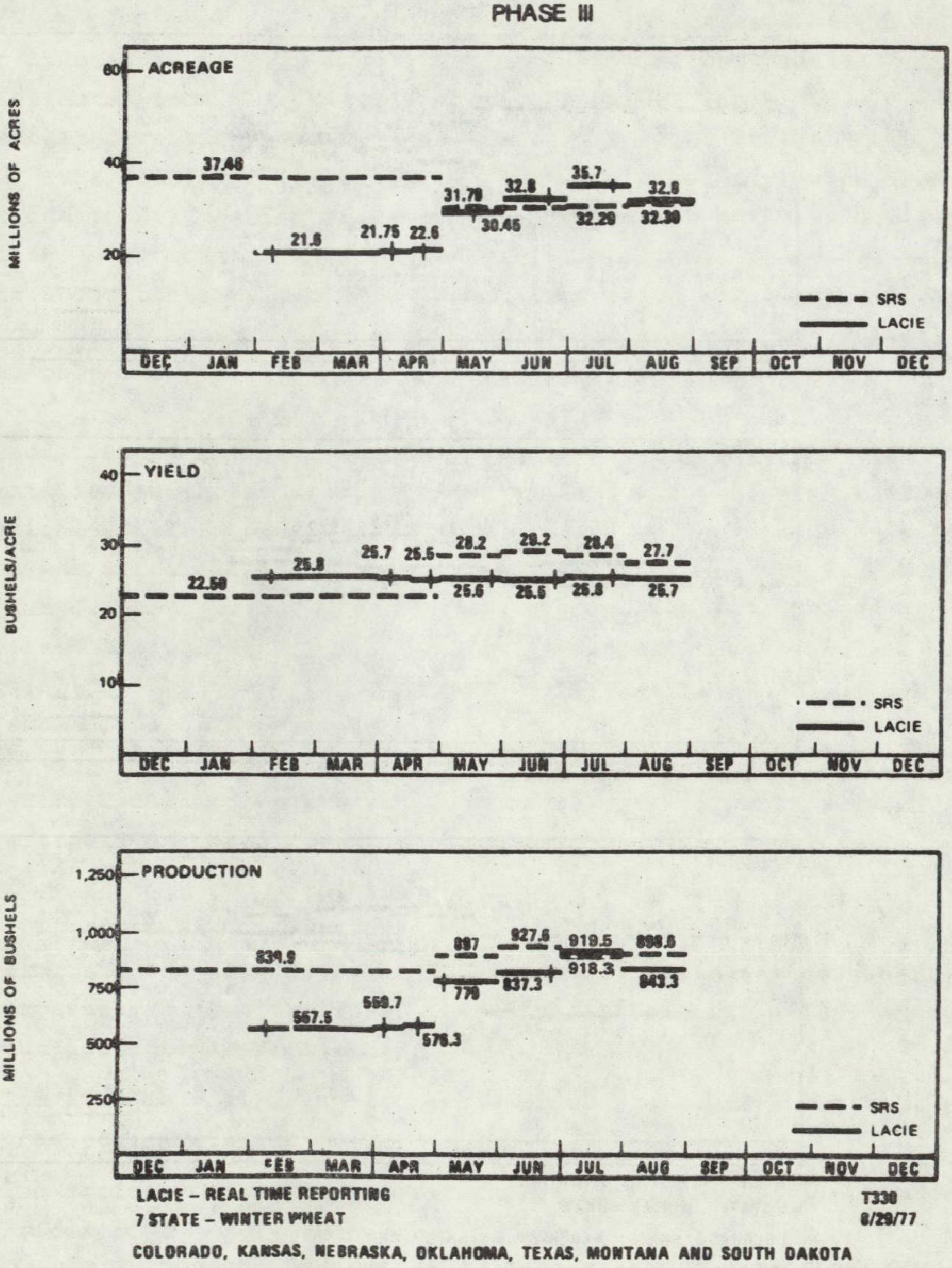


Figure 2-2.– Concluded.

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

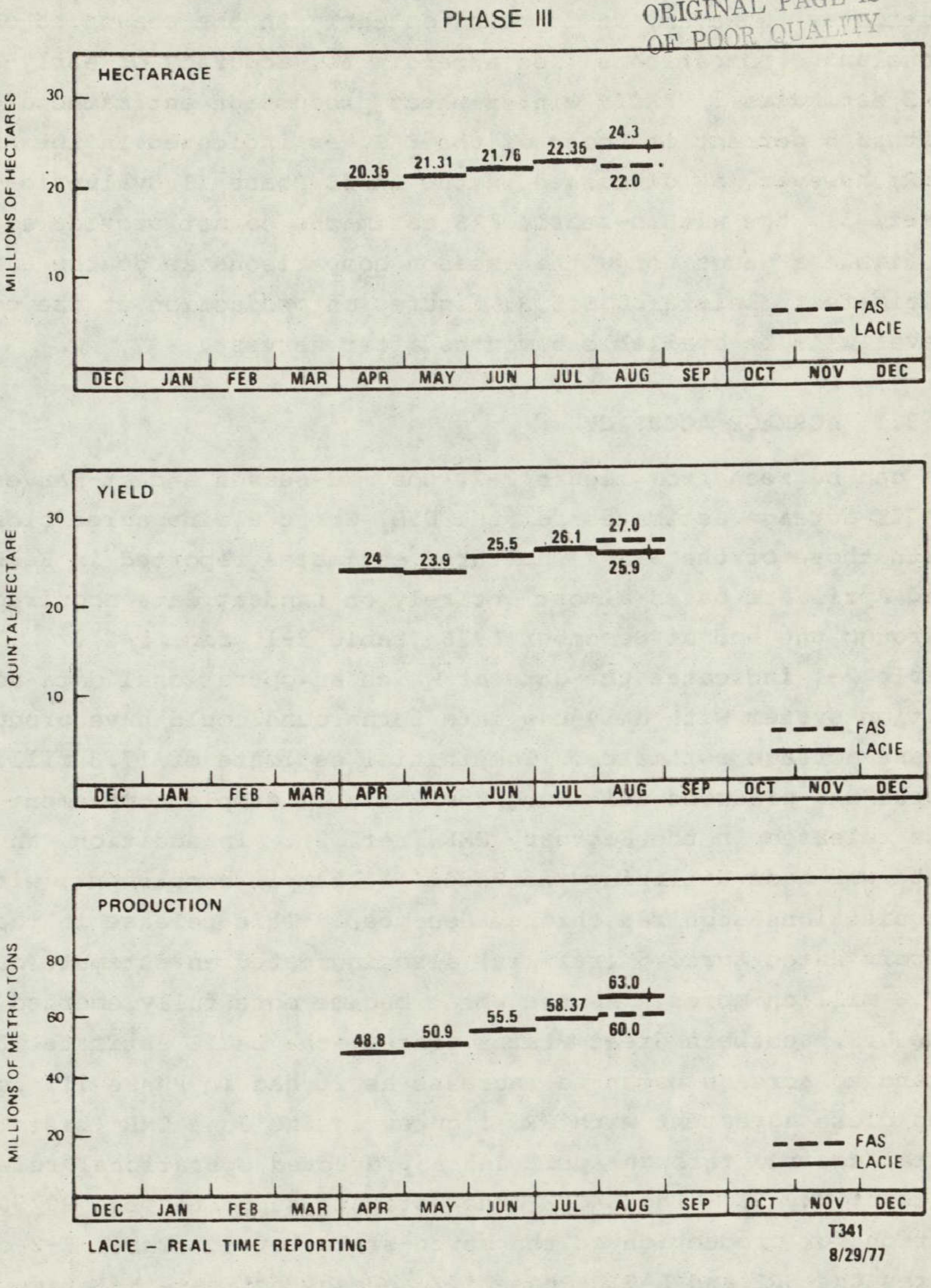


Figure 2-3.– Comparison of LACIE and FAS acreage, yield, and production estimates for U.S.S.R. winter wheat.

estimates; however, it is much too early in the season to obtain conclusive comparisons (see appendix for accuracy of early-season FAS estimates). LACIE winter-wheat production estimates are within 5 percent of those of the FAS, as indicated in the August CMR; however, as discussed in the LACIE Phase II Evaluation Report (ref. 5), the within-season FAS estimates do not provide as reliable a gauge for within-season comparisons as do the U.S.S.R. estimates. Initial U.S.S.R. figures on production at the country level will be available 5 months after harvest.

2.3.1 ACREAGE ACCURACY

As can be seen from figure 2-2, the mid-season and at-harvest LACIE acreage estimates for the U.S. Great Plains agree closely with those of the SRS. The three estimates reported in February and April are based almost entirely on Landsat data acquired through the end of December 1976 (table 2-1, fig. 1-2). Table 2-1 indicates the date at which an operational data acquisition system with a 14-day data turnaround could have produced these acreage estimates. The initial estimate of 17.8 million acres was produced utilizing the Phase II sample complement and was released in the February CMR (ref. 2). In addition, an estimate was made utilizing the Phase III sample complement, with acquisitions acquired through December. This release in the report dated April 6 (ref. 10) also indicated an estimate of 17.8 million acres. As the wheat became more fully emerged in the U.S. southern Great Plains states, the LACIE estimate of standing acreage began to increase as it had in Phase II, reaching close agreement with SRS figures in the June CMR (based on data acquired through April and a projected operational release date of May 5). The LACIE June estimate supported the 90/90 criterion for production at the seven-state level. Table 2-2 compares the SRS and LACIE Phase III acreage estimate relative differences and CV's for the U.S. southern Great Plains, as shown in the July CMR (June operational release date) with the

TABLE 2-1.- SCHEDULE OF CMR'S WITH OPERATIONAL DATA RELEASE DATES AND YIELDS

Country	Original CMR date	Projected operational release date	Yields used in report
United States	February 8	January 11	February 1
	April 6*	January 11	February 1 and March 1
	April 22	February 12	April 1
	May 9	April 11	May 1
	June 9	May 5	June 1
	July 11	June 14	July 1
U.S.S.R.	March 30	January 25	March 1
	May 2	March 2	April 1
	June 3	May 8	May 1
	July 1	May 25	June 1

---

*Release of February report.

TABLE 2-2.– COMPARISON OF COEFFICIENTS OF VARIATION AND RELATIVE DIFFERENCES BETWEEN LACIE AND SRS ACREAGE ESTIMATES AS OF JULY REPORT

Region	Phase II – July 1976		Phase III – July 1977	
	RD, % (a)	CV, %	RD, % (a)	CV, %
Colo.	23.3	25.0	20.3	13.2
Kans.	-2.8	6.0	-4.6	5.0
Nebr.	27.4	11.0	12.2	12.4
Okla.	-56.5	15.0	-23.5	8.5
Tex.	-8.9	15.0	-2.0	11.6
5-state average	-4.5	5.0	-3.0	3.9

[a] Relative difference.

equivalent Phase II information. The comparison for each state shows that the relative differences observed in Phase III either are comparable or are significantly reduced from the relative differences observed in Phase II.

2.3.2 YIELD ACCURACY

While statistical analysis indicates that the observed relative difference of -7.8 percent between the LACIE and the SRS August winter-wheat yield estimates was not significant, a tendency existed in the 1977 U.S. crop year to predict yields which were lower than those of the SRS. Table 2-3 shows the LACIE winter-wheat yield estimates in July are below those of the SRS in every winter-wheat state in the U.S. Great Plains except Kansas. The relative difference in July of -12.5 percent at the U.S. southern Great Plains (five-state) level is primarily the result of large relative differences of -30.7 percent in Oklahoma and -23.2 percent in Texas.

A term-by-term analysis of the CCEA yield model indicates two primary contributing factors to the underestimates in the States of Oklahoma and Texas. In both states, the trend term of the CCEA model has been selected to show no average increase in yield since 1960. On the contrary, ancillary data show that an irrigated winter-wheat area in Texas is now producing almost 25 percent of the total winter-wheat acreage. Nearly all of this additional irrigated acreage has been introduced since 1960. The weather terms in the Texas model did not alter the yield estimate significantly from trend. Thus, it is likely that the constant trend since 1960 is a major contributor to the underestimate in Texas. It is noteworthy that Texas yield also was underestimated by 17.6 percent in the 1976 crop year. In Oklahoma, the weather terms in the yield model were also, in addition to a constant trend term, a factor in the underestimate. The model underestimate in Oklahoma resulted mainly from below-normal

TABLE 2-3.– COMPARISON OF COEFFICIENTS OF VARIATION AND RELATIVE DIFFERENCES BETWEEN LACIE AND SRS YIELD ESTIMATES AS OF JULY REPORT

Region	Phase II – July 1976		Phase III – July 1977	
	RD, % (a)	CV, %	RD, % (a)	CV, %
Colo.	-22.2	17.	-2.2	14.8
Kans.	6.1	9.	-7.6	9.7
Nebr.	0	12.	-8.6	9.3
Okla.	-4.8	10.	-30.7	10.7
Tex.	-12.3	12.	-23.2	10.1
5-state average	0.8	5.	-12.5	5.5

[a] Relative difference.

precipitation between August and February (over the winter period), a March precipitation deficit relative to potential evapotranspiration, and an above-average May precipitation. The weather factors which most likely contributed to the improved Oklahoma yields and which were overlooked by the LACIE yield models were the above-normal April temperatures and precipitation and the temporal distribution of the May precipitation in Oklahoma. The April temperatures were about 5° above normal in Oklahoma, which would make them nearly ideal for wheat (upper 60°'s F); and 3 inches or more of well-distributed precipitation occurred in April and 4 inches fell in May. Good April rainfall amounts following moisture deficit periods, such as those which occurred during the preceding winter months and even during the previous season, typically give an extra stimulus to yield by encouraging more extensive crop rooting. This results in improved utilization of nutrients when moisture becomes available. The monthly averaging of precipitation in the Oklahoma model also created an unrealistic response to the rather well-distributed May rainfall, which nearly doubled the average May precipitation. Since Oklahoma wheat is harvested at the end of May and the first of June, large rainfall amounts near the end of May tended to reduce yields. However, a majority of the 1977 May precipitation came in mid-May, with lesser amounts in late May. The mid-May precipitation came during the heading to ripening period for the Oklahoma winter wheat and thus contributed to increased yields, as opposed to the decrease predicted by the LACIE models.

Thus, in the third year of LACIE, the performance of the LACIE models at subregional levels indicates that these models can and should be improved by the use of daily meteorological inputs, more complex model forms, and satellite data to augment the sparse ground station network.

2.3.3 PRODUCTION ACCURACY

As of the June CMR (based on Landsat data acquired through April), winter-wheat production estimates for the U.S. Great Plains supported the 90/90 criterion at the seven-state level. With a 14-day instead of the LACIE experimental 30-day Landsat data turnaround, the LACIE could have produced such an estimate as early as mid-May or approximately 1-1/2 to 2 months before completion of harvest. A comparison of Phase II and III CV's and relative differences between LACIE and SRS production estimates as of the July CMR (June operational release date) for each of the states and the five-state U.S. southern Great Plains area is presented in table 2-4.

As of the August CMR, the estimated CV of production for the U.S. Great Plains winter-wheat estimate was 6.4 percent, as compared with 7 percent reported for the same period in Phase II. The random error was divided between acreage and yield at this level; the acreage estimate had a CV of 4.0 percent and the yield a CV of 5.2 percent. The LACIE Phase III acreage estimate was in close agreement with SRS figures; but, in contrast to Phases I and II, the corresponding yield estimate was somewhat below that of the SRS. The relative difference between the LACIE acreage estimate and that of the SRS was 1.3 percent. This, combined with the negative relative difference in yield of -7.8 percent at the seven-state level, resulted in a -6.6 percent relative difference in production observed at the seven-state level. Statistical analysis indicates these differences are not significant.

2.3.4 PHASE III CLASSIFICATION USING PROCEDURE 1

Three major CAMS systems deliveries occurred during Phase III: System software to support Procedure 1 has been implemented on the CAMS IMAGE 100 Hybrid System utilizing the LACIE/ERIPS on the IBM 360/75 computer for batch processing. Procedure 1

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 2-4.— COMPARISON OF COEFFICIENTS OF VARIATION AND RELATIVE DIFFERENCES BETWEEN LACIE AND SRS PRODUCTION ESTIMATES AS OF JULY REPORT

Region	Phase II — July 1976		Phase III — July 1977	
	RD, % (a)	CV, %	RD, % (a)	CV, %
Colo.	6.0	30.	18.4	19.7
Kans.	3.7	11.	-12.4	10.9
Nebr.	27.3	16.	4.6	15.7
Okla.	-64.3	18.	-61.1	13.6
Tex.	-22.2	17.	-25.4	13.9
5-state average	-3.7	7.	-15.7	[b]6.6

[a] Relative difference.

[b] This is an approximation since the estimate is not given in the CMR.

ORIGINAL PAGE IS
OF POOR QUALITY

software was delivered in two stages on the LACIE/ERIPS during Phase III — once in January and again in June. The complete Procedure 1 software capability was achieved in November 1977.

Prior to the January delivery on the LACIE/ERIPS, Phase II procedures were employed to analyze Phase III data. A complete version of Procedure 1 with interactive displays was delivered on the IMAGE 100 in June. The IMAGE 100 served as the prototype and training system for the USDA Application Test System (USDA ATS) for LACIE Transition Year activities.

The IMAGE 100 will be used to classify 30 Canadian blind-site segments and 50 U.S.S.R. spring-wheat segments. Results will not be reported here. However, preliminary testing of Procedure 1 on the IMAGE 100 using Phase II blind-site data has produced encouraging results.

Procedure 1 was tested extensively utilizing a simulated system developed within the LACIE RT&E effort prior to its delivery on the LACIE/ERIPS and the IMAGE 100 in Phase III. The hybrid system test results verified (1) that, with ground-acquired training data, Procedure 1 was superior to the Phase II machine processing procedures and (2) that Procedure 1 produced estimates with significantly reduced bias and variance in comparison to the Phase II procedures. Preliminary indications are that operations utilizing Procedure 1 will significantly improve estimates in segments with small fields and that, generally, Procedure 1 is performing well. However, quantitative assessment of Phase III operational data has not yet been made comparing the performance of Procedure 1 with the Phase II field training processing procedures.

