GEORGE HANUSCHAK'S COPY

LACIE-00450

JSC-13736

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)

NASA NOAA USDA

LACIE

PHASE I AND PHASE II

ACCURACY ASSESSMENT FINAL REPORT

NASA
National Aeronautics and
Space Administration

Lyndon B. Johnson Space Center
Houston, Texas 77058

APRIL 1978

c. To study the various sources of error in the LACIE estimates of wheat production, area, and yield, quantify these errors where possible, and recommend procedures for reducing the error.

1.2 ACCURACY ASSESSMENT ACTIVITIES

In order to satisfy its objectives, AA carries out general types of evaluations and the results are presented in (1) monthly quick-look reports; (2) a number of interim reports leading up to a final report, and (3) certain special reports. The following paragraphs contain descriptions of the AA evaluations presented in the three types of reports.

1.2.1 ACTIVITIES REPORTED IN THE QUICK-LOOK REPORTS

The quick-look reports contain an evaluation by AA of the LACIE estimates reported in the Crop Assessment Subsystem (CAS) monthly reports (CMR's) and the CAS annual report (CAR). The quick-look reports are released one week following the release of a CMR or a CAR. The CMR's and CAR's contain the official LACIE estimates of wheat production, area, and yield, and the corresponding statistics. The true wheat production, area, and yield for the particular region or country are, of course, unknown. Therefore, to ascertain the accuracy of the LACIE estimates, comparisons are made with a reference standard. In the United States, the reference standard consists of the most recent (at the time of the comparison) estimates released by the Statistical Reporting Service of the USDA (USDA/SRS). In foreign countries, the reference consists of the most recent estimates released by the Foreign Agricultural Service of the USDA (USDA/FAS). The AA quick-look reports contain a comparison of the LACIE estimates of wheat production, area, and yield with the corresponding reference standard, as well as significance tests of no difference at the region or country level. The relative difference calculated at

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE) PHASE I AND PHASE II ACCURACY ASSESSMENT FINAL REPORT

APPROVED BY

D. E. Pitts, Manager
LACIE Accuracy Assessment, NASA/JSC

J. D. Erickson, Chief
RTEB, NASA/JSC

J. L. Dragg, Chief
AES, NASA/JSC

F. G. Hall, LACIE Project Scientist
NASA/JSC

J. Hill, Assistant Deputy Project Manager
for LACIE, NOAA

J. Murphy, Assistant Deputy Project Manager
for LACIE, USDA

R. B. Erb, Manager
LACIE Project Office, NASA/JSC

R. B. MacDonald,
LACIE Project Manager, NASA/JSC

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

February 1978

PREFACE

This document was prepared by Lockheed Electronics Company, Inc., Systems and Services Division, Houston, Texas, from materials provided by the National Aeronautics and Space Administration and Lockheed Electronics Company, Inc. This work was done under contract NAS 9-15200 for the Earth Observations Division, Space and Life Sciences Directorate, National Aeronautics and Space Administration, Lyndon B. Johnson Space Center.

if CV_A for that level was less than 4.30 percent and if the acreage estimate was unbiased. In Phase I the estimate of CV_A at the national level was 3.74. Therefore, the 90/90 criterion would have been satisfied if the acreage estimate were unbiased. In fact some bias would be allowed, since 3.74 is somewhat smaller than 4.30. The relative differences between the LACIE and USDA/SRS estimates indicated that some bias was indeed present, but no accurate estimate of this bias was performed in Phase I; therefore, it is not possible to say whether or not the results satisfied the 90/90 criterion at the national level.

The area of most concern in Phase I was North Dakota, which had a relative difference of -74.6 percent. Blind site investigations indicated that the primary source of this problem was sampling error.

The experience gained in Phase I was used in developing the CAMS system for Phase II. Several changes were made on the basis of this experience. In particular, more sample segments were allocated to North Dakota, and the classification procedures developed for the CAMS rework experiment became the basis for the Phase II CAMS operations.

2.1.2 PHASE II

In Phase II, estimates were made for acreage, yield, and production. Generally the LACIE yield estimates were quite close to the USDA/SRS estimates and therefore can be considered satisfactory. However, the acreage and production estimates at the USGP level were low compared to the USDA/SRS estimates, due primarily to significant underestimates for spring wheat in the four U.S. northern Great Plains (USNGP) states and for winter wheat in Oklahoma.

CONTENTS

Section | Page

Appendix | Page

TABLES

Table		Page

FIGURES

ACRONYMS, ABBREVIATIONS, AND DEFINITIONS

AA	Accuracy Assessment.
ACC	adjustable crop calendar.
agromet	agricultural/meteorological.
biowindow or biophase	biological window, biological phase – a Landsat data acquisition period that is related to the biostages of wheat development. The LACIE approach is based on the judgment that wheat can be separated adequately from other crops by analysis of up to four acquisitions of Landsat data during the growing season. The biowindow may be updated if there is a significant lag or advancement in the current crop calendar. The sequence chosen includes acquisitions during the following biowindows: 1. Crop establishment – from 50 percent tillering to 50 percent jointing (biostage 2.3 to 3.0). 2. Green – from 50 percent jointing to 50 percent heading (biostage 3.1 to 4.0). 3. Heading – from 50 percent heading to 50 percent soft dough (biostage 4.1 to 5.0). 4. Mature – from 50 percent soft dough to 50 percent harvest (biostage 5.1 to 6.0).
biostage	biological stage – the specific stage of development of a crop which can be recognized by a major change in plant structure; i.e., emergence after germination, jointing, heading, soft dough, ripening, and harvest, which are represented by integers on the Robertson Biometeorological Time Scale.
blind sites	LACIE sample segments chosen at random for which ground truth is obtained in order to test classification performance. The identity of the blind sites is withheld from the CAMS analysts so that these segments will be treated the same as the other segments.
BMTS	Biometeorological Time Scale.