It is apparent that, for the first time in LACIE, a means for successfully processing multitemporal data has been provided.

ORIGINAL PAGE IS
OF POOR QUALITY

Using Procedure 1, all segments are now processed multitemporally whereas, in Phase II, only limited manual multitemporal processing was done. The improved clustering capability developed for Procedure 1 is also functioning well; however, some problems have been observed for segments with more than two acquisitions. These could well be the result of misregistration.

Procedure 1 also offers other significant capabilities such as quantitative spectral aids and trajectory plots to assist the analyst in labeling. Preliminary results indicate that these aids have improved analyst labeling accuracy, not only for small grains but also in the discrimination of wheat from small grains. The evaluation of segment results also has been greatly aided by Procedure 1. The analyst examines the classification results for consistency between the classifier dot labels and the machine labels; this provides a quantitative procedure for judging a segment estimate as acceptable or nonacceptable.

Several technical issues regarding the use of Procedure 1 have been identified in Phase III and will be discussed in section 2.4.1.

2.3.5 ACCURACY OF BLIND-SITE AND INTENSIVE-TEST-SITE ESTIMATES

As was the case in Phase II, LACIE Phase III blind-site estimates of standing winter wheat tended to be lower than the ground-observed estimates of planted wheat during the early and mid-season. However, the Phase III mid-season standing winter-wheat estimates were considerably closer to the SRS ground-observed proportions than they were in Phase II. A comparison of LACIE area estimates of winter wheat with those of the SRS shows closer agreement during mid-season for both phases. In the early season, classification error contributed more to winter-wheat acreage estimation error than did sampling error. The objective thresholding procedure applied during mid-season improved classification accuracy. This brought the LACIE midseason acreage estimates closer to those of the SRS; and, as a result, sampling error became the greater contributor to acreage estimation error.

In sites with less than 10 percent winter wheat, LACIE tended to overestimate the winter-wheat proportions, as shown in the comparison of LACIE and SRS acreage estimates in South Dakota. Volunteer wheat, pastureland, and some spring small grains (such as barley and spring wheat) were misidentified as winter wheat. In the Kansas and Texas intensive test sites (ITS's), some of which are representative of western Oklahoma, atypical wheat signatures (purplish blue and mottled brown) were acquired, which caused early-acquisition signatures of late-planted and late-developing stands to be missed. In Texas, dryland winter wheatfields were also being misidentified whereas the irrigated, fertilized, winter wheatfields were identified correctly. These omission errors are reflected in the underestimation by LACIE of winter-wheat acreage for Texas when compared to SRS estimates.

2.3.6 CROP CALENDAR MODEL ACCURACY

Crop growth stage estimation based on current year weather conditions serves two vital components of the LACIE: the CAMS and the Yield Estimation Subsystem (YES). Initially, the CAMS utilizes the crop growth information early in the year to determine whether or not the wheat is emerged sufficiently to be detectable. Once the Robertson Biometeorological Time Scale (BMTS) model predicts the crop to have emerged (Robertson stage 2.0, ref. 13), analysis of the segment for wheat percentage is initiated. The winter-wheat crop is monitored also to ascertain whether or not it has emerged from dormancy. In some more northerly regions of the winter-wheat-producing states of the U.S. Great Plains, crop estimates are not attempted during dormancy because the canopy is too sparse. The next major growth period of interest to CAMS is the period after dormancy to heading, where the analyst relies on the Robertson stage to ascertain the approximate expected intensity of the wheat vegetation signature in comparison to other spring planted crops. Heading to senescence or maturity is another key stage in the separation of wheat from other vegetation. During this stage, the appearance of the wheat is significantly different from other vegetation types. Senescence to harvest and postharvest is very important to the analyst, inasmuch as the Landsat acquisitions during this period permit him to verify his early-season identifications of wheat. (Only wheat matures and is harvested during this period.)

This very general description of the crop calendar function in CAMS aids in qualitatively understanding the effects of growth stage prediction errors. For example, if the Robertson model predicts full emergence at a date earlier than crops are fully emerged (growth model is ahead of actual progress), CAMS will analyze the segment in a period when some amount (depending on the magnitude of the growth model prediction error) of the wheat is incompletely emerged. Since incompletely emerged wheatfields will go undetected by the analyst, the growth model prediction error can result in a negative bias in the segment proportion estimate. In all cases, if the model predictions run too far ahead of the actual growth stage, the analyst will anticipate an onset of changing signatures within the segment, which will not occur at the predicted rate. Thus, if the growth model predicts 90-percent senescence within the segment and the analyst bases his labeling decision on this fact, certain fields could be discarded as being nonwheat because a senescent signature was expected and the analyst did not observe a change.

Although the interactions between the growth model prediction errors and CAMS errors are not quantified, substantial prediction errors in the model can result in substantial errors in analyst labeling. The key issues for crop growth model research are addressed in section 2.6.6.

The currently implemented operational yield models in LACIE do not depend on the crop growth model. However, the response of wheat yield to meteorological conditions is known to depend quite strongly on the growth stage at which these conditions are present. For example, high temperatures after wheat maturity do not affect yields in the same way as they do during heading. The second-generation yield models being evaluated for LACIE in Phase III depend on the crop growth models; and the effects of certain meteorologically related variables are weighted differently, depending on the estimated growth stage of the plant. Thus, errors in the growth model can strongly influence the yield estimation error; e.g., if high temperatures are experienced the last 2 weeks in May in an area where heading is occurring and the growth model (running fast) is predicting that the crop is ripe, the second-generation yield models will fail to predict the actual reduction in yield.

As stated, the relationship between the growth model prediction errors and the yield estimation errors is not completely understood, and their effects have not been quantified.

The Accuracy Assessment effort within LACIE has designed an evaluation of the crop growth models, utilizing ground-acquired information from ITS's in the yardstick region. This evaluation was conducted over 8 winter-wheat ITS's in Kansas and Texas during Phase II and was expanded to include 22 ITS's throughout the United States in Phase III and 11 ITS's in Canada (figs. 2-4 and 2-5).

Within each of the U.S. ITS's, the average ground-observed growth stage for the wheat crop is calculated from periodic

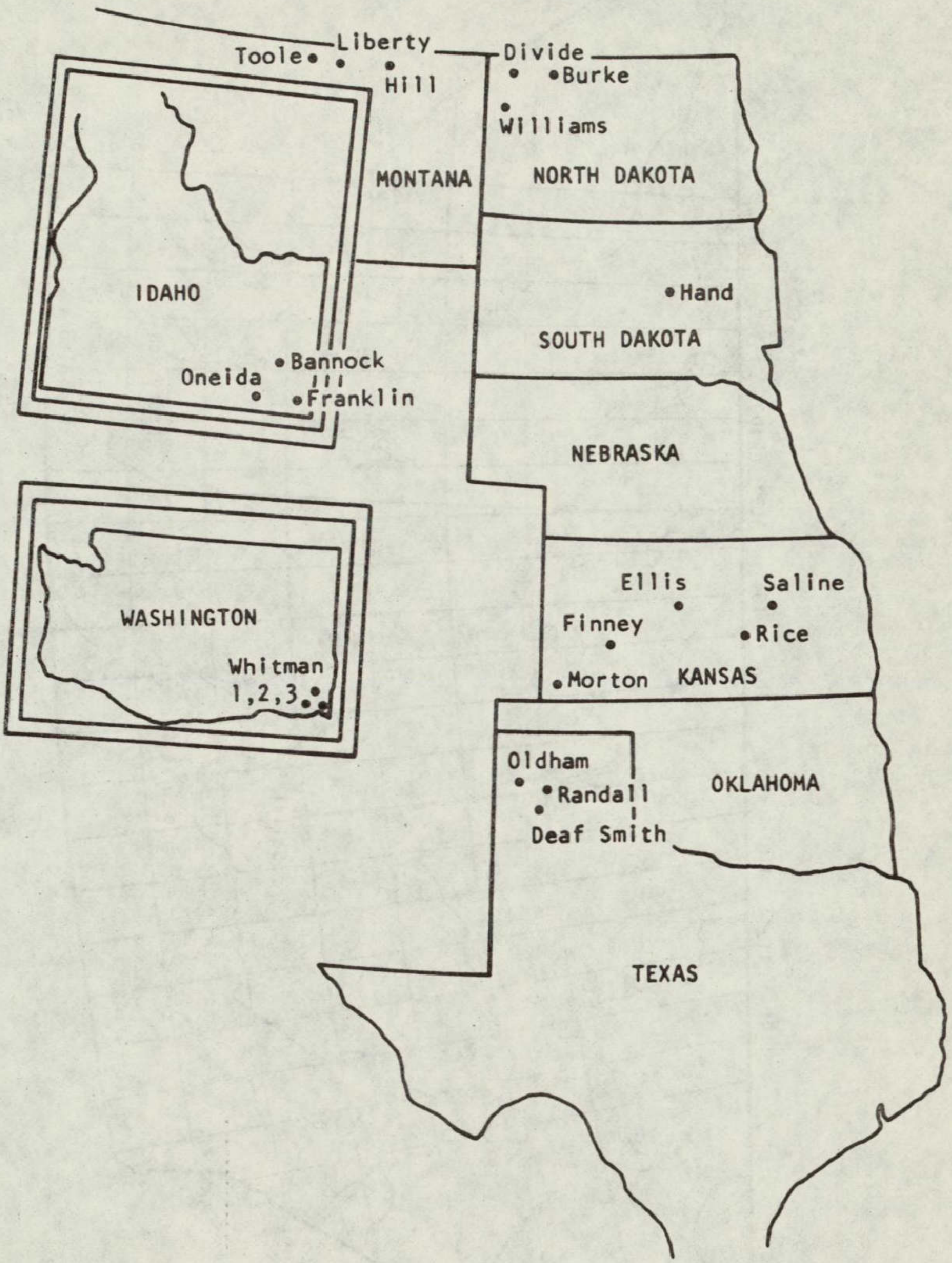


Figure 2-4.— Map of U.S. wheat-producing areas showing intensive test sites.

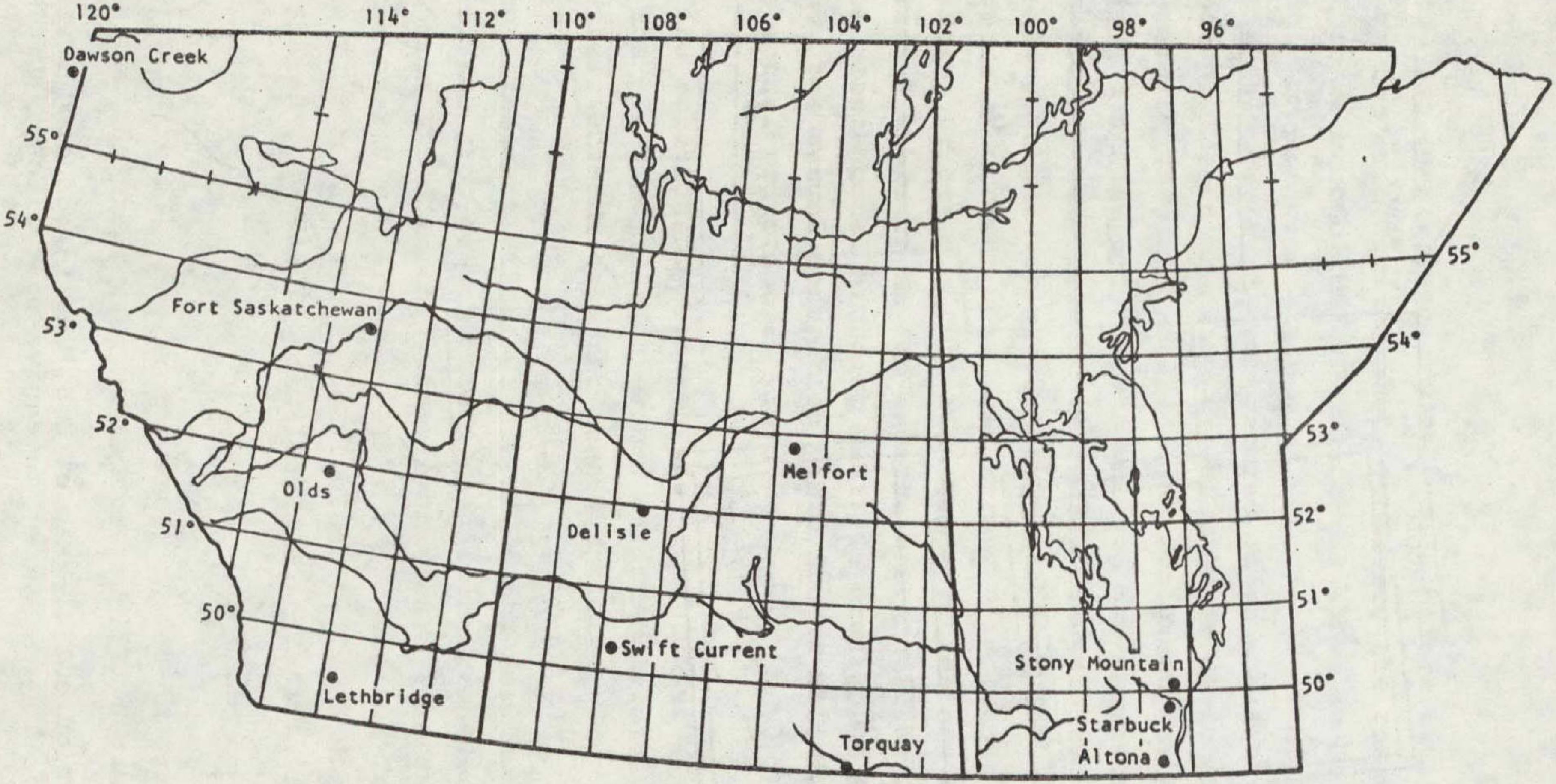


Figure 2-5.- Map of Canada showing intensive test sites.

field-by-field observations obtained by personnel from the USDA Agricultural Stabilization and Conservation Service (ASCS). ASCS personnel record detailed information regarding each field on the form shown in figure 2-6. The observer specifies the growth stage of each field to be one of the 10 stages listed on this form. All sites are visited each 18 days by ASCS field personnel, except for the Finney County, Kansas, and Hand County, South Dakota "supersites," which are visited each 9 days. The 11 ITS's in Canada are monitored each 18 days by personnel from the Canadian Agriculture Department.

The crop calendar model used by LACIE is a modification of the BMTS developed by Robertson (ref. 13). The Robertson BMTS estimates the stages for the progress of wheat crop development from planting to harvest (table 2-5). Daily maximum and minimum temperatures and day length are variables used to implement this model, which is often referred to as the ACC.

All of the growth stages defined by Robertson in the BMTS model development are not easily observable by field personnel. For example, BMTS 3.0, jointing, can be observed only by plant dissection. Thus, a different set of stages has been developed for ground observations. The ground-observed growth stage of each ITS must be developed by relating the ITS growth stage observations to the related BMTS stage. After planting, the earliest stage at which there is no ambiguity in this relationship is at heading. The BMTS stage 3.0, jointing, cannot be easily observed and is known to occur after tillering and before booting, which are observable by ground personnel. Thus, jointing is estimated by extrapolating between these observations. An error as large as a few days is customary in relating ground observations to BMTS stages. It should be kept in mind that heading is the most valid comparison as the results of the ACC are reviewed.

TABLE 2-5.— ROBERTSON BMTS AND OBSERVED ITS WHEAT PHENOLOGICAL STAGES

Stage	Robertson BMTS	ITS	Description
Planted	1.0	01	Planted
		02	Planted, no emergence
Emergence	2.0	03	Emergence
Jointing	3.0	04	Tillering, prebooting, prebudding
	3.5	05	Booted or budded
Heading	4.0	06	Beginning to head or flower
	4.5	07	Fully headed or flowered
Soft dough	5.0	08	Beginning to ripen
Ripening	6.0	09	Ripe to mature
Harvest	7.0	10	Harvest

GROUND TRUTH PERIODIC OBSERVATION FORM

TEST SITE # 40 (WHITMAN (2), WASHINGTON)
OBSERVATION 09

MONTH DAY
LANDSAT PASS DATE -8/24/77
OBSERVATION DATE -8/23/77
RAINFALL SINCE LAST OBSERVATION _____ IN.

LAND USE CODES
100-SPRING WHEAT
200-BARLEY
300-OATS
400-WINTER WHEAT
500-GRASSES/PASTURE
600-OTHER CROPS
601-RAPESEED
602-RYE
604-FLAX
607-CORN
617-SOYBEANS
618-COTTON
700-SUMMER FALLOW
900-UNKNOWN CROPS

GROWTH STAGES
01-NOT PLANTED
02-PLANTED NO EMERGENCE
03-EMERGENCE
04-TILLERING, PREBOOT, PREBUD
05-BOOTED OR BUDDED
06-BEGINNING TO HEAD OR FLOWER
07-FULLY HEADED OR FLOWERED
08-BEGINNING TO RIPEN
09-RIPE MATURE
10-HARVESTED
11-DOES NOT APPLY

GROUND COVER (%)
1- 0- 19
2-20- 39
3-40- 59
4-60- 79
5-80-100

SURFACE MOISTURE CONDITIONS
1-DRY
2-DAMP
3-WET
4-STANDING WATER

WEED GROWTH
1-NEGLIGIBLE
2-SLIGHT
3-MODERATE
4-HEAVY

FIELD OPERATIONS
01-BARE GROUND
02-BARE DISKED/CULTIVATED
03-BARE PLOWED
04-BARESEEDED
05-STANDING STUBBLE
06-STUBBLE DISKED/CULTIVATED
07-STUBBLE PLOWED
08-STUBBLE SEEDED
09-BURNED
10-GRAZED
11-WINDROWED OR SWATHED
12-MOWED OR COMBINED
13-STACKED OR BALED
14-OTHER

GROWTH/YIELD DETRACTANTS
01-SALINITY
02-INSECTS
03-DISEASE
04-DROUGHT
05-MOISTURE
06-WIND
07-HAIL
08-FROST
09-BIRDS
10-POT HOLES
11-UNEVEN STAND
12-WEEDS
13-WINTERKILL
14-LODGING
15-OTHER

STAND QUALITY
1-POOR
2-BELOW AVERAGE
3-AVERAGE
4-ABOVE AVERAGE
5-EXCELLENT
6-DOES NOT APPLY

FIELD NO.	ACREAGE	LAND USE CODE	GROWTH STAGE (CIRCLE ONE)	GROUND COVER (CIRCLE ONE)	PLANT HEIGHT (INCHES)	SURFACE MOISTURE (CIRCLE ONE)	WEED GROWTH (CIRCLE ONE)	FIELD OPERATIONS (CIRCLE ONE)	GROWTH/YIELD DETRACTANTS (CIRCLE ONE)	STAND QUALITY RATING (CIRCLE ONE)	COMMENTS
43	233.7	414	01 02 03 04 05 06 07 08 09 [✓ 10] 11	1 2 3 4 [✓ 5]	10	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 02 03 04 [✓ 05] 06 [✓ 07] 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	[✓ Y] N
20	511.6	700	01 02 03 04 05 06 07 08 09 10 [✓ 11]	[✓ 1] 2 3 4 5	-	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 [✓ 02] 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
31	455.0	414	01 02 03 04 05 06 07 08 09 [✓ 10] 11	1 2 3 4 [✓ 5]	10	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 02 03 04 [✓ 05] 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
30	306.6	700	01 02 03 04 05 06 07 08 09 10 [✓ 11]	[✓ 1] 2 3 4 5	-	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 [✓ 02] 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
29	401.0	414	01 02 03 04 05 06 07 08 09 [✓ 10] 11	1 2 3 4 [✓ 5]	10	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 02 03 04 [✓ 05] 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
27	89.1	700	01 02 03 04 05 06 07 08 09 10 [✓ 11]	[✓ 1] 2 3 4 5	-	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 [✓ 02] 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
25	377.1	414	01 02 03 04 05 06 07 08 09 [✓ 10] 11	1 2 3 4 [✓ 5]	10	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 02 03 04 [✓ 05] 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
8	304.8	414	01 02 03 04 05 06 07 08 09 [✓ 10] 11	1 2 3 4 [✓ 5]	10	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 02 03 04 [✓ 05] 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
49	265.0	700	01 02 03 04 05 06 07 08 09 10 [✓ 11]	[✓ 1] 2 3 4 5	-	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 [✓ 02] 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
50	160.0	614	01 02 03 04 05 06 07 08 09 10 [✓ 11]	[✓ 1] 2 3 4 5	-	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 [✓ 02] 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	[✓ Y] ~~N~~
2	126.0	400	01 02 03 04 05 06 07 08 09 [✓ 10] 11	1 2 3 4 [✓ 5]	10	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 02 03 04 [✓ 05] 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
6	108.9	400	01 02 03 04 05 06 07 08 09 [✓ 10] 11	1 2 3 4 [✓ 5]	10	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 02 03 04 [✓ 05] 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]
5	61.5	700	01 02 03 04 05 06 07 08 09 10 [✓ 11]	[✓ 1] 2 3 4 5	-	[✓ 1] 2 3 4	[✓ 1] 2 3 4	01 [✓ 02] 03 04 05 06 07 08 09 10 11 12 13 14	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	1 2 3 4 5 [✓ 6]	Y [✓ N]

Figure 2-6.– ASCS Ground Truth Periodic Observation form.

ORIGINAL PAGE IS OF POOR QUALITY

The ACC is published biweekly in a meteorological summary for all regions being examined by LACIE. The BMTS stages of wheat are based on inputs from each reporting meteorological station. These estimates are then utilized to develop BMTS contours as shown in figure 2-7. The ITS BMTS estimate is then determined from its location on this contour map and compared to that determined by ground observations. Such a comparison is shown for two ITS's (fig. 2-8). The standard deviation (±1σ) of these ground-observed estimates on a field-to-field basis is also shown in these figures. Note in the Oldham County, Texas, example that the ground-computed stage contains the ACC-estimated stage within one standard deviation in the periods from mid-jointing (3.5) to soft dough (5.0). Before stage 3.5 and after 5.0, the ACC was ahead of the ground truth by a few days and more than one standard deviation. However, in most cases, the ACC BMTS estimate was somewhat more accurate than assuming a normal or average growth stage. In Finney County, Kansas, the historic data operated about as well as the BMTS, and both were relatively close to the ground-observed information.

Tables 2-6, 2-7, and 2-8 display the differences in days at which each of the BMTS stages were estimated by ground observations and the LACIE ACC. At heading, the standard deviation of the ground observations is about 6 to 9 days. A difference between the ground-observed and ACC estimates larger than ±1σ occurred in only three of the U.S. ITS's. While statistical analyses of these data have not been concluded at this writing, it would appear that the computed differences between the ground-observed and ACC-estimated BMTS stages are not significant in terms of the experimental error. However, some trends were noted. In the winter-wheat region, the ACC was consistently ahead of the ground observations at BMTS stages 3.0 (jointing), 5.0 (soft dough), and 6.0 (ripening).

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 2-6.— COMPARISON OF LACIE ACC WITH OBSERVED STAGES IN THE WINTER-WHEAT ITS'S

[Monitoring ACC data (in days) between ITS and ACC development stages]

ITS, county/state	Jointing		Heading		Soft dough	Ripening
	3.0	3.5	4.0	4.5	5.0	6.0
Randall/Tex.	3	7	5	4	8	8
Deaf Smith/Tex.	(a)	(a)	(a)	(a)	(a)	(a)
Oldham/Tex.	-4	17	17	9	9	8
Finney/Kans.	4	5	-3	3	8	-5
Rice/Kans.	-12	0	-5	-14	0	7
Ellis/Kans.	-11	-3	-8	-15	1	-11
Saline/Kans.	4	0	-3	-3	6	11
Morton/Kans.	2	0	1	0	5	8
Shelby/Ind.	10	-1	-3	-1	-4	2
Madison/Ind.	10	6	1	0	8	5
Boone/Ind.	10	9	2	0	2	5
Oneida/Idaho	-11	-7	-7	-7	-5	(a)
Franklin/Idaho	(a)	(a)	(a)	1	4	(a)
Bannock/Idaho	15	3	0	-1	8	(a)
Whitman (1)/Wash.	(a)	(a)	(a)	(a)	(a)	(a)
Whitman (2)/Wash.	-5	10	-3	-9	2	7
Whitman (3)/Wash.	(a)	(a)	(a)	(a)	(a)	(a)
Hill/Mont.	3	-8	-9	-10	5	(a)
Liberty/Mont.	(b)	(b)	(b)	(b)	(b)	(b)
Hand (1)/S. Dak.	17	5	-5	0	(a)	(a)
Hand (2)/S. Dak.	17	(a)	(a)	(a)	(a)	(a)
Toole/Mont.	-4	-8	-6	-9	-8	(a)

[a] No data.

[b] No winter wheat.

TABLE 2-7. - COMPARISON OF LACIE ACC WITH OBSERVED STAGES IN THE SPRING-WHEAT ITS'S

ITS, county/state	Jointing		Heading		Soft dough	Ripening
	3.0	3.5	4.0	4.5	5.0	6.0
Hand (1)/S. Dak.	-10	-5	-2	-8	(a)	(a)
Hand (2)/S. Dak.	(a)	(a)	(a)	(a)	(a)	(a)
Burke/N. Dak.	(a)	(a)	(a)	(a)	(a)	(a)
Divide/N. Dak.	(a)	(a)	(a)	(a)	(a)	(a)
Williams/N. Dak.	(a)	5	2	4	12	(a)
Hill/Mont.	10	12	6	6	15	(a)
Liberty/Mont.	-19	(a)	(a)	(a)	(a)	(a)
Toole/Mont.	2	(a)	-1	6	(a)	(a)
West Polk/Mont.	-7	-5	-2	6	(a)	(a)

TABLE 2-8.— COMPARISON OF LACIE ACC WITH OBSERVED STAGES IN THE CANADIAN ITS'S

ITS, town/province	Jointing		Heading		Soft dough	Ripening
	3.0	3.5	4.0	4.5	5.0	6.0
Ft. Sask./Alta.	-1	0	-7	(a)	(a)	(a)
Olds/Alta.	10	7	(a)	(a)	(a)	(a)
Lethbridge/Alta.	12	13	10	(a)	(a)	(a)
Melfort/Sask.	9	9	7	(a)	(a)	(a)
Delisle/Sask.	11	5	0	(a)	(a)	(a)
Swift Current/Sask.	9	5	-4	(a)	(a)	(a)
Torquay/Sask.	7	3	-2	(a)	(a)	(a)
Stony Mt./Man.	6	3	1	2	(a)	(a)
Starbuck/Man.	4	0	-3	-3	(a)	(a)
Altona/Man.	3	-1	-8	-9	(a)	(a)
Dawson Creek/B.C.	-5	(a)	(a)	(a)	(a)	(a)

[a] No data.

ORIGINAL PAGE IS
OF POOR QUALITY

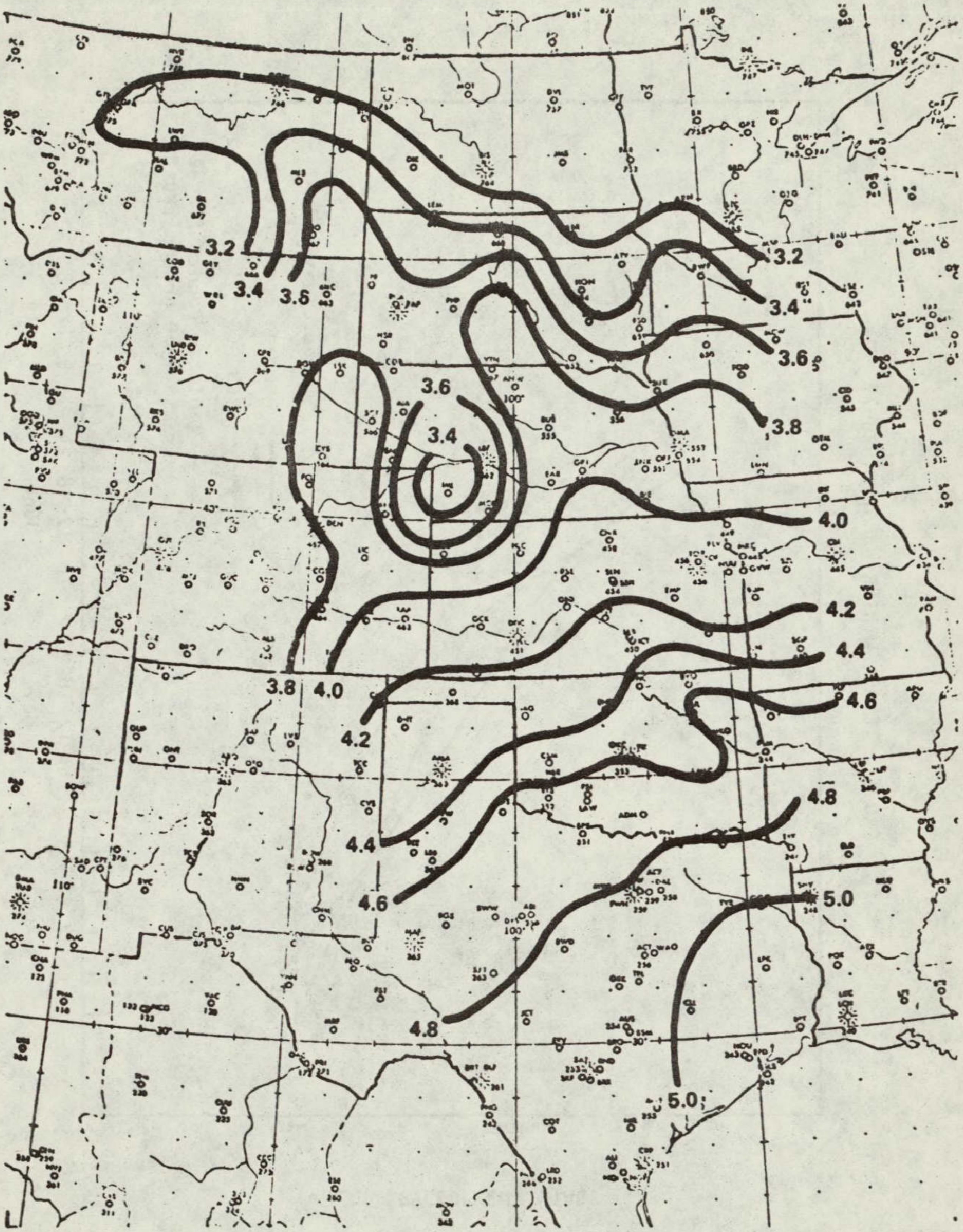


Figure 2-7.- Winter-wheat BMTS isolines as predicted by the LACIE ACC meteorological data through May 1, 1977.

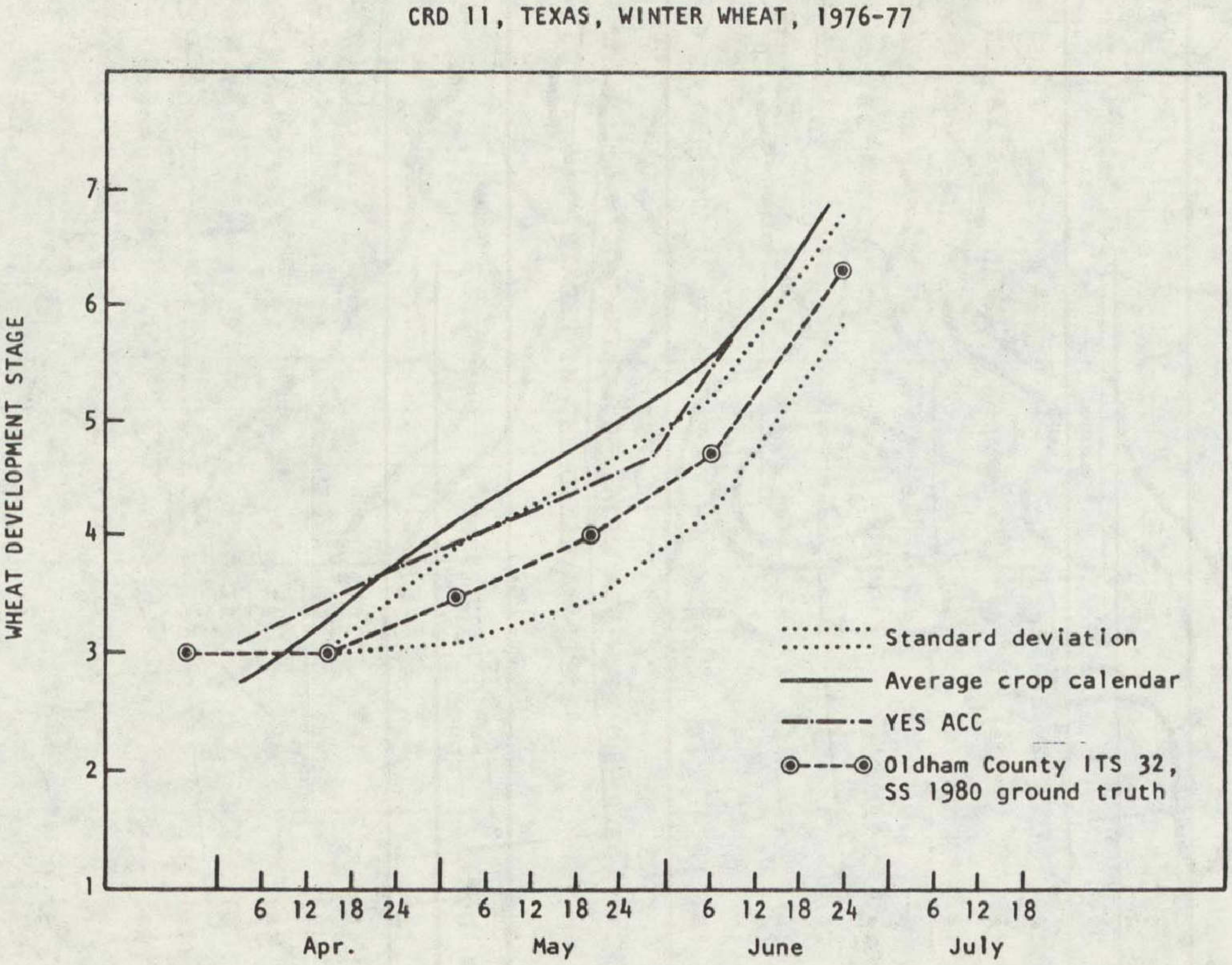


Figure 2-8.– Comparison of observed and predicted crop calendar stages for Oldham County, Texas, and Finney County, Kansas.