CAMS	Classification and Mensuration Subsystem.
CAS	Crop Assessment Subsystem.
CCEA	Center for Climatological and Environmental Assessment — an organization of the National Oceanic and Atmospheric Administration, Columbia, Missouri.
classification	in computer-aided analysis of remotely sensed data, the process of assigning data points to various classes by a testing process in which the spectral properties of each unknown data point are compared with spectral properties typical of these classes.
classification error	a measure of the degree to which the LACIE CAMS either overestimates or underestimates the wheat acreage in a specific area.
CMR	CAS Monthly Report.
CRD	Crop Reporting District — a geographical area used by the U.S. Department of Agriculture for the collection and reporting of agricultural information; each district consists of several counties.
crop calendar	a calendar depicting the biostages of the major crop types within a specified region during a calendar year.
crop calendar adjustment	an adjustment made to the normal crop calendar on the basis of current meteorological data.
CUR	CAS Unscheduled Report.
CV	coefficient of variation (standard deviation divided by the mean).
DAPTS	Data Acquisition, Preprocessing, and Transmission Subsystem.
Group 2 segment	LACIE segment in a county that historically produces small quantities of wheat/small grains; samples are allocated with probability proportional to size.
IE	Information Evaluation.
IMR	IE Monthly Report.

ITS	intensive test site — a LACIE segment in the United States or Canada on which detailed crop information is collected by using ground and airborne equipment.
JSC	Lyndon B. Johnson Space Center of NASA, Houston, Texas.
LACIE	Large Area Crop Inventory Experiment.
Landsat	Land Satellite — formerly called ERTS (Earth Resources Technology Satellite); operates in a circular, Sun-synchronous, near-polar orbit of Earth at an altitude of approximately 915 kilometers; orbits Earth about 14 times a day and views the same scene approximately every 18 days.
LEC	Lockheed Electronics Company, Inc.
MSE	mean square error.
MSS	Multispectral Scanner System or multispectral scanner — the remote sensing instrument on Landsat that measures reflected sunlight in various spectral bands or wavelengths.
NASA	National Aeronautics and Space Administration.
NOAA	National Oceanic and Atmospheric Administration.
90/90 criterion	criterion that the LACIE U.S. Great Plains at-harvest production estimate be within 10 percent of the true value with a probability of at least 0.9.
PFC	production film converter.
PPS	probability proportional to size.
Sample segments	the 5- by 6-nautical-mile areas used as samples in LACIE to make acreage estimates. They are selected by a sampling strategy which is described in appendix A.
Thresholding	a procedure in the CAMS classifier whereby pixels which have a very low probability of belonging to any class are not classified. These pixels are said to have been thresholded.
USDA	U.S. Department of Agriculture.

USDA/ASCS	USDA Agricultural Stabilization and Conservation Service.
USDA/SRS	USDA Statistical Reporting Service.
U.S. Great Plains (USGP) (USSGP) (USNGP)	The U.S. Great Plains (USGP), an area encompassing the nine states of Colorado, Kansas, Minnesota, Montana, Nebraska, North and South Dakota, Oklahoma, and Texas; it is divided geographically into (1) the U.S. southern Great Plains (USSGP), which includes Colorado, Kansas, Nebraska, Oklahoma, and Texas, and (2) the U.S. northern Great Plains (USNGP), which includes Minnesota, Montana, and North and South Dakota.

1. INTRODUCTION

The Large Area Crop Inventory Experiment (LACIE) is an interagency endeavor of the National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Administration (NOAA), and the United States Department of Agriculture (USDA). Its purposes are (1) to demonstrate the economical benefit to be obtained by using remotely sensed data from the Land Satellite (Landsat) for agricultural applications, (2) to test the capability of a system utilizing remote sensing in conjunction with climatological, meteorological, and conventional data to produce timely estimates of the production of a major world crop prior to harvest, and (3) to validate the technology and procedures for such a system.

In accordance with the objectives of LACIE, the Accuracy Assessment (AA) effort is designed to check the accuracy of the products from the experimental operations throughout the growing season and thereby determine if the procedures used are adequate to accomplish the above objectives.

1.1 OBJECTIVES

The objectives of AA are as follows:

a. To determine whether the accuracy goal of the LACIE estimate of wheat production for a region or country is being met. The LACIE accuracy goal is a 90/90 at-harvest criterion for wheat production. This specifies that the at-harvest wheat production estimate for the region or country be within 10 percent of the true production 90 percent of the time.

b. To determine the accuracy and reliability of early season estimates and estimates made at regular intervals throughout a crop season prior to harvest. This includes a determination of the degree to which the 90/90 criterion is supported at these intervals during the crop season.

c. To study the various sources of error in the LACIE estimates of wheat production, area, and yield, quantify these errors where possible, and recommend procedures for reducing the error.

1.2 ACCURACY ASSESSMENT ACTIVITIES

In order to satisfy its objectives, AA carries out general types of evaluations and the results are presented in (1) monthly quick-look reports; (2) a number of interim reports leading up to a final report, and (3) certain special reports. The following paragraphs contain descriptions of the AA evaluations presented in the three types of reports.

1.2.1 ACTIVITIES REPORTED IN THE QUICK-LOOK REPORTS

The quick-look reports contain an evaluation by AA of the LACIE estimates reported in the Crop Assessment Subsystem (CAS) monthly reports (CMR's) and the CAS annual report (CAR). The quick-look reports are released one week following the release of a CMR or a CAR. The CMR's and CAR's contain the official LACIE estimates of wheat production, area, and yield, and the corresponding statistics. The true wheat production, area, and yield for the particular region or country are, of course, unknown. Therefore, to ascertain the accuracy of the LACIE estimates, comparisons are made with a reference standard. In the United States, the reference standard consists of the most recent (at the time of the comparison) estimates released by the Statistical Reporting Service of the USDA (USDA/SRS). In foreign countries, the reference consists of the most recent estimates released by the Foreign Agricultural Service of the USDA (USDA/FAS). The AA quick-look reports contain a comparison of the LACIE estimates of wheat production, area, and yield with the corresponding reference standard, as well as significance tests of no difference at the region or country level. The relative difference calculated at the zone level (state in the U.S.) is used to indicate problem areas in zones.

1.2.2 ACTIVITIES REPORTED IN THE INTERIM AND FINAL REPORTS

The interim reports are released at regular intervals throughout the crop season. They contain the results of the previous quick-look reports, a discussion of the 90/90 criterion as it applies to the region for which the LACIE estimates of wheat production are available, and the results of investigations of the error sources in the LACIE wheat production estimate.*

Each interim report is built up from the previous one by including data that became available during the interim period. Technical comments on each report are solicited from a variety of sources and are used to upgrade subsequent reports. Early and mid-season evaluations are made in the first and second interim reports; late season and at-harvest evaluations are made in the third and fourth interim reports.

The fourth interim report also serves as a draft for the final report, which contains material which is similar to the interim reports but covers the entire year.