CRD 30, KANSAS, WINTER WHEAT, 1976-77

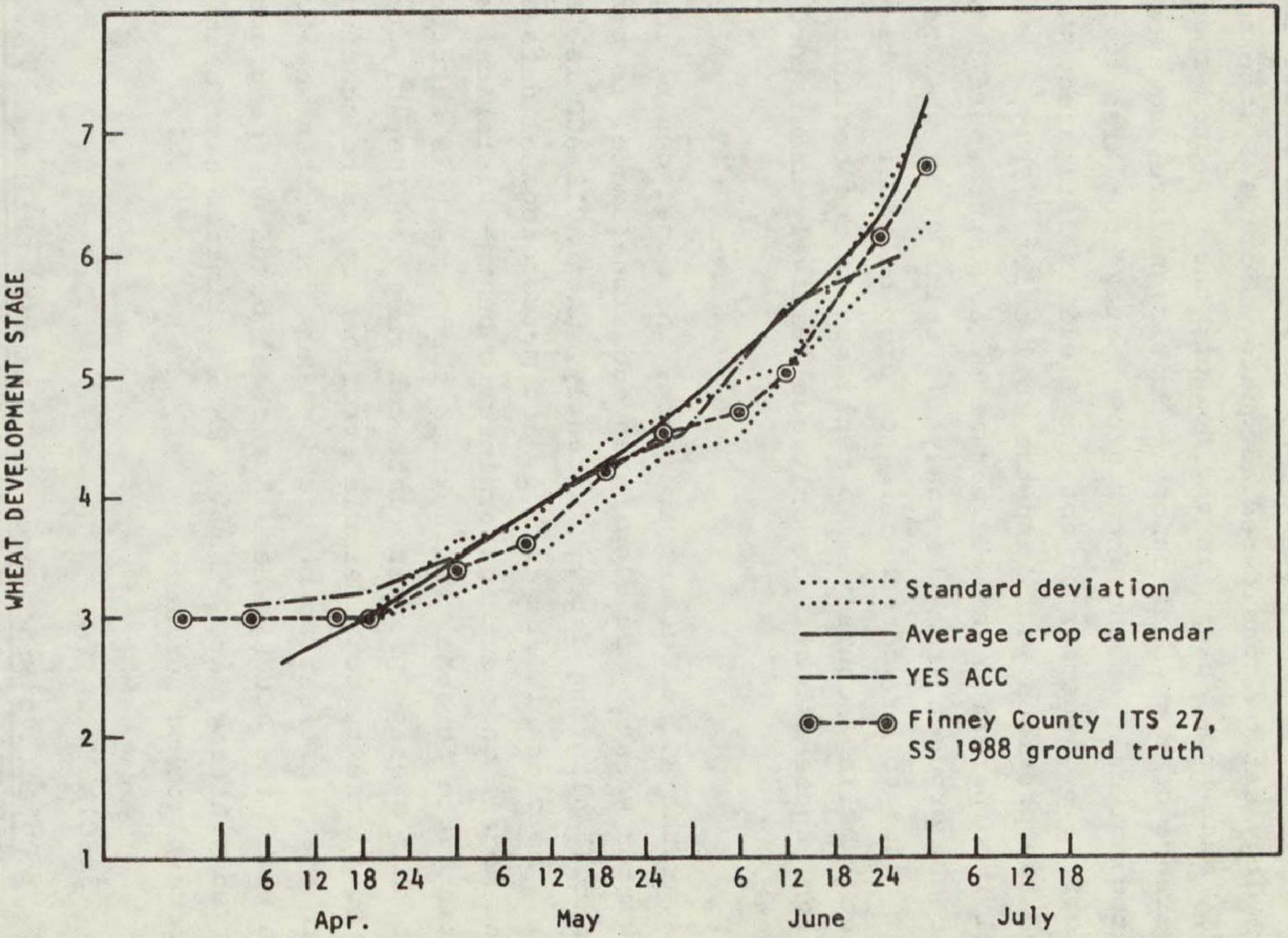


Figure 2-8.– Concluded.

While these results do not conclusively demonstrate crop calendar inadequacies, several issues must be addressed before the ACC technology can be considered adequate. For CAMS, the analyst really must know, early in the season, the expected spectral appearance of the wheat canopy. This signature, however, is related not only to the wheat growth stage but also to other factors; i.e., whether or not the field is irrigated, was fallowed the previous year, and the soil color. Thus, a signature model incorporating the ACC parameter as input would be a more desirable product from the analyst's point of view. Another major issue to be addressed is how crop calendar errors would affect labeling accuracy. As mentioned at the beginning of this section, these effects are only qualitatively understood at present.

Whatever the ACC model requirements, the model can be improved for winter wheat by developing an additional model to predict the actual planting date. Currently, the LACIE ACC is "started" (i.e., the clock is set to 1.0, and meteorological data are fed to the model) on a date determined to be the historical average planting date for the CRD in which the segment is situated. Since this average planting date can vary considerably from one year to the next, considerable error can be introduced into growth stage estimation before dormancy for winter wheat. In tests where the ACC has been "started" based on the ground-observed planting date, the ACC BMTS estimates have been more accurate prior to dormancy.

2.3.7 SPECIAL STUDIES

2.3.7.1 Objective Procedures To Eliminate Estimates From Segments With No Acquisitions Prior to Complete Emergence Thresholding

Investigations of the early-season estimates in Phase II disclosed the presence of an early-season bias or underestimate of harvestable wheat acres. This was caused by wheatfields with insufficient canopy development, which the LACIE procedures could not detect from the Landsat imagery. The LACIE began Phase III Landsat data processing when the normal crop calendar reached stage 2.0 (emergence) on the Robertson growth scale. As the season progressed, ground cover within the fields increased, and the LACIE acreage estimates converged toward the acres harvested. Because of cloud cover, some segments were not acquired after complete emergence. However, wheat estimates based on the early acquisitions for these segments were utilized to make acreage estimates throughout the season. This contributed to the tendency to underestimate wheat acreage at harvest.

In Phase III, an objective thresholding procedure was developed to eliminate from consideration in the overall acreage estimate estimates from segments with incomplete emergence. The thresholding procedure can be applied only at mid-season after several opportunities to acquire and estimate wheat percentages have occurred. Basically, the procedure consists of monitoring the rate of change of segment wheat percentage estimates within each of several segments with multiple estimates. At the average date when the rate of change is small, the crop growth stage of wheat at that date is computed, and all segment wheat percentage estimates based on Landsat acquisitions at dates before the occurrence of that stage are deleted from the acreage estimate. Counties for which all segment estimates are deleted are treated as group III counties in the aggregation. This procedure was tested in Phase III and was demonstrated to decrease the magnitude of the underestimate throughout the season. Therefore, in addition to the normal (nonthresholded) LACIE estimates, the CAS also provided the thresholded estimates in the June and July CMR's (refs. 7, 8). These thresholds were applied to the Landsat data, and no segments acquired before the detection threshold

ORIGINAL PAGE IS
OF POOR QUALITY

were included in the thresholded aggregation. Robertson stage 2.55, as determined from the ACC for crop year 1977, was estimated as the wheat detection threshold of the LACIE for the winter-wheat states.

In table 2-9, the LACIE thresholded and nonthresholded estimates of winter-wheat acreage for the seven states and for the regional levels are compared with the SRS estimates. In June, estimates from all regions and states except Nebraska increased after the thresholding procedure was utilized. Nebraska showed a slight decrease in the acreage estimate. These changes in the estimates brought them closer to agreement with SRS estimates in four of the seven U.S. Great Plains winter-wheat states but increased the relative difference at the seven-state level. This was caused by a sampling problem in the mixed-wheat states and will be discussed momentarily. The CV's were increased only slightly by the procedure except in South Dakota (where the greatest increase in the estimate occurred) and at the seven-state level. The CV of acreage for South Dakota jumped from 22 to 60 percent and for the yardstick estimate went from 4 to 18 percent. This increase occurred because fewer segments were used in August and September than in July because of reallocation based on wheat.

As shown in the July CMR (ref. 8), estimates for the five U.S. Great Plains winter-wheat producing states changed only slightly after thresholding. The CV's for these states remained constant. The small observed differences between the thresholded and nonthresholded estimates resulted from a large number of segment acquisitions after emergence and, therefore, minimal thresholding. Recorded changes were in the forms of mixed increases and decreases among the seven states. The thresholding technique has been approved for operational use in LACIE, and Phase III estimates will be thresholded.

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 2-9.– COMPARISON OF THRESHOLDED WITH NONTHRESHOLDED ACREAGE ESTIMATES IN THE U.S. GREAT PLAINS

Region	Thresholded				Nonthresholded			
	Segments acquired	Total allocation	RD, %	CV, %	Segments acquired	Total allocation	RD, %	CV, %
				June CMR				
Colo.	12	32	32.6	16	28	32	23.0	15.8
Kans.	82	121	-2.2	6	112	121	-9.9	5.8
Nebr.	22	67	15.3	15	50	67	15.5	12.1
Okla.	40	46	-22.5	7	45	46	-33.3	9.0
Tex.	26	38	10.4	14	34	38	2.8	11.9
5-state	182	304	2.6	5	269	304	-4.9	4.2
Mont.	3	80	33.2	25	41	80	13.9	19.2
S. Dak.	5	56	95.1	60	28	56	73.9	34.0
2-state[a]	8	136	80.6	47	69	136	40.6	18.5
7-state	190	440	32.4	18	338	440	3.2	4.8
				July CMR				
Colo.	25	32	15.1	15.5	30	32	20.3	13.2
Kans.	98	121	1.8	4.8	111	121	-4.6	5.0
Nebr.	34	67	18.6	11.6	52	67	12.2	12.4
Okla.	37	46	-15.5	7.5	42	46	-23.5	8.5
Tex.	29	38	.5	12.8	34	38	-2.0	11.6
5-state	223	304	1.7	3.8	269	304	-3.0	3.9
Mont.	44	80	-6.5	11.9	58	80	9.6	12.3
S. Dak.	32	56	88.0	13.8	39	56	85.3	12.6
2-state	76	136	58.1	10.1	97	136	55.0	9.0
7-state	229	440	14.1	3.7	366	440	9.6	3.6

[a] Mixed wheat.

While the thresholding of acreage also improved the production estimates as reported in the June CMR (see table 2-10), a problem in South Dakota caused the thresholding procedure to degrade the estimate reported in July. An investigation into this problem showed that the South Dakota overestimate for winter wheat was caused by CAMS overestimates of small grains in the mixed-wheat areas. In this strategy, each segment that was allocated on the basis of total small grains was estimated separately for both winter- and spring-wheat acreage. Many South Dakota segments had enough spring wheat to have a segment allocated, but some of these segments contained almost no winter wheat.

The CAMS overestimated the winter-wheat acreage in these segments, confusing pasture with winter wheat. Although these errors were reasonably small in an absolute sense (1 to 2 percent), the relative overestimate in these low-acreage segments greatly inflated the South Dakota winter-wheat estimate.

For the August report, the procedure was corrected to allocate segments in mixed areas: Those areas with little or no winter wheat were not analyzed by CAMS for winter wheat, and the corresponding counties were treated as group III counties in the aggregation. This procedure greatly improved both the South Dakota and the U.S. Great Plains acreage estimates for August.

A decision was made to utilize thresholding methodology for the U.S.S.R. keyed to the wheat-tillering growth stage. The validity of this approach can be determined, at least in part, by comparing the April through July acreage estimates for Phase II with Phase III. The early-season estimate in Phase II (where the thresholding technique was not used) reflected an extremely low acreage estimate, which was increased by 45 percent for July; whereas the acreage increase in Phase III during the same time frame amounted to less than 10 percent. The more stable

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 2-10.– COMPARISON OF THRESHOLDED WITH NONTHRESHOLDED PRODUCTION ESTIMATES IN THE U.S. GREAT PLAINS

Region	Thresholded				Nonthresholded			
	Segments acquired	Total allocation	RD, %	CV, %	Segments acquired	Total allocation	RD, %	CV, %
June CMR								
Colo.	12	32	31.6	21.8	28	32	21.8	21.9
Kans.	82	121	-19.4	11.6	112	121	-28.4	11.5
Nebr.	22	67	3.0	18.7	50	67	1.9	16.2
Okla.	40	46	-61.0	13.1	45	46	-75.0	14.0
Tex.	26	38	-10.7	15.9	34	38	-19.6	14.2
5-state	182	304	-15.0	7.1	269	304	-23.6	6.9
Mont.	3	80	35.8	28.1	41	80	17.3	23.2
S. Dak.	5	56	96.2	62.2	28	56	79.9	38.3
2-state	8	136	81.2	47.3	69	136	43.9	21.1
7-state	190	440	23.0	19.0	388	440	-10.8	7.0
July CMR								
Colo.	25	32	13.1	21.3	30	32	18.4	19.7
Kans.	98	121	-5.5	10.8	111	121	-12.4	10.9
Nebr.	34	67	11.3	15.0	52	67	4.6	15.7
Okla.	37	46	-50.8	13.1	42	46	-61.1	13.6
Tex.	29	38	-22.6	14.8	34	38	-25.4	13.9
5-state	223	304	-10.2	(a)	269	304	-15.7	(a)
Mont.	44	80	-8.7	16.9	58	80	7.8	17.2
S. Dak.	32	56	89.2	23.2	39	56	86.8	22.6
2-state	76	136	58.3	(a)	97	136	55.2	(a)
7-state	299	440	5.4	6.6	366	440	.1	6.4

[a] No data.

acreage estimates through the season suggest that the U.S.S.R. thresholding procedure is valid, given that the at-harvest acreage estimates are the most accurate seasonal estimates. Furthermore, the thresholding dates established for the United States corresponded well with crop stages which are considered as approximately tillering.

2.3.7.2 Second-Generation Sampling Strategy and Yield Model Evaluation

The second-generation sampling strategy and the second-generation yield models were implemented for Phase III in an offline mode for two U.S. Great Plains states (Kansas and North Dakota) and two oblasts in the U.S.S.R. spring-wheat indicator region (Kurgan and Tselinograd). The yield model was implemented for a U.S.S.R. winter-wheat oblast, Khmel-Nitsky; and the sampling strategy included a third spring-wheat oblast, Kustanay. The first-generation sampling strategy and the first-generation yield models were retained in an operational mode over these areas for the purpose of comparing the hectarage (acreage), yield, and production estimates obtained from the two technologies and to evaluate operational designs and impacts.

The second-generation sampling strategy design for the United States was developed using procedures and data input requirements similar to those in the U.S.S.R. so that the performance parameters obtained from the U.S. evaluation would be as applicable as possible to the U.S.S.R. region. In summary, the Phase III scope of the second-generation strategy is to test the sampling scheme and procedures for aggregating estimates of wheat hectarage (acreage), yield, and production in LACIE foreign areas using Kansas and North Dakota as quantifiers from the yardstick region. The testing is in the initial stages at this time with initial Kansas aggregations being the only ones completed.

Several tests to evaluate the effectiveness and efficiency of the second-generation strategy are being carried out during Phase III. These include tests of the degree of homogeneity of yield and agricultural density achieved by restratification and comparisons of various aggregations.

Comparisons are being made of one aggregation with another using the first-generation strategy and each corresponding aggregation and using the second-generation strategy, including comparisons on all statistics. The following inputs to the formulas for second-generation strategy aggregation were used.

- CAMS estimates from second-generation strategy segments only.
- CAMS estimates from first-generation strategy segments only.
- CAMS estimates from a statistically feasible mixture of first- and second-generation strategy segments (i.e., choices of the first-generation strategy segments and certain subsets of the second-generation strategy segments which result in a sample statistically equivalent to the second-generation strategy within each stratum). This mixture of the first- and second-generation strategy segments permits utilization of the collected history available on first-generation segments.

The above inputs are made in combination with the following:

- Use of the Feyerherm yield model, which is applied at the natural stratum level (this refers to the stratum resulting from restratification in support of the second-generation strategy).
- Use of the CCEA yield model, which is applied at the political subdivision (state or oblast) level.

Two aggregations have been completed over Kansas at this time. Aggregations were made on first-generation segments acquired and analyzed as of June 7 and July 11, 1977. Second-generation segments were aggregated on June 20 and July 11.

Based on the limited comparisons carried out in Kansas, preliminary indications are that the second-generation sampling strategy is significantly more efficient than the first-generation strategy. In Kansas, the second-generation strategy gave a wheat production and acreage estimate with about the same CV as the first-generation strategy; however, 81 segments were required for the second-generation, as compared to 121 for the first-generation, strategy and only 84 in the 1975 allocations based on wheat. It was also proven that, in South Dakota (section 2.3.7.1), first-generation sampling based on wheat is more efficient than first-generation sampling based on small grains.

2.3.7.3 Evaluation of Wheat From Small-Grains Procedures

A major technical issue within the LACIE has been the inability to reliably differentiate wheat from small grains directly from the Landsat data. Specifically, in Phase I, analyses of 20 North Dakota blind sites revealed that spring barley, a crop very similar in appearance and growth cycle to spring wheat, was not being distinguished accurately from spring wheat. In some segments, spectral separation did exist. This separation was not observed in enough segments to permit sufficiently accurate overall analysis. Efforts were begun late in Phase I to develop improved analysis procedures which could take advantage of the spectral separability between these crops. For Phase II, however, the classification and mensuration procedures were used to estimate total small grains, and ratios based on the historic proportions of spring wheat to other small grains were used to convert Landsat-based estimates of small grains to spring-wheat estimates.

In Phase II, given the Landsat-based estimates of total small grains, the ratios from the latest year for which data were available were used to estimate spring wheat. In most cases, the current-year prevalence of wheat had increased considerably over the historic value. In Canada, where the latest available crop district data were for 1971, the ratios had increased by as much as 50 percent. In the United States, the increase over 1975 averaged approximately 10 percent. Thus, the use of the historic ratios in Phase II contributed to an underestimate of about 10 percent in the four yardstick spring-wheat states and by larger percentages in Canada.

For Phase III, priority was assigned to technological improvements for identifying spring wheat directly from the Landsat data. Procedures utilizing improved analyst aids, such as interpretation keys and displays of quantitative spectral data, were developed. In addition, econometric models for the prediction of wheat to small-grains ratios were developed, tested, and utilized. These models predict the current ratios of wheat to small grains resulting from influential factors such as historical crop and livestock patterns, current-year growing conditions (such as available soil moisture), economic conditions, and prevailing government farm programs.

Utilizing blind-site ground truth, investigations were made into the spectral and temporal differences between spring wheat and spring barley. These studies revealed that the following characteristic differences might provide sufficiently different spectral/temporal patterns to permit reliable differentiation between these two crops.

- On the average, barley is planted after wheat.
- Barley "greens up" more and develops faster than wheat.

- Barley ripens and is harvested earlier than wheat.
- Barley is more reflective than wheat.

In addition, it was noted that rye is greener than wheat and that oats are not as green as wheat and may mature earlier.

Analyst procedures were developed for using the quantitative spectral aids developed in Procedure 1 to identify barley fields and wheatfields based on these general differences. The analyst was required first to execute the standard CAMS procedures for obtaining an acceptable segment total small-grains proportion estimate and then to examine the computer-classified labels for the preselected 209 dots and to label the small-grains dots as to wheat, barley, and other small grains. The 1975 general production statistics (ranges) were furnished the analyst so he could obtain a crude estimate of the ratio of wheat to small grains. He then studied the spectral crop calendar to ascertain the expected spectral characteristics of each small grain. Following this, he observed the Procedure 1 spectral plots, which are a temporal sequence of two-dimensional plots of the dot radiance values transformed into the Kauth-Thomas (ref. 3, p. 52) coordinate representation (greenness-brightness axes), to determine if the small-grains dot values tended to cluster in groups. The ancillary data on relative abundances gave the analyst a firsthand impression as to the relative size of the spectral groupings; however, the separation was based on natural breaks in the computer-classified small-grains data. Then, based on the above procedure, the analyst labeled the groups as wheat, spring barley, or other crops. The fraction of the small-grains dots labeled wheat was then multiplied by the Procedure 1 bias-corrected small-grains estimate to obtain one wheat estimate for the segment.

This procedure is being evaluated over North Dakota in Phase III. Each segment was processed for both small grains and spring wheat.

The small-grains estimates were ratioed utilizing spring-wheat to small-grains ratios predicted by the econometric models developed for Phase III, and both wheat estimates for each segment were aggregated to determine the North Dakota wheat acreage estimate. The ratioed results were utilized in the operational reporting and were compared also to the results obtained by the CAMS direct estimation procedure. Preliminary results indicate that direct estimates of wheat were larger than ratioed estimates and that, on a segment-by-segment basis, there is a correlation of 0.89 between the estimates. However, this evidence does not infer that the direct estimates are superior, inasmuch as this spring-wheat acreage estimate for North Dakota is an early-season estimate with significant early-season bias. Final evaluation must await the completion of the spring-wheat season and a comparison to the blind-site ground observations.

2.4 SYSTEMS PERFORMANCE

2.4.1 DATA RATES

Phase III of the LACIE required another significant expansion in increased throughput rates over Phases I and II for processing wheat segment data acquired by Landsat through the quasi-operational element of the LACIE. In order to handle the peak rates projected for the May through September time frame, more efficient procedures were required for the segment data analyses. In addition to accuracy requirements, this was an integral part of the Procedure 1 design rationale. For the first time, complete multitemporal machine processing was routine for all segment wheat estimates.

By August 1, the number of acquisitions and analyses had approximately doubled that of all of Phase II; the analyst time had been reduced from 6 to 4 hours per segment; and segments were analyzed at the rate of approximately 55 per day (table 2-11).

TABLE 2-11.- COMPARISON OF PHASE I, II, AND III DATA RATES

Phase	Number of segments	Number of acquisitions			Analysis time per segment, hr.
		By Landsat (a)	Available for analysis	Analyzed	
I	692	7 500	2 649	1 627	12
II	1 683	27 000	9 148	9 148	6
[b]III	2 900	50 000	14 600	14 600	4

[a] Goddard Space Flight Center (GSFC) processing.

[b] Approximately to August 1.

2.4.2 LANDSAT DATA ACQUISITION

The acquisition of Landsat data has proceeded for the most part as anticipated. Retro-orders were required based on relocated and additional samples in the United States and the U.S.S.R. Real-time data were backlogged while the retro-order was acquired at capacity rates. As a result, the GSFC has been operating at or near capacity for all of Phase III. Data rejections for Phase III were about the same as for Phase II. The rejection rates were slightly larger than 50 percent because of cloud cover and 15 percent caused by correlation and other technical difficulties.

Originally, it has been planned to acquire most of the foreign data through the Pakistan and Italian ground stations in order to conserve the onboard tape recorder. Because of problems with the tape records from these stations in late spring and some data loss over the U.S.S.R., it was decided that all U.S.S.R. data would be acquired using the onboard recorder. The data normally acquired by the Italian station were recorded on board for only 1 week. At that time, the problem was isolated and fixed, and the ground station mode of data collection was resumed. The onboard tape recorder was used to supplant the Pakistan ground station for the remainder of Phase III. In addition, full-frame data were acquired over China through use of the onboard recorder in order to build a historical data base, even though China was not a formal part of Phase III operations.

The use of the onboard tape recorder decreased the transfer time from acquisition to receipt of the data by JSC. However, by late June when the maximum Phase III data loads began to peak, the GSFC processing system became saturated with data, and backlogs began to increase significantly. By mid-July, even though GSFC was operating at a greater than projected capacity, typically it took 3 weeks from acquisition of a segment until it was received at JSC.

2.4.3 INITIAL PHASE III ACREAGE ESTIMATION PROCEDURES

The procedures employed early in Phase III, which were essentially those utilized in Phase II, required 6 hours of analyst time for each processed segment. The period of processing from December 1976 through early February 1977 involved the delayed early-season processing of the Phase II sample allocations. The Phase III data acquisition windows were opened 45 days earlier than those of Phase II to obtain acquisitions during seedbed preparation. These preemergence acquisitions were accumulated, along with all subsequent acquisitions, through the end of December. All segments available were then analyzed once to support the February 8 CMR. Consequently, the average turnaround time from Landsat acquisition to CAS processing was 64 days. However, the analysis in-house time was only 18 days, which would have supported the Phase III turnaround goal of 30 days. In fact, those acquisitions of Landsat data at the end of December were turned around in 30 days.

The period from February through April 1977 was impacted by the retro-order of Landsat data caused by the relocation and reallocation of segments; as a result, the turnaround time from Landsat to CAS had no meaning as an indicator of operational performance.

2.4.4 SMALL FIELDS PROCEDURE

In late March 1977, the Small Fields Procedure was implemented. Although developed primarily as an improvement in the processing of small field areas, it was an initial step toward implementation of the full Procedure 1. With the Small Fields Procedure, clustering to support multitemporal machine processing and machine processing of segments with small proportions (less than 5 percent) of small grains were provided for the first time.

Some problems were encountered during the 2 months (April and May) following transition from the Phase II procedure to the Small Fields Procedure. Most of these were attributed to the newness of the basic technological concept and the incompleteness of the detailed design implementation. The most severe problem was that an overly stringent evaluation criterion was used and initially required most segments to be reworked. This resulted in severe congestion at the end point of the processing cycle. When the problem associated with the evaluation criterion was solved in early May with an improved evaluation procedure, adequate numbers of segment estimates were available to support the May CMR (ref. 6). In spite of these problems, the per-segment analyst throughput met and, in some cases, exceeded that of the previous procedures. In addition, there were indications of improvement in the quality of the estimates, and the data processing was current to approximately 30 days. However, by June, the increasing data loads and backlog at GSFC had increased the turnaround time from Landsat acquisition to CAS to approximately 45 days.

2.4.5 PROCEDURE 1 — MAINLINE DATA SYSTEM

Beginning with the processing of spring-wheat data on June 6, 1977, a majority of the Procedure 1 software had been delivered, and segment processing utilized the Procedure 1 concept. Because of their basic similarities, the transition from the Small Fields Procedure to Procedure 1 was executed fairly smoothly; however, the expected improved throughput did not occur immediately, and backlogs began to develop. An operations analysis was performed to isolate the problem areas. The key findings were that the analyst contact had been reduced significantly (from 6 to 4 hours); that the support functions were not adequately prepared to handle the many new products associated with Procedure 1; and that the quality control group could not handle the throughput. In order to handle the data load to support the July reports, adjustments of resources and extensive overtime were required for the first 4 to 6 weeks. After mid-July, procedures and software were developed to support Procedure 1 operations, restrictions to the data flow began to lessen, and backlogs receded.

Even with the data handling problems, the throughput rate for the first full month of Procedure 1 operations (July) averaged 56 segments per day, 45 of which were considered suitable for aggregation. The unsuitable segments were caused by a variety of causes, including preemergence of spring wheat, consecutive-day acquisitions, cloud cover, and misregistration. Another very important operational aspect of Procedure 1 is that less than 1 percent of interactive rework was required, resulting in complete elimination of the need to maintain a special rework team and allowing computer time to be utilized for the more efficient batch operations.

2.4.6 PROCEDURE 1 — HYBRID SYSTEM

The implementation of Procedure 1 on the integrated CAMS IMAGE 100 Hybrid System was completed on May 31. Analyst "hands on" training began on June 6 and was completed on June 17 in parallel with a verification test which was designed to fully exercise the new system and analysts. It provided an operational shakedown of the system and data flow interfaces. On June 20, operational data processing began with four USDA analysts processing a U.S.S.R. spring-wheat oblast (to support the U.S.S.R. reports) and the Canadian blind sites and ITS's for Accuracy Assessment evaluations.

As anticipated, several minor software, hardware, and procedural problems were encountered in the early days of operational use, causing some slowdown in throughput. However, the analyses required to support the CMR's were accomplished on schedule, and most of the early goals were met. The operational throughput for the first 30 days averaged 3.3 segments per 12-hour shift, slightly better than the forecasted 3 per shift. The U.S.S.R. segment average turnaround time for analysis was 7 to 8 working days, slightly longer than the target of 6 working days but a factor of 2 better than the mainline operations. Analyst contact time for the initial analyses averaged 4.5 hours per segment, similar to the mainline operations.

2.4.7 METEOROLOGICAL DATA PROCESSING

In Phase III, the LACIE/ERIPS again ran CCEA yield models using meteorological data gathered from the NWS, the ETAC, and the WMO network. Yield models for the yardstick area were revised to eliminate data overlap areas, and additional models were developed for nine regions in the U.S.S.R. Output from these revised and new models began in April 1977, even though yield estimates for the yardstick area and the U.S.S.R. began in November 1976. Yield models for 16 states in India were developed, and estimates for all states were prepared in December 1976. Estimates for one Indian state (Madhya Pradesh) were continued for the remainder of the season.

Feyerherm yield models were run operationally for one state (Kansas) and one oblast in the U.S.S.R. from April 1977 through the crop season. Daily meteorological data, along with crop calendars for Canada, the U.S. Great Plains, and the U.S.S.R., were obtained to run these models. Maximum and minimum temperatures for model operations in the U.S.S.R. are actually the highest and lowest observed temperatures from the three hourly synoptic reports and are used operationally to approximate the true maximum and minimum temperatures.

ORIGINAL PAGE IS
OF POOR QUALITY

The winter-wheat crop calendars were run throughout the winter season, and a programming error was detected. Crop calendars were not run from mid-January through mid-March because of this problem. Wheat was dormant in most areas for this entire period. A procedure was developed in October 1976 to provide CAMS analyst feedback to the YES whereby the crop calendar could be adjusted from analyst input. This procedure was used to correct the U.S.S.R. spring-wheat crop calendar.

A weekly meteorological summary was published for use by the CAMS analysts. These summaries were prepared by LACIE personnel from data furnished by the NWS, the ETAC, the CCEA, and foreign newspaper reports.

2.4.8 RESULTS REPORTING

The first U.S. Great Plains winter-wheat CMR was produced on February 8, 1977, as scheduled, based on the Phase II allocation of 431 segments. An additional 170 segments, allocated to attain a sample density for the yardstick area to support the 90/90 criterion, plus 40 relocated nonagricultural segments were retro-ordered on January 31, 1977. These data were received at JSC on March 3 and were processed to support an April 6 release (ref. 10) of the early-season February report (ref. 2) for the 601-segment allocation. A CAS unscheduled report was released April 22 (ref. 11). It included some of the data acquired in January, February, and March, which were held in abeyance until the retro-order was completed. However, because of slow throughput time encountered with the implementation of the Small Fields Procedure in March, the April 22 report did not completely represent the nominal April 6 report as was originally intended. Winter-wheat reports were released also on May 9, June 7, and July 11 as scheduled. Although all available spring-wheat data were processed in time for the July CMR (ref. 8), it was deemed that insufficient samples were available to support an estimate.

The first U.S.S.R. report was originally planned for January 21 but was released on March 30. This delay was caused primarily by the retro-order and secondarily by an adjustment in schedules to have the releases nearer the scheduled USDA/FAS task force meetings early in the month. Subsequent reports were made for U.S.S.R. winter wheat on May 2 and June 3 as scheduled, but little additional data were available after dormancy because many segments in the U.S.S.R. were closed for dormancy through the winter. The July 1 report included only winter wheat and contained significant amounts of data after winter dormancy. However, sufficient data after spring-wheat emergence were not available to support a spring-wheat estimate.

Data processing and reporting on the State of Madhya Pradesh in India were suspended indefinitely because of the impact caused by the retro-orders.

2.4.9 EARLY-SEASON THRESHOLDING

LACIE estimates made early in the season tend to be biased low, which is caused by an inability to detect wheatfields with insufficient canopy development from Landsat data. As the season progresses, ground cover within the fields increases and the LACIE estimates converge toward the acres harvested. In order to reduce the bias introduced by these early-season estimates, thresholding procedures were employed which delete these early-season estimates from aggregation.

Because the average tillering date for each oblast (stratum) was readily available through newspaper reports for the U.S.S.R. and because wheat in the tillering stage appeared to be detectable using spectral data, it was decided to use these dates as a criterion for the thresholding effort in the U.S.S.R. The estimates for segments acquired prior to these tillering dates then

ORIGINAL PAGE IS
OF POOR QUALITY

were removed from the CAS aggregation that supported the March 30 and subsequent reports for U.S.S.R. winter wheat.

Reports of tillering dates were not available for the yardstick area; however, an objective method of establishing early-season LACIE detection thresholds was developed and applied to the June and July yardstick winter-wheat estimates, which were carried as estimates in the appendixes of those reports. The threshold method applied to the Landsat data for these estimates was derived by examining the CAMS estimates for segments classified more than once and is oriented toward determining the growth stage/calendar dates at which the estimates stabilize. This method is more fully discussed in section 2.3.7.1.

The thresholding procedure for the U.S. Great Plains has been applied to data at the state level and varies slightly from one report to another based on the available data. Multiple estimates were not available for South Dakota and Montana winter wheat for these reports. The threshold stage for those states was established at 2.55 on the Robertson BMTS, which was consistent with the stages in the other five states which had multiple estimates. This stage effectively eliminated South Dakota and Montana estimates for the June and July reports because it thresholded nearly all the segment estimates which were based on data in the fall prior to the long winter dormancy period. In prior LACIE phases, no attempt had been made to make estimates for those states in this time period.

2.4.10 OPERATIONAL SYSTEMS PROJECTIONS

a. Turnaround Time: Because of the various technology modifications, the average turnaround time observed in Phase III cannot be used to project turnaround time for an operational system. However, a relatively few cases were observed where backlogs and other system or resource problems did not exist

ORIGINAL PAGE IS
OF POOR QUALITY

and segments were processed through the system from Landsat acquisition to analysis completion in 16 to 18 days. This indicates that a 14-day turnaround is achievable. The U.S.S.R. segment analysis time on the CAMS IMAGE 100 Hybrid System of 7 to 8 days would also seem to support this contention. Furthermore, even in the presence of the problems discussed, the reports issued to date in Phase III generally included most of the high-priority data acquired from 30 to 40 days before the deadlines established for data input to the reports.

b. Analyst Contact Time: Analyst contact time for Procedure 1 averaged 4 hours per segment in Phase III. Because of limited experience to date, it is probably premature to project this to an operational system. However, methods of reducing the time by eliminating some of the mechanical steps now involved are now being considered.

c. Reporting Dates: An analysis was performed on the data available for each report issued through July for the purpose of establishing when each report could have been issued based on a 14-day turnaround from Landsat acquisition (table 2-1). This assumes that the LACIE yield estimates are not significantly different over a 30-day period, which generally is true for the U.S. Great Plains; however, differences of 4 to 5 percent in production for a crop type at the country level have been observed as a result of the monthly changes in U.S.S.R. yield estimates.

2.5 RESEARCH, TEST, AND EVALUATION

The function of the LACIE RT&E effort is twofold: The mainline operation identifies key technology problems, defines their nature and magnitude, and prioritizes their relative importance; and the research, which is keyed to the prioritized problem list, develops alternative approaches. Test and evaluation is an offline element to test and analyze alternative approaches to the mainline wheat survey technology.

The following key technology issues were defined through two phases of LACIE operations and identified in the LACIE Phase II Evaluation Report (ref. 5):

a. Inability to reliably differentiate wheat from small grains directly from Landsat data

b. Subsequent need for econometric models to predict the ratios of wheat to small grains

c. Observed classification underestimates of small grains

d. Improved yield models

e. Improved sample design

f. Need for signature extension technology

Overall, the RT&E program through Phase III has been responsive to these efforts; and the responsiveness is gradually increasing as the university and industrial communities are alining their organizations to pursue focused, large-scale research efforts as opposed to small, fragmented, individual efforts. In addition, the 2-year operation of the LACIE survey system has defined more clearly the nature of agricultural remote sensing problems so that such efforts are possible.

2.5.1 TEST AND EVALUATION

Two major test and evaluation efforts were conducted during Phase III: the test and evaluation of Procedure 1 (see section 2.4.3) and the test and evaluation of the modified first-generation yield models. The Phase III test and evaluation of Procedure 1 accomplished the following two major tasks.

a. Studies over limited numbers of data sets were conducted to determine a workable set of parameters for Procedure 1, such as the required number of dots to be labeled by the analyst and certain clustering parameters.

b. Procedure 1 performance was evaluated.

In this latter effort, Procedure 1 was tested over several blind sites scattered throughout the yardstick region. These tests, which were based on ground-truth labeled dots, showed (1) that Procedure 1 produces accurate and unbiased estimates of wheat proportions and (2) that the machine classification part of the procedure compared favorably and, in fact, did better on the average when compared with the field-trained classification methods used in Phases I and II. These classifier comparison results are tabulated in table 2-12.

From the analyst/interpreter's point of view, Procedure 1 has proven to be an efficient method. The transition from a field-trained to a dot-trained classification procedure has proven to be no more vulnerable to pixel misregistration problems than the field-trained classifier. However, because Procedure 1 has permitted multitemporal classification, registration is seen to be a problem when more than two passes are analyzed, particularly in small field areas such as the strip-fallow fields in the U.S. northern Great Plains.

Another task within the Phase III test and evaluation effort was the assessment, using a test data set, of modified first-generation yield models for Phase III. Preliminary evaluation of the 10-year test for the U.S. Great Plains model has been completed. Model boundary revisions since Phase II (removal of historical yield and weather overlap between pseudozones) removed the biases in North Dakota, Nebraska, and Oklahoma. A bias was

TABLE 2-12.– COMPARISON OF CLASSIFIER RELATIVE BIAS AND COEFFICIENT OF VARIATION

[Analysis based on average performance across four LACIE segments in Kansas]

Pass combinations	Field-trained classifier		Dot-trained classifier (a)	
	Relative bias	CV, %	Relative bias	CV, %
1	36	92	15	12
1, 2	26	62	-1	12
1, 2, 3	16	54	-2	17
1, 2, 3	20	57	-.4	17

[a] Table values are for classifier results without bias correction.

observed in only one model — the Texas West Edwards Plateau model. For this model, an average difference between LACIE and SRS yield prediction of -2.4 bushels per acre was observed.