The above schedule was followed in Phase II. In Phase I there were no interim reports and the Phase I final report will be incorporated into the Phase II final report.

1.2.3 ACTIVITIES REPORTED IN AA UNSCHEDULED REPORTS

From time to time, special investigations are carried out that are of interest to LACIE but which are not required on a regular basis such as those mentioned above. These investigations are reported in AA unscheduled reports.

*A detailed description of the error sources in LACIE is given in appendix A.

2. SUMMARY AND RECOMMENDATIONS

2.1 SUMMARY

2.1.1 PHASE I

Phase I of the LACIE project concentrated on the estimation of wheat acreage. Yield and production feasibility studies were also carried out but the Accuracy Assessment team investigated only the accuracy of acreage estimation.

The initial CAS estimates, which were made for each month from April through August, were considerably higher than the USDA/SRS estimates. This was attributed to (1) the practice of considering bare ground as "potential wheat" and counting it as wheat, (2) overestimation of the wheat proportions in segments having only a small amount of wheat, and (3) the classification of confusion crops as wheat. At the end of the season most of the segments were reworked using improved methods based on experience gained during the season. In particular, new procedures were developed to solve the three problems listed above.

These and other improvements used in the rework experiment resulted in at-harvest estimates that were much closer to the USDA/SRS estimates than those obtained during the regular season. At the U.S. Great Plains (USGP) level the relative difference* was -11 percent. An approach was developed to evaluate whether the acreage results could support the 90/90 criterion. For this purpose it was assumed that the acreage and yield estimates were unbiased and independent, and that the coefficients of variation (CV) for acreage (CV_A) and for yield (CV_Y) were equal. If this were true, the 90/90 criterion applied at a given level** would be satisfied

---

*Relative difference is defined as $\frac{\text{LACIE-SRS}}{\text{LACIE}}$.

**In Phase I the 90/90 criterion was applied at the national level; in Phase II it was applied at the USGP level.

if CV_A for that level was less than 4.30 percent and if the acreage estimate was unbiased. In Phase I the estimate of CV_A at the national level was 3.74. Therefore, the 90/90 criterion would have been satisfied if the acreage estimate were unbiased. In fact some bias would be allowed, since 3.74 is somewhat smaller than 4.30. The relative differences between the LACIE and USDA/SRS estimates indicated that some bias was indeed present, but no accurate estimate of this bias was performed in Phase I; therefore, it is not possible to say whether or not the results satisfied the 90/90 criterion at the national level.

The area of most concern in Phase I was North Dakota, which had a relative difference of -74.6 percent. Blind site investigations indicated that the primary source of this problem was sampling error.

The experience gained in Phase I was used in developing the CAMS system for Phase II. Several changes were made on the basis of this experience. In particular, more sample segments were allocated to North Dakota, and the classification procedures developed for the CAMS rework experiment became the basis for the Phase II CAMS operations.

2.1.2 PHASE II

In Phase II, estimates were made for acreage, yield, and production. Generally the LACIE yield estimates were quite close to the USDA/SRS estimates and therefore can be considered satisfactory. However, the acreage and production estimates at the USGP level were low compared to the USDA/SRS estimates, due primarily to significant underestimates for spring wheat in the four U.S. northern Great Plains (USNGP) states and for winter wheat in Oklahoma.

For winter wheat in the USGP, the relative difference between the final LACIE production estimate and the USDA/SRS estimate was -7.2 percent. A significance test indicated that the LACIE estimate was not significantly different from the USDA/SRS estimate at the 10-percent level of significance. However, underestimation problems were still evident in Oklahoma. Investigations indicated that this underestimate was partially due to drought conditions and grazing of cattle which caused wheat signatures to differ significantly from those of normal wheat. In particular there was late "greening up" of the winter wheat crop, which caused the crop development to vary considerably from the crop calendar for "normal" winter wheat.

For spring wheat production, the relative difference between the final LACIE and USDA/SRS estimates for the USGP region was -22.3 percent. North Dakota had a relative difference of -6.6 percent, indicating that the sampling problems encountered with this state in Phase I largely had been solved. The major contributors to the spring wheat underestimate in Phase II were Minnesota (relative difference -89.6) and Montana (relative difference -67.4). The spring wheat proportions were obtained from small-grains proportion estimates produced by CAMS by using historical wheat/small-grains ratios. Spring wheat blind site investigation indicated that there was underestimation of the small-grains proportions in Minnesota and Montana. One of the major causes for this was that strip fallow fields were not classified well. Also, the blind site investigations indicated that sampling errors and incorrect estimates of wheat/small-grains ratios further contributed to the underestimation. (Several other reasons are discussed later, in section 4.2.2.2.)

For total wheat in the USGP, the relative difference between the final LACIE production estimate and the USDA estimate was -12.3 percent, a statistically significant difference. The LACIE estimate was evaluated in terms of the 90/90 criterion using an estimate for the relative bias in the LACIE production estimate; it was found that the 90/90 criterion was not met. The CV for production, estimated to be 5 percent, was sufficiently small for the 90/90 criterion to be satisfied if the production estimate had a relative bias whose absolute value was less than approximately 4 percent. However, the estimates obtained were much larger than this. Two methods of estimating the bias were used. One gave a bias of -24.0 percent which resulted in LACIE satisfying a 90/75 criterion (i.e., one was 90 percent confident that the LACIE estimate was within ±25 percent of the true wheat production of the USGP). The other method of estimating the bias gave a value of -12.3 percent which resulted in LACIE satisfying a 90/84 criterion. In both cases the large bias was due to acreage underestimation, particularly for spring wheat, and this problem will have to be solved for LACIE to meet its goals.

In Phase III, several steps have been taken to solve the problems outlined above. In particular, (1) new classification procedures have been instituted which hopefully will reduce the bias in the classification results, (2) the number of sample segments has been increased from 431 to 601, and (3) an effort will be made to estimate spring wheat directly instead of spring small grains and thereby avoid the error due to ratioing of wheat to small grains.

2.2 RECOMMENDATIONS

On the basis of the experience gained in Phase I and Phase II, the following recommendations are made.

a. Techniques shall be developed to avoid consistently underestimating spring wheat.

b. CAMS should develop procedures to solve the problem of underestimation in areas where there is drought, grazing, and late green-up such as occurred in Oklahoma.

c. Improved techniques should be developed for classifying strip-fallow fields.

d. The proportion error in CAMS estimates shows a striking dependence on the amount of wheat in the segment. Further attempts should be made to understand the cause of this effect.

e. More sample segments should be allocated to the state of Minnesota since wheat acreage in that state has increased considerably since the epoch year.

f. Accuracy Assessment should develop a data processing system to fully exploit the information in the blind-site ground truth.