Over the 10 years used in testing, the mean-square error between the predicted and the SRS yields increased slightly for the modified Phase II yield models when compared to results for the yield models utilized in Phase II. Some of the reasons encountered for the increased variance are:

a. Changes in the way trends were used

b. Changes in weather censoring

c. Differences in the meteorological data base

The hypothesis that the 10 years of simulated yield predictions meet the LACIE 90/90 criterion was sign tested on the observed errors (predicted yield minus SRS yield) relative to the tolerable errors. The decision to accept or reject the hypothesis was based on a binomially distributed test statistic. The hypothesis was accepted at the 0.07 level of significance, which required that eight or more cases fall within the tolerance limits. These results are shown in figure 2-9.

2.5.2 RESEARCH

To date, the research program through Phase III has accomplished the following major tasks.

- The development of an automated, statistically based, multitemporal classification procedure designed to be trained with or without ground observations (Procedure 1, which is discussed in detail in section 2.4)
- The development and initial testing of a multitemporal signature extension procedure (Procedure B, which is an extension of the Procedure 1 concept)

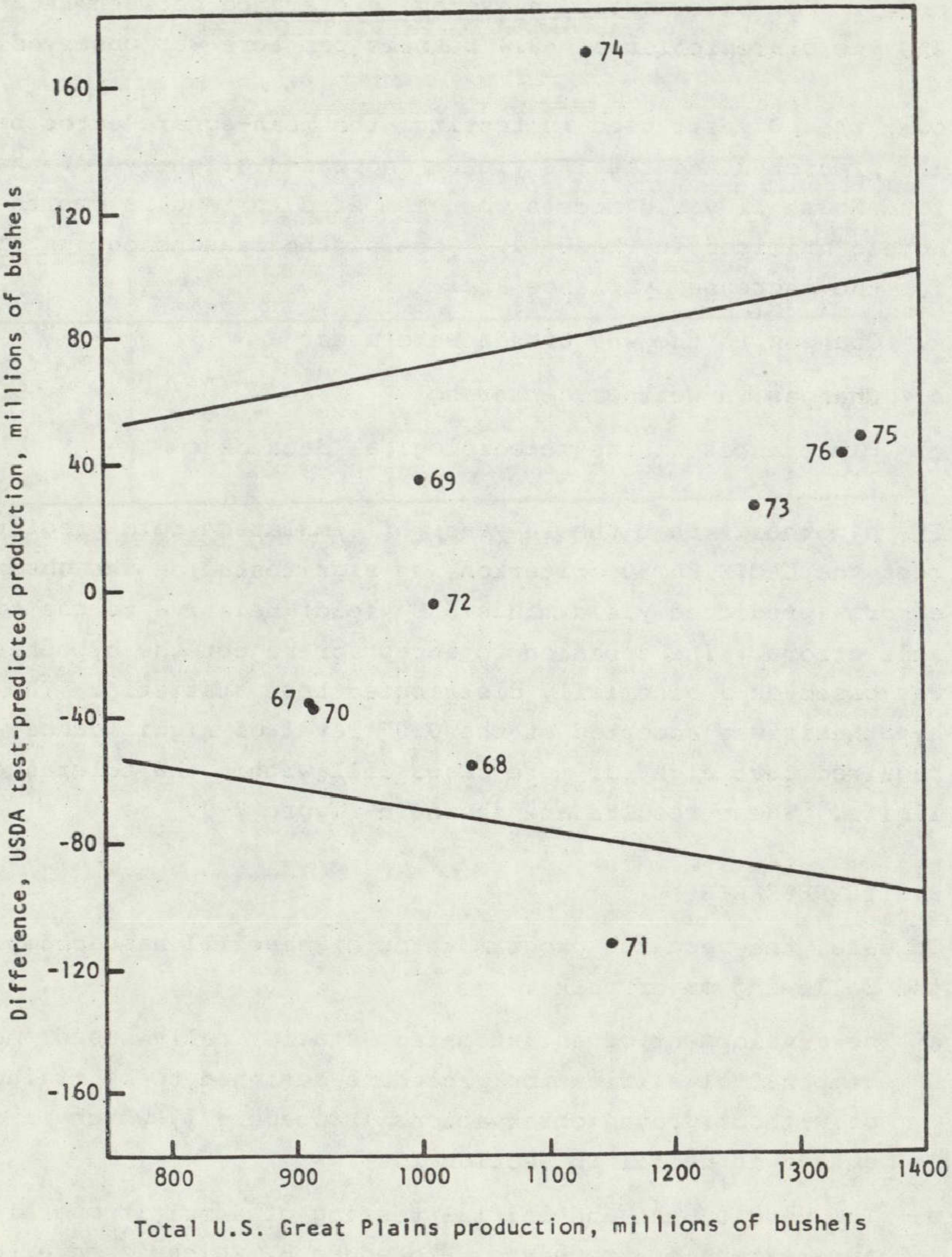


Figure 2-9.— Distribution of U.S. Great Plains yield-related production errors with respect to LACIE tolerance bounds. Errors are based on Phase III test yields and SRS acreages, and tolerance limits assume permissible error is equally divided between yield and acreage.

- The application of quantitative displays and investigations to apply statistical pattern recognition to image interpretation.
- The development of a globally applicable, efficient sampling strategy
- The development of a yield model with potential global applicability
- The construction of a data base from an ongoing field measurements program

2.5.2.1 Improved Machine Processing Procedures

The LACIE experience in the analysis of Landsat data has vastly improved the technology for the automatic machine processing of complex data structures inherent in the multitemporal acquisition of multispectral data.

The evaluation of the improved technology has resulted in the development of a nearly optimum automatic processing procedure which will be implemented by mid-Phase III of LACIE. The procedure can be described as nearly optimum in the sense that (1) the need for manual intervention is almost eliminated from the machine processing sequence; (2) every measurement in the scene, as well as the full dimensionality of the spectral data, is utilized in statistical computations prior to maximum likelihood classification; and (3) with correct analyst determinations of crop identity for a very small sample of the segment, the machine processing procedure will provide an unbiased estimate of segment crop proportions.

This Phase III procedure has automated many of the manual functions performed previously and incorporates many new features. Specifically, the important features are as follows.

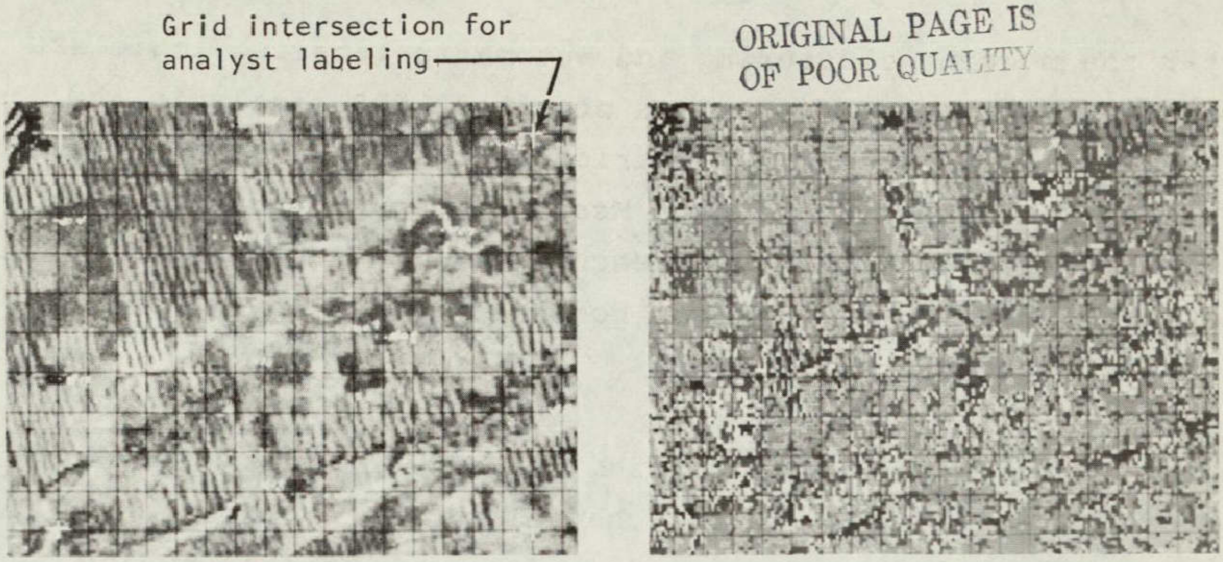
a. As shown in figure 2-10(a), pixels (grid intersections) are randomly selected within the segment and presented to the analyst for labeling as wheat or nonwheat using image interpretation techniques. The analyst submits these labels to the machine, which, without further intervention by the analyst, executes the remaining functions.

b. Machine clustering is performed to delineate the spectrally homogeneous modes within the multispectral/multitemporal segment data, and a color map is generated displaying the cluster groups [fig. 2-10(b)].

c. The machine automatically compares the spectral properties of these homogeneous groups to the spectral properties of the randomly selected pixels which have been identified and labeled by the analyst. Based on its "closeness" or "similarity" to the analyst-labeled pixels, each cluster is labeled wheat or nonwheat.

d. "Conditional" clusters, the properties of which are significantly different from any signatures labeled by the analyst, are automatically flagged for more intense examination; a color map is generated to display these conditional clusters. All unconditionally labeled wheat clusters are displayed in a single color and the nonwheat clusters in a different color, as shown in figure 2-11(a). If a later examination of the spectral and spatial properties of these conditional clusters by the analyst does not agree with the label assigned by the automatic labeling logic, the analyst may change the label. If the cluster comprises only a small part of the scene, as in figure 2-11(b), the analyst may assume that the automatic bias correction will account for any significant error introduced. In cases where significant numbers of conditional clusters occur, the analyst would be required to resubmit the segment data for additional analysis.

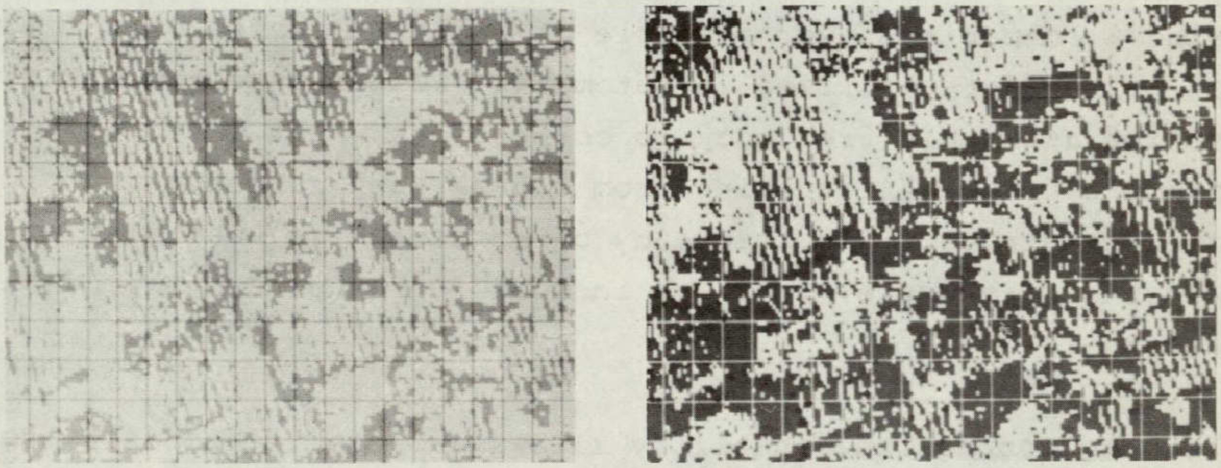
ORIGINAL PAGE IS
OF POOR QUALITY



(a) Color infrared image. Wheat emergent stage: W = winter grains; N = nonwinter grains.

(b) Cluster map. Bright blue, and cyan = winter grains; other colors = nonwinter grains.

Figure 2-10.— Landsat color imagery and cluster map of Fergus County, Montana, November 11, 1976.



(a) Conditional cluster map. Green = winter grains; yellow = nonwinter grains.

(b) Classification map. White = winter grains; gray = nonwinter grains; black = thresholded.

Figure 2-11.— Conditional cluster and classification maps, Fergus County, Montana, November 11, 1976

C-2

After the machine clustering and automatic labeling logic are completed, the labeled clusters of all 22 500 pixels in the scene are characterized parametrically by the machine as multivariate normal distributions. Means and covariances are computed utilizing all measurements in each cluster. Each pixel is then machine classified as wheat or nonwheat utilizing a maximum likelihood decision rule.

This machine processing algorithm sequence processes up to four temporal acquisitions of four-channel Landsat multispectral data. The four-channel four-date Landsat data are treated by the machine as a 16-element measurement vector. In the event a fifth acquisition is obtained, a feature selection algorithm automatically selects the "best" three of the four acquisitions residing on the data base and replaces the "worst" acquisition by the incoming (fifth) acquisition. Upon completion of classification, the frequency of agreement between the machine-assigned and the analyst-assigned labels is computed automatically using a comparison over a sample of analyst-labeled dots independent of the dots utilized in automatic cluster labeling. The machine uses this frequency to correct its wheat proportion estimate for bias resulting from causes such as automatic cluster labeling errors. The frequency of agreement is used also as a performance measure; i.e., an indicator of the need for possible rework.

The bias correction capability allows an incoming Landsat acquisition to be processed automatically utilizing analyst labels from an earlier acquisition. If the analyst reviews the labels and decides no significant change has been made, an automatic estimate is obtained utilizing more recent Landsat data with potentially improved spectral separability. If the analyst review indicates the need for a modest number of label changes,

ORIGINAL PAGE IS
OF POOR QUALITY

the estimate can be updated without reprocessing, simply by utilizing the bias correction procedure to account for shifts in dot labels.

In summary, once the analyst has assigned labels to each spectral class, machine processing furnishes the bias-corrected wheat proportion estimate without further intervention by the analyst. In addition, the analyst receives many products which allow him to quantitatively assess the quality of the segment estimate. In cases where problems are encountered, several diagnostic products are provided to the analyst to facilitate rework.

From an operational viewpoint, much less intensive labor will be required using second-generation rather than the first-generation procedures. Analyst contact time for segment analysis has steadily declined from approximately 12 hours per segment in Phase I to 6 hours in Phase II. An analyst contact time of 3 to 4 hours is projected for Phase III using the new procedures; this reflects an efficiency increase by a factor of 4 from Phase I performance. In addition, the Phase III procedures will provide the analyst improved and more repeatable decision-making procedures. The spectral differences between wheat and nonwheat and between small grains and nonsmall grains, as observable on multiple Landsat acquisitions, are an invaluable aid to LACIE analysts in manually identifying wheat or small grains in order to train the classifier. In addition, Procedure 1 permits the extensive use of multitemporal processing for the first time in LACIE.

2.5.2.2 Signature Extension

The signature extension program is based on expanding the single-segment training concept presently used in the LACIE to a multisegment training concept. In the single-segment training concept, training data from a given segment are used to classify only that segment. In the multisegment training concept, training data from several segments are used to classify these and other segments. The approach for the design of this signature extension concept is based (1) on a stratified statistical sampling design for efficiently selecting a small set of training segments and (2) on research into correction procedures for minimizing "noise" effects due to haze and soil variations.

The first design, called Procedure B, has been developed and partially tested. The key steps in this procedure are:

a. The segment data are corrected for Sun angle and haze depth by methods not requiring ground-truth data.

b. Segments are assigned to areal strata which have been constructed based on ancillary variables such as soil type, climatology, and cropping practices.

c. Segments in a given areal stratum are clustered to produce a spectral stratification. The smallest set of segments which adequately sample these clusters is picked for training segments.

d. Training samples within the training segments are labeled.

e. The proportion of wheat in the segments assigned to a given areal stratum is computed as:

$$\sum_{N} \frac{\text{Number of samples from cluster N labeled wheat}}{\text{Number of samples in cluster N}}$$

$$\times \frac{\text{Number of samples in cluster N}}{\text{Total number of pixels}}$$

Preliminary tests of this procedure showed that reasonably good estimation results could be achieved with a training gain of 3; i.e., only one-third of the segments classified were used for training. The tests were run on 17 Phase II LACIE segments in Kansas which averaged 23.12 percent wheat. The procedure produced an estimate of 20.93 percent with a standard error of

ORIGINAL PAGE IS
OF POOR QUALITY

9.1 percent at the segment level. This accuracy magnitude is commensurate with that obtained utilizing single-segment training.

2.5.2.3 Quantitative Displays and the Application of Statistical Pattern Recognition to Image Interpretation

As part of Procedure 1, numerical displays of spectral data were developed. One display, the trajectory plot, provides the analyst a temporal (time history) account of the spectral changes of a point in an image. The plot displays the position of a dot represented as a vector of brightness and greenness at the times used in the analysis (generally within the four biowindows). Another series of displays includes scatter plots of all the dots plotted as brightness versus greenness. Separate displays show the scatter of unlabeled dots, of the dots labeled by the classifier on a previous pass, and the dots as labeled by the analyst again using previous pass information.

Research was begun also on a more quantitative way of assigning labels to a given point. This method is one in which a series of questions is asked of the analyst requiring in general YES, NO, or an indeterminate type of response. These questions are based on ancillary data (such as crop calendar data, cropping practices, and climatology) and on spectral data in the form of the above-mentioned spectral plots and color imagery. The responses to these questions are scored first by assigning a weight to each question and then by adding the weighted responses. Scores above a certain number lead to the decision that a dot is wheat, scores below another number imply nonwheat, and scores between these two numbers lead to a classification of indeterminate for a dot. The weights are derived by regressing actual dot labels against weight scores over selected blind-site data from a previous year.

2.5.2.4 Development of a New Sampling Strategy

The original LACIE sampling strategy required that the segment be allocated based on historical data at the county level. Such data generally are not available at such a small geographic level for countries outside the United States. Therefore, a new sampling strategy was developed using available global information. Such data include Landsat full-frame imagery, soil type maps, limited historical data, and previous LACIE sample-segment estimates. Based on such data, areas to be mensurated are stratified into what is called agrophysical units. Segments are then allocated within the strata based on an optimum allocation computed from production and production variance estimates.

As discussed in section 2.3.7.2, comparison of this new sampling strategy with the old strategy in Kansas shows that, with the new strategy, comparable results can be obtained with fewer segments allocated. With 121 segments allocated (of which 113 were aggregable) according to the old strategy, the relative difference between LACIE and SRS was -5 percent with a CV of 5 percent. With the new sampling strategy, 84 segments were allocated (of which 75 were aggregable), and the resulting relative difference and CV were -2 percent and 6 percent, respectively.

2.5.2.5 Development and Testing of Wheat Yield Models

The original LACIE yield models were developed for specific regions by regressing trend and weather-related variables against several years of historical yields. During Phase II of the project, a new modeling effort was undertaken to develop a model that would be more applicable to foreign countries. The approach, which would require a shorter series of historical data, was first to develop a yield predictor based on regressing weather-related variables against historical agricultural plot data acquired across many experiment stations in the United States and then, through the use of a local adjustment multiplier, to

ORIGINAL PAGE IS
OF POOR QUALITY

apply that predictor to a given region. The estimation of the local adjustment, called the Management and Productivity (MAP) factor, is believed to require less historical data than would be required to develop the above-mentioned regression models. This is based on the reasoning that the MAP factors do not vary greatly over time and therefore do not require large amounts of historical data to obtain a good estimate.

Preliminary tests of this Feyerherm model indicate that its performance would be commensurate with the earlier developed LACIE models in the United States. The preliminary results of this effort are discussed in section 2.3.7.2.

2.5.2.6 Status of Second-Generation Tests

The Test and Evaluation Plan for the Feyerherm (KSU) and CCEA Phase III Yield Models (ref. 4) specified three tests to be made of the Phase III yield models:

a. An evaluation of Feyerherm yield models for the States of Kansas and North Dakota and three U.S.S.R. oblasts (Khmel-Nitsky, Kurgan, and Tselinograd)

b. A comparative test and evaluation of Feyerherm and CCEA yield models for Kansas and North Dakota and the three U.S.S.R. oblasts

c. An evaluation of the CCEA foreign yield models

The status of each test as of mid-Phase III is presented as follows.

The data for evaluating the Feyerherm yield model over a 10-year period for Kansas, North Dakota, and the three U.S.S.R. oblasts were received in April of 1977. These data were evaluated and, as a result, the models for spring wheat have been revised. The U.S.S.R. winter- and spring-wheat model boundaries which resulted from these tests are given in figures 1-5 and 1-6 (section 1.10). Test data for the revised spring-wheat model for North Dakota are being evaluated at this time.

A comparison of the KSU and CCEA models was made; but, as a result of an erroneous procedure used in the evaluation, a new comparison will be made with results provided in the next evaluation report.

The third test, which evaluated the CCEA foreign models, is complete.

CCEA models for other foreign countries are available; i.e., Australia, Argentina, India, and Canada.

- One model for one state has been evaluated for Brazil. Another model has been requested for an additional state in Brazil, but no test data have been received.
- The models for Argentina have been completed and tested.
- The Australia models consist of five state models, and test data for all have been evaluated.
- India has only one state modeled, and the test data have been evaluated.
- In Canada, 16 models are available, and the data have been evaluated.

The testing of the Phase III models is only partially completed. Completion is expected the first of October 1977.

2.6 TECHNICAL ISSUFS THROUGH MID-PHASE III

As of the completion of winter-wheat processing in the yardstick region for Phase III, the following technical issues have surfaced and are being worked.

2.6.1 PROCEDURE 1

ORIGINAL PAGE IS
OF POOR QUALITY

While the initial version of Procedure 1 has performed quite well, much has been learned through processing Phase III data. Many improvements which would lead to even greater efficiency and accuracy have been defined.

- Labeling – The largest single error source in acreage estimation is analyst labeling. Although preliminary indications are that the Phase III analyst procedures and labeling aids provided by Procedure 1 have improved labeling accuracy, a significant amount of confusion of wheat with other categories still exists. Blind-site analyses in Phase III indicate that pasture has been a major source of confusion in low-density wheat segments. Preliminary test results of analyst procedures for separating spring wheat from spring small grains were encouraging and demonstrated the feasibility of estimating wheat directly from Landsat data. The use of quantitative spectral aids such as the green number are critical elements of these techniques. Utilizing Procedure 1, two major classes of dots are labeled. One set of dots, called type I dots, is used to initialize clustering and in cluster labeling; dots which are judged by the analyst to be on field or spectral boundaries are eliminated from this set. Thus, the type I dots are "pure pixels." Type II dots are utilized in bias correction and estimate evaluation. These dots include both boundary and nonboundary pixels and contain no dots belonging to the type I set. Labeling accuracy tests conducted to date indicate that analyst labeling accuracy is significantly better for type I than for type II dots. Further, these tests indicate that the labeling errors on the type II dots contribute significantly to the proportion estimation bias. As reported in the LACIE Phase II Evaluation Report (ref. 5), in cases where there is a reasonable probability that a pixel can be either wheat or nonwheat (e.g., a boundary pixel), in order to refrain from guessing, the analyst is faced with two alternate labeling procedures: One, a "wheat conservative" procedure, in which case the analyst decides wheat only in the event he is certain the pixel is wheat; otherwise, the pixel is labeled nonwheat. This procedure will obviously lead to underestimates of wheat. The alternate, called the "wheat liberal" procedure, in which the pixel is labeled wheat if there is a reasonable chance it is wheat, will result in overestimates of wheat.

The analyst must be given a procedure for objectively labeling these "border" pixels (i.e., either field of spectral boundary pixels or pixels for which the signatures could reasonably be associated with either wheat or nonwheat). Otherwise, the border pixel labeling errors arising from either a wheat conservative or a wheat liberal procedure will cause proportion estimation bias. To remove the bias arising from the border pixels, a procedure is needed which permits an unbiased estimate of the proportion of wheat contained in a border pixel. Once these proportions are known, they can be utilized in Procedure 1 to remove wheat proportion estimation bias. Such bias correction procedures are being investigated within the LACIE RT&E program.

- Dot Labeling Allocation Strategy — Since the amount of analyst time is a key operant in the cost effectiveness of a system and since analyst labeling errors are currently the largest source of estimation error, it is extremely important to develop a dot sample strategy which requires the fewest dots to be labeled for a given accuracy. In the initial version of Procedure 1, the labeling dots of both type I and type II were allocated randomly within the segment. Furthermore, the number of dots allocated was the same for each segment. The statistical theory borne out by experience with Procedure 1 indicates that the dots should be allocated among segments in proportion to the amount of wheat contained in a segment.

Furthermore, the allocation should be a systematic, stratified, random sample within the segment. Efforts are currently underway to develop and test improved dot allocation strategies for the LACIE Transition Year.

- Clustering – At the start of LACIE Phase I, an existing clustering algorithm, the Iterative Self-Organizing Clustering System (ISOCLS, ref. 14), was utilized to delineate the spectral structure of the multispectral data. Early in Phase I, a number of problems forced the abandonment of the algorithm, and the analyst had to delineate the spectral structure from the color infrared imagery. This approach worked reasonably well for the single-pass analysis; however, in multidate cases, the data structure became too complex. Thus, through Phase II, only limited multidate machine processing was available.

  The primary problem with the algorithm was its parametric sensitivity to scene properties and its dependence on acquisition history. In order to use the algorithm, a number of parameters (such as number of clusters, the maximum permissible standard deviation, and the minimum distance between clusters) needed to be specified. Experience with the Landsat data proved that the ISOCLS parameter set which would produce good cluster results was so scene dependent that a new set of parameters was required for almost every segment or for each combination of acquisition dates. This made the algorithm unusable in a highly automated fashion. Clustering investigations by the LACIE RT&E effort through Phase II uncovered two basic problems: (1) a number of mathematical errors in the algorithm and (2) excessive variance introduced into the MSS data by Sun angle and haze effects. By Phase III, the ISOCLS algorithm had been modified to remove many of the mathematical deficiencies, and a Sun-angle-correction algorithm had been implemented. Phase II testing indicated the algorithm to be workable; and the design of a highly automated, cluster-based, machine processing procedure was begun, culminating in Procedure 1.

Experience with the modified clustering algorithm has shown that it performs well and, in addition, it has raised new issues involving (1) the treatment of border pixels and (2) the requirement for extremely efficient procedures requiring a minimum of analyst hours for a given accuracy. Regarding (1), ISOCLS can be used in two distinct modes, both of which initialize clustering with the MSS vectors of a subset of type I dots. This subset is called the starting dots. One mode, referred to as the iterative mode, clusters the MSS vectors in the segment around the starting vectors based on their proximities in spectral space. The mean vector for each cluster is then computed and used as a new starting vector, and the MSS vectors are reassigned based on proximities to the new starting vector set. This process continues for a predetermined number of iterations, after which a split sequence and a combine sequence are initiated for these clusters. Clusters are split if their standard deviation exceeds a certain preselected value. Clusters are combined if their intercluster distance is smaller than a prescribed value. In the nearest neighbor mode of ISOCLS, MSS vectors are assigned to clusters, the centers of which are defined by the initial starting vectors. This assignment again is based on the proximity in spectral space. Essentially, nearest neighbor is the iterative mode aborted after the initial step. Preliminary results to date seem to indicate that the nearest neighbor mode of ISOCLS produces less biased wheat estimates than does the iterative mode. Preliminary investigations have shown this to be the result of three major factors: (1) The clustering is initiated using type I dots, which include no boundary pixels; (2) the iterative mode of ISOCLS clusters boundary pixels into separate clusters from pure pixels; and

(3) these clusters are labeled using a nearest neighbor approach and the type I dots. As a result of these factors, the boundary pixels, which include both wheat and nonwheat, are clustered by the ISOCLS iterative mode and are labeled by the automatic labeling logic as either totally wheat or totally nonwheat. The nearest neighbor clustering mode, on the other hand, tends to arbitrarily assign the boundary pixel to a cluster based on its proximity to the labeled type I starting vector. It is hypothesized that this explains why this latter mode produces less biased estimates. During Phase III, Procedure 1 has utilized ISOCLS in the iterative mode. Further tests are being conducted to support a change to the less expensive and potentially more accurate nearest neighbor mode for Phase III.

The second issue defined for clustering is the need to improve the algorithm to require fewer starting and labeling vectors, in order to obtain accuracy at a lower cost in analyst labeling hours. Currently, the starting and labeling dots are chosen at random from within the segment and therefore at random within the spectral domain. Because at least one starting vector, as well as at least one labeling vector, is needed for each cluster and because multidate segments can have in the range of 40 to 60 clusters, a spectral sampling strategy must minimize the minimum number of vectors initially allocated in order to provide a starting vector for each cluster.

ORIGINAL PAGE IS
OF POOR QUALITY

2.6.2 EVALUATION

The ultimate objective of evaluating the results of a segment analysis is to detect analyst labeling or machine analysis errors which would create unacceptably large proportion estimation errors. Analyst labeling errors are the most difficult to detect. In Phase III, the analyst depends primarily on a review of the spectral properties of the labeled dots to check for consistency in labeling. Procedure 1 has provided several new quantitative spectral aids to facilitate this process: spectral and trajectory plots, which display the Landsat measurement values in two-dimensional displays. However, spectral consistency does not guarantee correct labels; and, to date, there is no practical method of assuring the correctness of the analyst labels without ground truth.

Disagreement between the machine and analyst tables indicates problems with either the automatic labeling of clusters or insufficient spectral separability between certain generic classes. These frequencies of disagreement can be related to both the bias and variance of the proportion estimate. In practice, however, labeling errors by the analyst confound these relationships.

If the analyst has mislabeled a certain dot and consequently the cluster containing the dot, then it is likely that the machine will also follow suit. This results in a good frequency of agreement between man and machine but a biased wheat proportion estimate. Thus, the issue of developing indicators of the accuracy of the wheat proportion estimate at the segment level remains an open one. The evaluation procedure currently is utilized to cross-check for consistency between man and machine in an attempt to ferret out machine labeling problems or, in some cases, analyst labeling errors. However, the degree of consistency is not sufficient to establish the accuracy of the proportion estimate.

2.6.3 REGISTRATION

With the advent of Procedure 1 in Phase III, more regions with significantly smaller field sizes have been successfully analyzed than in Phase II. However, with multitemporal classification, a

ORIGINAL PAGE IS
OF POOR QUALITY

ORIGINAL PAGE IS
OF POOR QUALITY

fundamental limit has been observed to result from misregistration in strip-fallow fields in the U.S. northern Great Plains. Basically, the root mean square ±1 pixel registration specification is being met by the LACIE registration technology. However, this is proving inadequate in many of the strip-fallow areas and is projected to lead to substantial proportion estimation error in these areas. In some segments, multitemporal classification has been abandoned in favor of single-pass data in order to work the smaller fields. This leads to higher error rates, also, as a result of the accompanying drop in spectral separability when going from multitemporal to single-date data.

2.6.4 YIELD ESTIMATION

For the first time in LACIE, the first-generation yield model estimates are noticeably below the SRS estimates of yield. Although the 10-year tests and the 3 years of experience in LACIE operations indicate that the yield models are performing adequately in support of the 90/90 criterion for production, investigations of model performance at the subregional levels have indicated that the models could and should be improved. These studies showed that, in a year with extended episodic conditions, the first-generation yield models would not be adequately responsive to extremely high or low yields and that, during such years, considerable yield estimation error would result. Therefore, a second-generation model was developed to overcome some of the first-generation model deficiencies and to permit operation in a foreign country with a much shorter time series of historical data.

As discussed in section 2.3.7.2, Phase III preliminary testing of this model over limited U.S. geographic regions has indicated that, based on aggregation results, the second-generation yield model performs approximately as well as the initial models and requires a reduced data base for model development. The results of these evaluations will be considered further after testing is complete and before implementation decisions are made for the Transition Year.

2.6.5 SAMPLING IN MIXED-WHEAT REGIONS

With regard to sampling mixed spring and winter wheat in LACIE Phases I and II, segments were allocated based on total wheat statistics, and areas containing both spring and winter wheat (mixed-wheat areas) were arbitrarily designated either winter or spring in proportion to the historical percentage of winter or spring grains grown in the county. Once these segments were so designated, each segment was analyzed for spring wheat only or for winter wheat only, and data were collected only during the growing season appropriate to either the winter or the spring wheat crop calendar. This strategy created a problem for those segments which had significant amounts of both spring and winter wheat. In Phase III, data were collected in the mixed-wheat areas for the "total wheat" growing season — essentially the entire crop year. This was based on the definition that a mixed-wheat area has a probability of both winter and spring wheat being grown in a sample segment. The Phase III data collection scheme acquired the satellite data to estimate both spring and winter wheat grown in all segments, as opposed to the Phase II mode of utilizing one set of segments for winter wheat and a different set for spring wheat. Aggregation and variance estimation methodology was developed and implemented to permit operation in this mode.

Utilizing this mixed spring/winter plan caused a problem in the mixed-wheat area of South Dakota in Phase III, as recorded in the June and July CMR's. The LACIE South Dakota winter-wheat acreage estimates were significantly larger than reasonable based on historical estimates. An investigation disclosed the fact that the mixed-wheat strategy had resulted in many South Dakota segments with almost no winter wheat. In these very low-density segments, CAMS errors tend to be overestimates. These segments were indeed being overestimated. While the absolute difference was not large, the relative difference was, and it created the large South Dakota overestimate. To remedy this problem, the South Dakota segments were redesignated based on historical county statistics to eliminate mixed-wheat designations for segments in counties which typically grow no winter wheat. This redesignation greatly reduced the magnitude of the overestimate, and the August CMR carried the estimates with the revised designation. Montana also was redesignated using the same procedure but with minimal effect because of a larger proportion of both spring and winter wheat. The modification is also being applied; but the effect is expected to be minimal as was the case in Montana.

2.6.6 CROP CALENDAR MODELS

LACIE currently utilizes the Robertson model to operationally predict wheat growth stages. This model utilizes daily maximum and minimum temperature inputs from the WMO ground network. Currently, no growth model is available to LACIE for crops other than wheat. Other crops are assumed to experience the same delay or advance from nominal as wheat. It is certain that, with key confusion crops such as spring barley and native grasses, a real-time growth model would improve the analyst's ability to identify wheat. Regarding the wheat crop calendar model, three key issues remain.

a. General model improvements are required, particularly the development of a planting date prediction model to improve the accuracy of growth stage estimation prior to dormancy. This item may be particularly crucial to improved early-season estimation.

b. Increased understanding and quantification of the manner in which growth stage prediction errors affect yield and acreage estimation error are needed.

c. A more accurate and efficient method for making ground observations of growth stage prior to heading for the purpose of improved model evaluation should be developed.

ORIGINAL PAGE IS
OF POOR QUALITY

3. OUTLOOK FOR TRANSITION YEAR AND BEYOND

As currently envisioned, the LACIE is a major step toward developing a remote sensing survey technology capable of global food and fiber monitoring. The contribution of the LACIE will be a demonstration of "proof of concept" of this new technology for significantly improving currently available information on one major global crop — wheat. By the end of LACIE Phase III, it is anticipated that the experiment will have (1) demonstrated the utility of remote-sensing-survey technology over the U.S. Great Plains and an important foreign country, (2) identified key areas where the technology needs improvement, and (3) brought the USDA advanced system to a point of initial testing. At this time, a transition period will be required to complete, document, and transfer the LACIE technology to an evolving USDA system to exploit the experimental accomplishments of the LACIE. This overall development, demonstration, and application program will be focused on a global food and fiber monitoring system. The next logical steps are (1) to continue refining the technology for subsequent transfer of both skills and technology to the USDA ATS and (2) to adapt the LACIE experience and technology to multicrop food and fiber inventory applications.

Early in LACIE Phase II, an effort was initiated to accomplish the transfer of technology to the USDA for further evaluation. This effort is now an approved follow-on to LACIE and is officially designated LACIE Transition Year. The objective of the Transition Year is the orderly transfer of proven technology to USDA facilities and personnel for further test and evaluation. The Transition Year represents a culmination of various improvements, expansion to the Southern Hemisphere, and a final test in the LACIE context prior to transferring the latest baseline technology to the USDA for application testing. It also represents the initial operation of a USDA test system on an important region where the LACIE has already demonstrated the applicability of the technology. Specifically, the areas to be studied, the level of technology to be used, and the learning expected to accrue are provided in the following subsections.