3. PHASE I ACCURACY ASSESSMENT

LACIE Phase I AA investigations conducted during the 1975 crop year concentrated on assessing the accuracy of wheat acreage estimates.

3.1 COMPARISON OF LACIE AND USDA/SRS ACREAGE ESTIMATES

Three different data bases were used to generate acreage estimates in Phase I; the results obtained with these data bases are described in sections 3.1.1 through 3.1.3.

3.1.1 THE CAS 1A DATA BASE

The 1A data base contained all the sample segments processed by CAMS. It was used with the initial quasi-operational system to produce acreage estimates for April through August. This operation was concerned primarily with "debugging" the system. The results are shown in table 3-1.

The LACIE estimates for April through July are for winter wheat only. Thus, the estimates listed under "Mixed Wheat" for these months should not be compared with the corresponding USDA/SRS estimates, which include spring wheat. The LACIE estimates for August include spring wheat and therefore all can be compared with the USDA/SRS values.

It will be seen that there is a large positive bias relative to the USDA results for all months. The overestimates were attributed to the following causes:

a. Most of the Landsat data acquired early in the growing season were acquired before the wheat had emerged, since real-time crop calendars were not available to use for computing acquisition dates until May of 1975. This period in the growing season was called biowindow 1A and covered the period from

50-percent planted to dormancy. The 1A data base received this name because it included data from this period. Area estimates were attempted using these data by declaring areas of seed bed preparation (i.e., bare ground) as "potential wheat" and including them in the estimates. Since fall plowing is done for various reasons other than for planting wheat, this produced overestimates. Also, other bare soil categories (river bottoms, etc.) were confused with plowed ground. The biowindow 1A data represented the largest percentage by biowindow that was used in the April through July aggregations. It also influenced the August aggregation, but to a lesser extent.

b. There was a marked tendency to overestimate the proportion of wheat in Group II counties. This led to a thorough review of Group II aggregation in LACIE. It was determined that the Group II aggregation was satisfactory and that the problem was due to overestimation of sample segment proportions for segments having only a small amount of wheat. Most Group II segments fell into this category. Therefore, a new procedure, consisting of hand-counting all the wheat pixels for segments with a small amount of wheat, was instituted and was used in the CAMS rework procedure described below.

c. The classification of confusion crops as wheat also led to overestimates. This effect is particularly important in the spring and mixed wheat states where there are large quantities of other small grains which are difficult to distinguish from spring wheat. Each acquisition had an estimate for wheat alone and sometimes had an estimate for small grains (i.e., wheat plus confusion crops). If both were given, the small grains estimate was used.

In order to avoid the problems caused by the data from biowindow 1A, the 1B data base was formed.

3.1.2 AGGREGATIONS WITH THE 1B DATA BASE

The 1B data base was obtained by eliminating the data from biowindow 1A from the 1A data base. The remaining portion of biowindow 1 was called biowindow 1B and covered the period from dormancy to jointing. The 1B data base therefore consisted of all the data in the 1B biowindow plus all of the data for biowindows 2, 3, and 4.

Aggregations with the 1B data base were carried out for July and August. The results are given in table 3-1. In July the 1B estimates are all lower than the 1A estimates with the exception of those for Oklahoma. At the U.S. southern Great Plains (USSGP) level, the 1B estimate was 4.0×10^6 acres lower than the 1A estimate but was still 14.4×10^6 acres larger than the USDA/SRS estimate. At the USGP level, the 1B estimate was 12.3×10^6 acres lower than the 1A estimate but it cannot be compared with the USDA/SRS estimate since the latter includes spring wheat and the LACIE estimates for July do not.

In August, the differences between the estimates from the 1A and 1B data bases were smaller than in July. This was probably due to the smaller influence of biowindow 1 acquisitions for the 1A data base in August. In July, 106 acquisitions out of 232 were from biowindow 1; in August 87 out of 340 were from biowindow 1. The August estimates all can be compared with the USDA/SRS estimates. At the USSGP and USGP levels, the 1B estimates are slightly lower than the 1A estimates but are still much higher than the USDA/SRS estimates.

The improvements obtained from using the 1B data base were probably due mainly to a reduction in the amount of bare ground classified as wheat. However, bare ground was still classified as wheat in the 1B aggregations, and this probably accounted for a substantial part of the remaining overestimates. Also, factors b and c (section 3.1.1) are expected to have contributed to the 1B aggregations in the same way they did with the 1A aggregations.

3.1.3 THE CAMS REWORK EXPERIMENT

At the end of the season a new at-harvest estimate of wheat acreage was obtained by reworking the data using techniques based on experience acquired throughout the season. In particular:

a. Bare ground was not counted as wheat.

b. Acquisitions that appeared very difficult to interpret were not used.

c. All segments used had at least two acquisitions, of which one was biostage 2 or 3.

d. Multitemporal classification was used for selected segments.

e. CAMS gave estimates for small grains proportions for the spring wheat segments. These estimates were converted to estimates of spring wheat acreage by ratioing, using 1974 SRS statistics for spring wheat and small grains in the appropriate states.

f. The procedure of hand-counting pixels was used for classifying low wheat acreage segments. Usually, Group II segments fell into this category.

Two at-harvest estimates were made using the CAMS rework data. These two estimates differed only in regard to the inclusion of Group II segments. The results for both cases are shown in table 3-2. As can be seen, the area estimates are significantly better when the Group II segments are used in the aggregation.

In Phase I, the 90/90 criterion was applied at the national level. An approximate relation was derived which expressed the CV of production (CV_P) in terms of the CV of the area estimate (CV_A)

and the CV of the yield estimate (CV_Y), namely

$$(CV_P)^2 = (CV_A)^2 + (CV_Y)^2 + (CV_A \times CV_Y)^2.$$

If one further assumes $CV_A = CV_Y$, then the 90/90 criterion could be satisfied if $CV_A = CV_Y \leq 4.30$ percent.

It will be seen from table 3-2 that the CV for acreage projected to the national level was 3.74. Since this percentage was smaller than 4.30, it was possible to satisfy the 90/90 criterion even if there was a small amount of bias. However, since there was no ground truth available in Phase I, no estimate was made of the bias, and therefore it is not possible to say whether the results satisfied the 90/90 criterion.

An evaluation of the Phase I 90/90 criterion using production estimates was given in the LACIE Phase I Evaluation Report but is not reported here since in Phase I, AA evaluated acreage estimation only.

From the results presented in table 3-2, the area of most concern was North Dakota. More detailed error analysis based on ground truth and ancillary data in Kansas, North Dakota, Nebraska, and South Dakota permitted a further assessment of the sampling and classification errors. These analyses, discussed in section 3.2, indicated the major source of the North Dakota problem to be sampling error.

After the regular CAMS rework estimates given in table 3-2 were made, there was a revision of the area in the pseudo counties (i.e., the part of the counties that is classified as agricultural as distinguished from nonagricultural). This caused a change in the estimates and CV's. The revised results are presented in table 3-3. Note that in most cases the CV's are smaller.

3.2 ESTIMATION OF AREA ERROR USING BLIND SITE DATA

The expression "blind site" is merely a designation applied to selected operational segments for which, unknown to the analyst, ground truth data were acquired for evaluation purposes. The implementation of this approach occurred late in the growing season of LACIE Phase I. Thus, all of the selected sites were in the northern spring wheat regions.

High-resolution color infrared aerial photography over 29 LACIE segments in North Dakota and Montana was acquired in mid-August 1975. (The results from only 16 of these segments in North Dakota are relevant to the basic discussion which follows.) Simultaneously, field teams were collecting ground information for a substantial portion of these segments. These data were combined to obtain both field and total segment ground truth data. The small grain proportion estimates were compared statistically to the LACIE estimates for the 16 segments in North Dakota. This resulted in a direct computation of the classification error, CV_C, for segments in the state of North Dakota, as listed in table 3-4.

This table indicates a relative difference of -18 percent between the average LACIE proportion and the average ground-observed proportion. This is not indicative of a significant bias in view of the standard error. However, the difference between the ground-observed proportions and the SRS county proportions is commensurate with the underestimate obtained in North Dakota. Thus, for North Dakota it was concluded that sampling error resulting from nonrepresentative sample segments was the major source of the observed bias. Other investigations with full frame imagery confirmed that agriculture is very heterogeneous in this region and many of the LACIE segments did not adequately represent their county.

TABLE 3-4.— LACIE BLIND SITE DATA

[North Dakota spring small grains]

County	Fraction of area in small grains, percent		
	Ground truth (5x6 n. mi. segment)	LACIE (5x6 n. mi. segment)	SRS county (entire county)
Ward 1	13.2	17.1	33.8
Ward 2	26.8	8.2	33.8
Williams	3.7	0.0	27.5
McHenry 1	0.0	0.0	25.9
McHenry 2	0.3	0.0	25.9
Rolette	4.9	---	18.8
Ramsey	38.4	49.5	41.5
McKenzie 1	1.3	---	10.6
McKenzie 2	1.0	0.3	10.6
McLean	29.3	28.4	31.7
Mercer	16.3	18.0	19.9
Oliver	15.6	---	16.2
Kidder	16.4	---	19.4
Sheridan	12.9	0.0	30.9
Adams	26.1	24.4	22.8
Hettinger	21.7	24.1	35.7
Burleigh	18.2	12.0	20.7
Morton	4.6	6.7	15.7
Richland	31.6	15.6	36.2
Sargent	35.0	32.3	34.7
	17.46 LACIE 16	14.78	---
Average	15.87 ALL 20	---	26.00

Variance of LACIE estimates is within allowable range, CV = 50 percent.
No apparent bias in LACIE estimate.

3.3 RESULTS OF PHASE I

Phase I comparisons of LACIE wheat acreage estimates with ground truth indicated that the LACIE classification technology was working fairly well and may have been adequate to support the 90/90 criterion applied at the national level. However, a definitive answer to the question of whether the 90/90 criterion was satisfied at the national level would require an estimate of the bias in the acreage estimate, which was not done in Phase I. The experience gained in Phase I was valuable in developing the system for Phase II. Several changes were made on the basis of this experience. In particular, more segments were allocated to North Dakota, and the classification procedures developed for the CAMS rework experiment became a basis for the Phase II CAMS operations.

4. PHASE II ACCURACY ASSESSMENT

In Phase II, LACIE produced operational estimates for acreage, yield, and production. Each of these is discussed below in a separate section.

4.1 ASSESSMENT OF PRODUCTION ESTIMATION

This section consists of three parts: an assessment of how well LACIE met the 90/90 criterion (section 4.1.1), a comparison of LACIE and USDA/SRS wheat production estimates (section 4.1.2), and an investigation of the contribution of the first-order error sources to the production CV (section 4.1.3)

4.1.1 THE 90/90 CRITERION

The LACIE accuracy goal for the USGP region is a 90/90 at-harvest criterion for wheat production. This specifies that for any given year the probability shall be at least 0.90 that the at-harvest wheat production estimate for the USGP will be within 10 percent of the true production.

Let $\hat{P}$ be the LACIE at-harvest estimate of wheat production for the USGP and let P be the true wheat production for the USGP. Then the 90/90 criterion may be expressed by the following probability statement:

$$\Pr[|\hat{P} - P| \leq 0.1P] \geq 0.90 \qquad (4\text{-}1)$$

It is reasonable to assume for large sample sizes that $\hat{P}$ is normally distributed with mean P + B and variance $\sigma^2_{\hat{P}}$, where B is the bias of the estimator, $\hat{P}$. Under this assumption, it is shown in appendix A (section A.3.3.4) that equation (4-1) is satisfied for a range of values of the relative bias of $\hat{P}$, $\frac{B}{(P + B)}$, and the coefficient of variation of the estimator $\hat{P}$,

$$CV(\hat{P}) = \frac{\sigma_{\hat{P}}}{E(\hat{P})} = \frac{\sigma_{\hat{P}}}{(P + B)}.$$

Inference as to whether the LACIE accuracy goal has been met is made by estimating $\frac{B}{P + B}$ and $CV(\hat{P})$ and then ascertaining whether these values fall in the range which satisfies equation (4-1). Now, $CV(\hat{P})$ is estimated by $\frac{\hat{\sigma}_{\hat{P}}}{\hat{P}}$ where $\hat{\sigma}_{\hat{P}}$ is an estimate of the standard deviation of $\hat{P}$, and $\hat{P}$ is an unbiased estimate of P + B. If the true wheat production for the USGP were known, then $\frac{B}{P + B}$ could be estimated simply by $\frac{\hat{P} - P}{\hat{P}}$. However, P is unknown so the relative bias in the production estimate must be estimated by some other method.