3.1 THE U.S. GREAT PLAINS AND INTENSIVE TEST SITES

The yardstick region and other ITS's will be worked with the most advanced technology available in real time, and reports will be made monthly on the following subjects:

- Benefits and efficiencies of new sampling strategies with stratification according to agrophysical units
- A full year's experience using the entire Procedure 1
- An understanding of the vagaries of yet another year (All 3 years thus far have been different.)
- Adequacy of direct discrimination of wheat if results in Phase III are encouraging
- Better understanding of the bias in LACIE estimates (possible because of longer series of data)

3.2 INDIA

One or two states in India will be worked with baseline technology in real time or approximate real time and reported at monthly intervals through the growing season (after adequate emergence) in the following areas:

- Adequacy of yield models and the ACC in an area where much wheat is of the dwarf variety and a significant amount is irrigated
- Adequacy of the LACIE technology in an area where small fields predominate

3.3 THE SOUTHERN HEMISPHERE

Countries of the Southern Hemisphere (including Australia, Argentina, and Brazil) will be studied using baseline technology. Two (mid-season and at-harvest) or three (early-season, mid-season, and at-harvest) estimates will be simulated during the analysis period to determine:

- The applicability of the LACIE baseline technology to new situations (different climatic conditions and low latitudes)
- The performance of yield models using different parameters for moisture assessment
- The adequacy of technology in areas where ancillary data are sparse
- Experience gained in areas with minimal ancillary data

3.4 CANADA

The USDA ATS will assess 30 blind sites and ITS's for:

- Replication of Phase III analysis
- Quantification of USDA ATS performance in an area where ground data are available

3.5 THE U.S.S.R. BY LACIE

An area (still to be determined) of the winter- and mixed-wheat growing region of the U.S.S.R. will be studied by the LACIE, using baseline technology and reporting early-season, mid-season, and at-harvest estimates to obtain:

- Better understanding of the bias in estimates from the extended length of analysis
- Improved understanding of U.S.S.R. statistics, which are believed to be unreasonably stable in acreage with variations attributed to yield

- Possibly a test of the efficiency of the new sampling strategy in a second country

3.6 THE U.S.S.R. BY THE USDA APPLICATION TEST SYSTEM

The USDA will investigate the pure spring-wheat region of the U.S.S.R. using prototype advanced technology and reporting monthly estimates:

- For the winter wheat areas
- For initial evaluation of the USDA operational test system performance

3.7 STATUS OF THE ADVANCED SYSTEM

The scope of the USDA ATS for the Transition Year includes 70 U.S.S.R. winter-wheat segments for system startup processing in the November-to-April 1978 time frame and as many as 800 segments in the U.S.S.R. spring-wheat area to test system loading capabilities and region estimating in the post-April period. Transition Year activities will not represent an attempt at an operational system but rather a test of procedures and technology transferred from LACIE.

A request for proposal (RFP) for hardware to support Transition Year processing was written and issued on January 19, 1977. The technical evaluation of bids received began on April 18, 1977, with an award on June 16, 1977, to Ford Aerospace and Communications Corporation to deliver the system by October 14, 1977. The configuration selected includes a Digital Equipment Corporation (DEC) PDP 11/70 computer; a Floating Point, Inc., AP-120B Array Processor; and an analyst station with International Imagery Systems (I^2S) color cathode-ray tubes (CRT's). Analyst station software, as well as general purpose data processing software, will be provided.

The CCEA yield model is being written for execution on the USDA ATS, along with the crop calendar presently being used at JSC. The LACIE CAS software will be utilized for aggregation and reporting capabilities during the application test. This requires interfaces to input the classification results to the CAS and to return aggregated results to the ATS. These interfaces are being developed at this time.

The Small Fields Procedure for classification will be used for early startup processing (fall 1977) on the USDA ATS. Efforts are currently underway to define Procedure 1 modules for inclusion in the system in the April 1978 time frame. These selected modules, along with other optional tasks in the area of data base generation, may be partially implemented by contractor resources under a software RFP to be issued in late summer 1977.

3.8 STATUS OF USDA PRODUCTION, ACREAGE, AND YIELD ESTIMATION SYSTEM

A preliminary RFP for the acquisition component of the USDA Production, Acreage, and Yield Estimation System (PAYES) has been prepared. The RFP will be completed when agreements as to the sources and forms of the data are finalized. The schedule calls for release of the RFP in early November 1977, with an award sometime in March 1978. Delivery of the acquisition hardware is scheduled for late fall 1978.

The system will not become fully operational until spring of 1979 because of the time required to procure, install, and make operational an extremely wide-band data link to drive the acquisition component. This data link is being funded with fiscal year 1979 funds. The overall schedule is now critical and will slip on a day-to-day basis until data source agreements are final.

ORIGINAL PAGE IS
OF POOR QUALITY

3.9 GOALS OF LACIE TRANSITION YEAR

In addition to the Transition Year efforts, the technology developed in LACIE will be adapted to inventory production of other food and fiber crops. These may include corn, rice, soybeans, and nonfood crops such as forest and timber. It will also be adapted to monitor foraging conditions within the world's important rangelands. This increased capability conceivably could be developed and incorporated in the middle to late 1980's in a second-generation global food and fiber monitoring system.

The goals of the LACIE, the Transition Year, and the technology expansion to a multicrop application will continue to require strong supporting research and technology development efforts within the research community. In this regard, LACIE can be considered as a paradigm for multicrop applications. That is, estimation of production for other crops will involve estimation of the same fundamental elements involved in wheat production estimation: crop acreage, average plant or producing unit population per acre, and average productivity per producing unit. It should be emphasized that the estimation approach utilized to date in LACIE is not the only approach which can be taken to estimate these quantities. Quite possibly, modifications of the LACIE approach will produce a more optimum survey approach for applications different from global wheat estimation. However, to a large extent, all such approaches will involve the same data input and analysis systems required for the LACIE, along with many of the same solutions to technology problems.

More specifically, the LACIE approach to date has utilized Landsat data primarily to estimate wheat acreage for harvest and meteorological data primarily to estimate the average productivity or yield for each acre harvested. In a sense, this separation is artificial; much information is available in the spectral data relating not only to total acreage but also to the plant population density within the acreage. In addition, information relating to plant condition and thus average yield is also furnished, along with plant environment and plant characteristics which can be measured well in advance of harvest and are known to be correlated with final yield. Therefore, a model which includes the effects on yield not only of the plant environment but also its physical characteristics (height and stand density, from which early yield estimates based on soil moisture may be made) will be a significant improvement over models utilizing only meteorological data. Potential quantitative connections through modeling involve efforts which relate the leaf-area index to evapotranspiration, the leaf-area duration to yield, and the leaf-area index to Landsat spectral response. With the advent of thermal sensing on Landsat-C, additional information will be available as potential predictor variables for crop yields.

Conversely, meteorological data also contain much information relevant not only to average productivity but also to planted and harvested acreage. For example, the LACIE early-season estimates of emerged acreage are a function both of the total wheat planted and that expected to be harvested. This fraction within a segment is related to the average growth stage within the segment, which, in turn, is strongly related to the segment temperature and precipitation history. Thus, the early-season LACIE estimates of emerged acreage could be used in a regression model involving both temperature and precipitation inputs to predict the total acreage to emerge at a later date. The emerged detectable acreage is related also, through meteorological and economic factors, to the acreage to be harvested. Based on an analysis of these factors, models which relate acreage at any one point in time to that anticipated for harvest could be developed.

Since meteorological and spectral data are both strongly related to total area, plant population density, plant condition, and (as a consequence) total production, it is anticipated that the survey models utilized for the LACIE will evolve toward forms which simultaneously account for these effects in a more integral fashion. In such a form, the production, acreage, and yield estimators would each involve predictor variables based on both spectral and meteorological and even agronomic and economic data, such as fertilizer application rates, cropping practices, and prices.

Another area for development within the near future is improved sensing and measurement of the basic predictor variables themselves. To date, the LACIE has utilized first-generation Earth-resources satellite data and meteorological data obtained from the ground stations. With the advent of the second-generation Earth-resources satellite, Landsat-C, and the development of the capability to utilize environmental satellite data to obtain more complete coverage for temperature and precipitation estimates, the survey estimates should improve significantly. The LACIE analysis experience has indicated that the Landsat data itself contains information regarding temperature and moisture, as these factors are manifested in crop condition and loss of vigor resulting from drought. Parameters such as soil moisture or, alternatively, precipitation and temperature can probably be more reliably and accurately estimated from a combination of Landsat-type and meteorological satellites.

The direction for the future, then, is the development of crop production estimation models based on both agromet and spectral data, which account for the influence of these data on both acreage and productivity. In addition, these models and the approach must be adapted to the other major global food and fiber crops. Improvements in survey estimates will be derived from basic improvements of the predictor variables themselves as a second generation of land satellites becomes available and as environmental satellite data, along with Landsat data, are used in estimating these parameters.

The LACIE participants have begun to plan a technology development program required to support the future implementation of global food and fiber monitoring systems. The methodology to best ensure a suitable technology base, together with an adequate understanding of its use, needs to be developed over the next year or two and vigorously implemented, if its output is to be available for the middle to late 1980's.

4. REFERENCES

1. MacDonald, R. B.; and Hall, F. G.: LACIE: A Look to the Future. Presented to the 11th Int. Symp. on Remote Sensing of Environment, Environmental Research Institute of Mich. (ERIM, Ann Arbor), Apr. 25-29, 1977.

2. USDA: LACIE CAS Monthly Crop Report, U.S. Great Plains. NASA/JSC (Houston), Feb. 8, 1977.

3. Kauth, R. J.; and Thomas, G. S.: System for Analysis of Landsat Agricultural Data. ERIM, NAS 9-14123, May 1976.

4. Test and Evaluation Plan for KSU and CCEA Phase III Yield Models. JSC-11677, LACIE-00622, NASA/JSC (Houston), Apr. 1977.

5. LACIE Phase II Evaluation Report. JSC-11694, LACIE-00453, NASA/JSC (Houston), July 1977.

6. USDA: LACIE CAS Monthly Crop Report, U.S. Great Plains. NASA/JSC (Houston), May 9, 1977.

7. USDA: LACIE CAS Monthly Crop Report, U.S. Great Plains. NASA/JSC (Houston), June 7, 1977.

8. USDA: LACIE CAS Monthly Crop Report, U.S. Great Plains. NASA/JSC (Houston), July 11, 1977.

9. USDA: LACIE CAS Monthly Crop Report, U.S. Great Plains. NASA/JSC (Houston), Aug. 10, 1977.

10. USDA: LACIE CAS Unscheduled Crop Report, U.S. Great Plains. NASA/JSC (Houston), Apr. 6, 1977.

11. USDA: LACIE CAS Unscheduled Crop Report, U.S. Great Plains. NASA/JSC (Houston), Apr. 22, 1977.

12. Thompson, D. R.; and Wehmanen, O. A.: The Use of Landsat Data To Detect and Monitor Vegetation Water Deficiencies. Proc. 11th Int. Symp. on Remote Sensing of Environment, ERIM (Ann Arbor), Apr. 25-29, 1977.

13. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiods. Int. J. of Biometeorology 12, 1968, pp. 191-223.

14. Minter, R. T.: Computer Program Documentation, ISOCLS, Iterative Self-Organizing Clustering Program. LEC/HASD, NASA/JSC program MSC-C094 (Houston), Oct. 1972.

APPENDIX

DATA USED FOR ASSESSMENT OF LACIE ACCURACY

APPENDIX

DATA USED FOR ASSESSMENT OF LACIE ACCURACY

A.1 ESTIMATES OF THE STATISTICAL REPORTING SERVICE

The SRS makes estimates throughout the growing season in the United States for a large number of agricultural commodities. For winter wheat, the estimates have different bases at different times of the season as follows:

1. December-April — Estimates are for seeded areas and come from the December enumerative survey of fall-planted crops and the fall mail survey. The yield for a seeded area is derived from mail survey estimates of condition made by farm operators. Such condition estimates are correlated to historical records of harvested production per unit of the seeded area to relate estimated condition to expected production per unit of the seeded area.
2. May-June — At this point in the season, the SRS normally uses the mail survey and the objective yield survey to estimate acreage and yield for harvested areas.
3. July-September -- In the June 30 enumeration, the first accurate estimate of acreage for harvest is made, and yield for harvested acreage is estimated from the objective yield survey (actual field measurements of such factors as plant density, etc.).
4. December — This report reflects revised estimates of acreage harvested, yield, and production. Estimates are based on mail surveys, farm census data from each state, grain shipments, and various other sources of check data.

For spring wheat, a similar sequence of estimates is made as follows.

1. January – First report of intentions to plant; data in this report are based on mail surveys.
2. April – Second report of planted area and intention; data in the report are based on mail surveys.
3. June – First estimate of area planted; data in this report are based on the June enumerative survey and the June area survey.
4. October-December – Same reports for winter wheat.

A.2 ESTIMATES OF THE FOREIGN AGRICULTURAL SERVICE

The FAS makes estimates throughout the growing season in various foreign countries for various agricultural commodities. For wheat in the U.S.S.R., different bases are available at different times of the year as follows:

1. February time frame – The production of winter wheat is scaled from the planned production of small grains using historical data. Acreage is similarly scaled, and yield is computed; this provides an informal figure internal to USDA and is not a published estimate.
2. June – The initial estimate of small grains production and area is published and includes inputs from attached reports, historical trends, meteorological data, etc. In late June, an initial estimate of winter wheat is made using the same data sources.
3. July and later – Refined estimates are made for all small grains, based on the same sources used for June estimates, additional field observations by visiting USDA teams, and U.S.S.R. data as available.

These FAS estimates are not considered sufficiently reliable for a comparison standard, not even in the final production estimates (see fig. A-1). Moderately reliable production estimates based

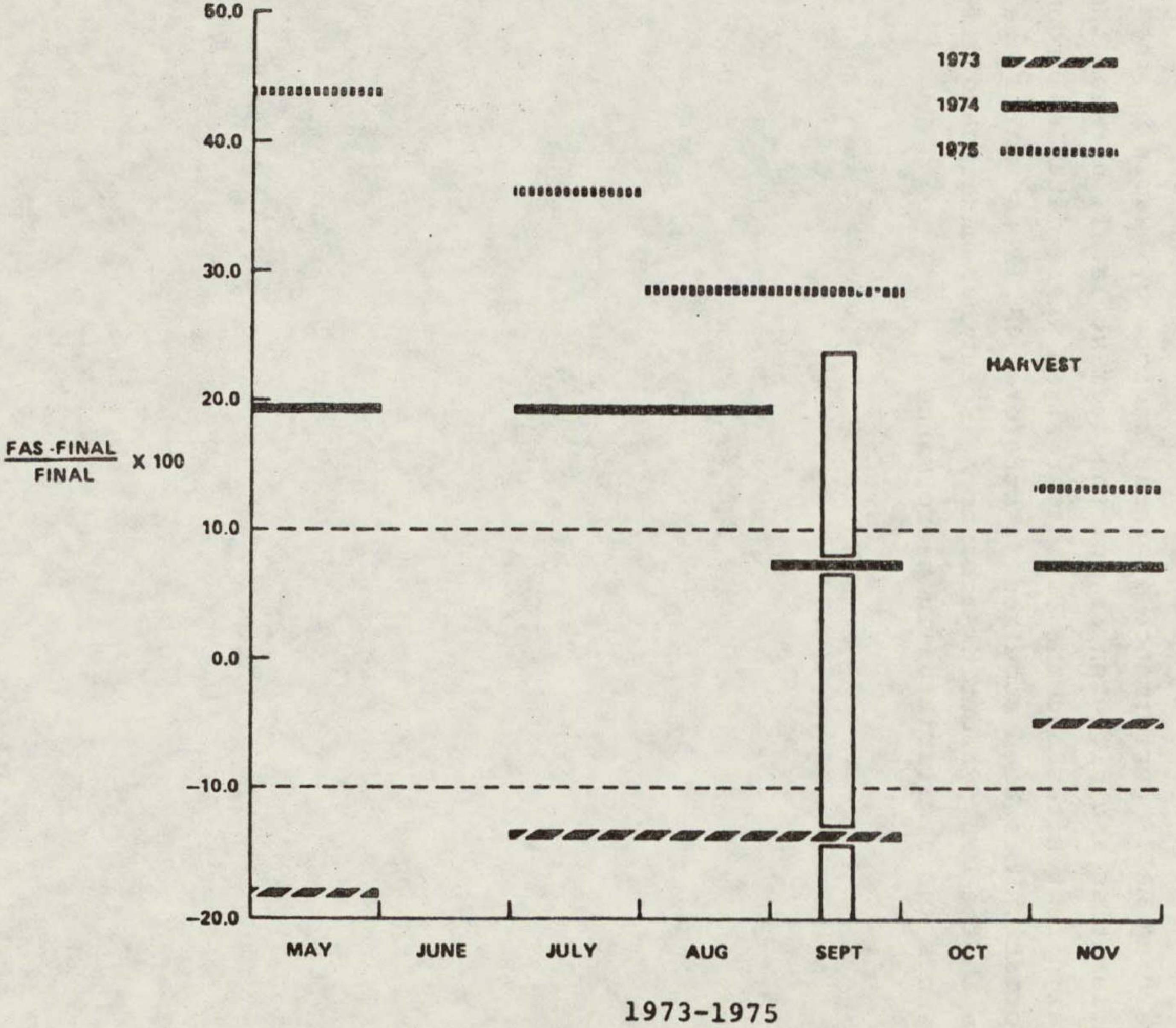


Figure A-1.— Relative difference between USDA/U.S.S.R. seasonal and U.S.S.R. final wheat production estimates.

ORIGINAL PAGE IS
OF POOR QUALITY

on U.S.S.R. reports are available at the country level about 6 months after harvest and at the indicator level about 1 year after harvest. Even though real-time information is unavailable in the U.S.S.R. and other foreign countries, much can be inferred regarding LACIE performance in these regions by examining the similarities and differences, at the segment level, between the foreign test sites and the U.S. test sites where detailed ground information has been acquired. Therefore, LACIE estimates are make in the U.S. yardstick area to help further understand differences and similarities in performance.

ORIGINAL PAGE IS
OF POOR QUALITY