One such method is described in appendix A (section A.3.3.3). This leads to an estimate of -24.0 percent for the relative bias. The 90-percent confidence limits for the bias in the production estimate, expressed as a percentage of the LACIE production estimate, are given by (-32.0, -16.6). From figure A-1 in appendix A it can be seen that the 90/90 accuracy goal cannot be achieved for any value of the relative bias within these confidence limits.

It can be shown, however, that an accuracy of 90/75 is achievable with a relative bias of -24.0 percent and a CV of 5.0 percent. That is, the probability that the LACIE estimate is within ±25 percent of the true wheat production for the USGP is 0.9.

A second method of estimating the relative bias is to estimate it by (LACIE - SRS_F)/LACIE, where SRS_F is the final SRS estimate and LACIE is the LACIE estimate for a given month. Then, for the data given in the August, September, October, and final reports, LACIE satisfies the following criteria:

a. August – 90/78

b. September – 90/83

c. October – 90/83

d. Final – 90/84

4.1.2 COMPARISON OF LACIE AND USDA/SRS PRODUCTION ESTIMATES

These comparisons are designed to monitor how well LACIE is performing relative to the USDA/SRS estimates, and also to detect any problems that may exist.

The LACIE and USDA/SRS production estimates are shown in figure 4-1 and table 4-1. In table 4-1, estimates are given for each state in the nine-state USGP region and for the following regions:

a. The USSGP region consisting of Colorado, Kansas, Nebraska, Oklahoma, and Texas. LACIE makes only winter wheat estimates for these states. The estimates are available for February through October.

b. The spring wheat (SW) states of Minnesota and North Dakota. These states have very little winter wheat so LACIE makes estimates for spring wheat only. The estimates are available for August through October.

c. The mixed wheat (MW) states of Montana and South Dakota. These states have both spring and winter wheat. LACIE estimates of wheat production are available from August through October for spring wheat and from June through October for winter wheat.

d. The U.S. northern Great Plains (USNGP) region made up of the two spring wheat states and the two mixed wheat states.

e. The USGP region made up of the nine states of the USSGP and the USNGP.

In the following discussion winter wheat is considered first, followed by spring wheat, then total wheat (winter wheat plus spring wheat). Figure 4-1 and table 4-1 are arranged in this order.

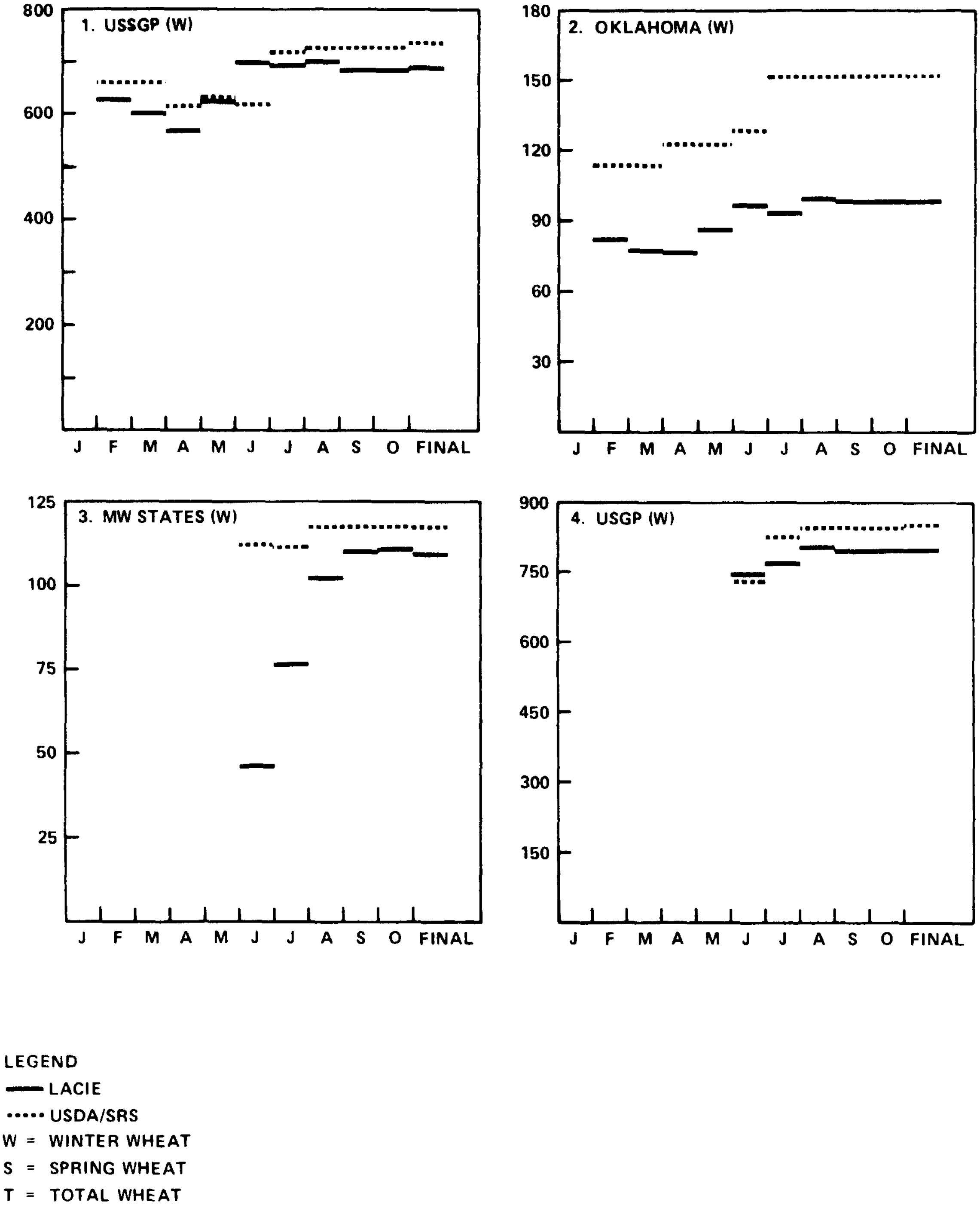


Figure 4-1.– LACIE and USDA/SRS production estimates [bushels × 10^6].

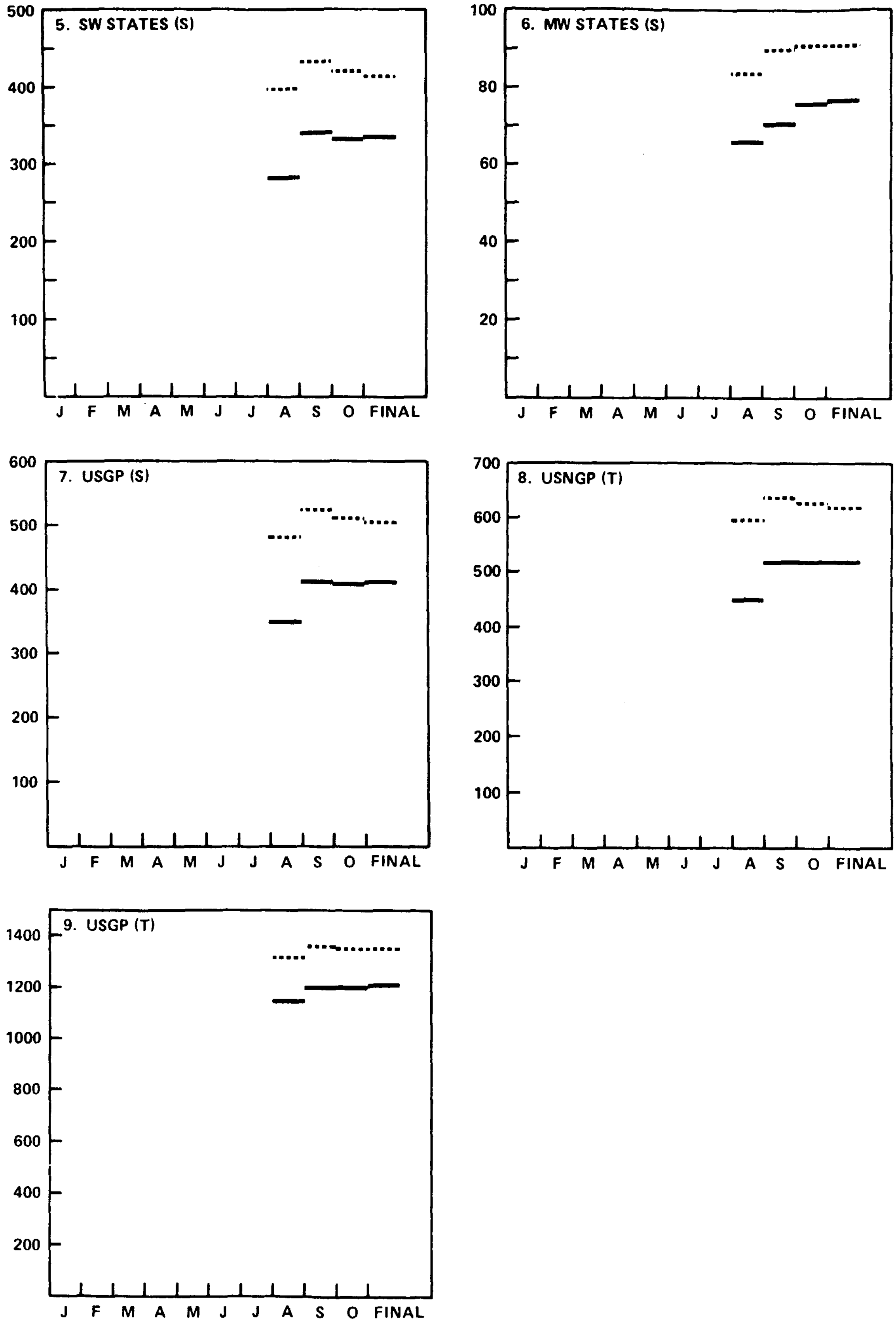


Figure 4-1.— Concluded.

TABLE 4-1.– COMPARISON OF USDA/SRS AND LACIE PRODUCTION ESTIMATES

[Bushels × 10^3]

Region	USDA/SRS (a)	LACIE	Relative difference (%)	CV (%)	Test statistic
February					
Winter Wheat					
Colorado	48 110	76 418	37.0	33	
Kansas	327 500	258 074	-26.9	17	
Nebraska	92 200	151 762	39.2	23	
Oklahoma	113 250	80 264	-41.1	29	
Texas	75 600	59 550	-26.9	28	
[b]USSGP	656 660	626 068	- 4.9	11	-.45[N]
March					
Winter Wheat					
Colorado	48 110	60 759	20.8	32	
Kansas	327 500	269 638	-21.5	14	
Nebraska	92 200	124 342	25.8	19	
Oklahoma	113 250	76 041	-48.9	25	
Texas	75 600	66 676	-13.4	32	
[b]USSGP	656 660	597 456	- 9.9	10	-.90[N]

[a]The USDA/SRS estimates for February and March are the December 1, 1975 estimates.

[b]The five-state USSGP region.

[N]The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-1.– Continued.

Region	USDA/SRS	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
April					
Winter Wheat					
Colorado	42 840	56 089	23.6	32	
Kansas	286 000	255 147	-12.1	13	
Nebraska	95 200	118 458	19.6	19	
Oklahoma	121 800	74 823	-62.8	22	
Texas	66 300	59 559	-11.3	22	
USSGP	612 140	564 076	- 8.5	8	-1.06^{N}
May					
Winter Wheat					
Colorado	41 800	55 285	24.4	31	
Kansas	302 400	283 124	- 6.8	12	
Nebraska	94 400	110 496	14.6	19	
Oklahoma	121 800	84 699	-43.8	21	
Texas	70 200	86 910	19.2	17	
USSGP	630 600	620 514	- 1.6	8	-0.2^{N}

TABLE 4-1.– Continued.

Region	USDA/SRS	LACIE	Relative difference (%)	CV (%)	Test statistics
June					
Winter Wheat					
Colorado	41 800	61 191	31.7	28	
Kansas	279 500	326 677	14.4	11	
Nebraska	97 350	128 692	24.4	17	
Oklahoma	127 600	94 975	-34.4	17	
Texas	70 200	84 094	16.5	17	
USSGP	616 450	695 629	11.4	7	1.63*
Montana	90 600	13 527	-569.8	192	
S. Dakota	20 800	31 553	34.1	46	
[c] MW states	111 400	45 080	-147.1	63	
[d] USGP	727 850	740 709	1.7	8	.21[N]
July					
Winter Wheat					
Colorado	48 400	51 492	6.0	30	
Kansas	321 900	334 107	3.7	11	
Nebraska	96 000	132 118	27.3	16	
Oklahoma	151 200	92 052	-64.3	18	
Texas	98 700	80 797	-22.2	17	
USSGP	716 200	690 566	- 3.7	7	.53[N]
Montana	93 620	30 082	-211.2	53	
S. Dakota	16 640	45 096	63.1	27	
MW states	110 260	75 178	-46.7	27	
USGP	826 460	765 744	- 7.9	7	-1.13[N]

*The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

[c]The mixed wheat states, Montana and South Dakota.

[d]The nine-state United States Great Plains region.

TABLE 4-1.– Continued.

Region	USDA/SRS	LACIE	Relative difference (%)	CV (%)	Test statistic
August					
Winter Wheat					
Colorado	48 400	50 024	3.2	29	
Kansas	327 450	338 078	3.1	10	
Nebraska	96 000	130 547	26.5	16	
Oklahoma	151 200	98 156	-54.0	18	
Texas	103 400	80 637	-28.2	18	
USSGP	726 450	697 442	-4.2	7	.60[N]
Montana	96 640	55 788	-73.2	36	
S. Dakota	19 760	45 096	56.2	26	
MW states	116 400	100 884	-15.4	23	
USGP	842 850	798 326	-5.6	7	-.80[N]
Spring Wheat					
Minnesota	122 518	55 490	-120.8	42	
N. Dakota	272 700	226 034	-20.6	17	
[e] SW states	395 218	281 524	-40.4	16	
Montana	63 095	29 188	-116.2	29	
S. Dakota	20 350	36 719	44.6	18	
MW states	83 409	65 907	-26.6	17	
USGP	478 663	347 431	-37.8	13	-2.91*
[f] Total Wheat					
Montana	159 735	84 976	-88.0	20	
S. Dakota	40 110	81 815	51.0	14	
MW states	199 845	166 791	-19.8	12	
[g] USNGP	595 063	448 315	-32.7	11	-2.97*
USGP	1 321 513	1 145 757	-15.3	6	-2.55*

[e] The spring wheat states, Minnesota and North Dakota.

[f] Spring wheat plus winter wheat.

[g] The four-state United States northern Great Plains region.

TABLE 4-1.– Continued.

Region	USDA/SRS	LACIE	Relative difference (%)	CV (%)	Test statistic
September					
Winter Wheat					
Colorado	48 400	52 924	8.5	29	
Kansas	327 450	339 974	3.7	10	
Nebraska	96 00	110 972	13.5	16	
Oklahoma	151 200	96 491	-56.7	18	
Texas	103 400	81 312	-27.2	18	
USSGP	726 450	681 673	-6.6	7	$-.94^{N}$
Montana	96 640	62 877	-53.7	30	
S. Dakota	19 760	45 904	57.0	26	
MW states	116 400	108 781	-7.0	21	
USGP	842 850	790 454	-6.6	7	$-.94^{N}$
Spring Wheat					
Minnesota	130 256	77 230	-68.7	29	
N. Dakota	300 040	261 197	-14.9	12	
SW states	430 296	338 427	-27.1	11	
Montana	65 410	35 064	-86.5	25	
S. Dakota	24 300	35 908	32.3	19	
MW states	89 710	70 972	-26.4	15	
USGP	520 006	409 399	-27.0	10	-2.70*
Total Wheat					
Montana	162 050	97 941	-65.5	15	
S. Dakota	44 060	81 812	46.1	13	
MW states	206 110	179 753	-14.7	10	
USNGP	636 406	518 180	-22.8	10	-2.28*
USGP	1 362 856	1 199 853	-13.6	5	-2.72*

TABLE 4-1.— Continued.

Region	USDA/SRS	LACIE	Relative difference (%)	CV (%)	Test statistic
October					
Winter Wheat					
Colorado	48 400	52 924	8.5	29	
Kansas	327 450	339 974	3.7	10	
Nebraska	96 000	110 972	13.5	16	
Oklahoma	151 200	96 491	-56.7	18	
Texas	103 400	81 312	-27.2	18	
USSGP	726 450	681 673	-6.6	7	$-.94^{N}$
Montana	96 640	63 758	-51.6	29	
S. Dakota	19 760	45 904	57.0	26	
MW states	116 400	109 662	-6.1	20	
USGP	842 850	791 335	-6.5	7	$-.94^{N}$
Spring Wheat					
Minnesota	126 344	66 589	-89.7	32	
N. Dakota	290 320	263 703	-10.1	12	
SW states	416 664	330 292	-26.2	11	
Montana	66 658	40 240	-65.7	25.	
S. Dakota	24 300	35 675	31.9	18	
MW states	90 958	75 915	-19.8	.16	
USGP	507 532	406 207	-24.9	10	-2.49*
Total Wheat					
Montana	163 208	103 998	-56.9	13	
S. Dakota	44 060	81 579	46.0	13	
MW states	207 268	185 577	-11.7	9	
USNGP	623 932	515 869	-20.9	8	-2.61*
USGP	1 350 382	1 197 542	-12.8	5	-2.56*

TABLE 4-1.– Concluded.

Region	USDA/SRS	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
Final					
Winter Wheat					
Colorado	47 300	52 924	10.6	29	
Kansas	339 000	344 472	1.6	10	
Nebraska	94 400	110 972	14.9	16	
Oklahoma	151 200	96 491	-56.7	18	
Texas	103 400	81 312	-27.2	18	
USSGP	735 300	686 171	-7.2	7	-1.03^N
Montana	98 560	62 167	-58.5	30	
S. Dakota	17 460	45 904	62.0	26	
MW states	116 020	108 071	-7.4	20	
USGP	851 320	794 242	-7.2	7	-1.03^N
Spring Wheat					
Minnesota	126 244	66 589	-89.6	32	
N. Dakota	284 050	266 529	-6.6	12	
SW states	410 294	333 118	-23.2	11	
Montana	68 735	41 058	-67.4	24	
S. Dakota	22 060	35 675	38.2	18	
MW states	90 795	76 733	-18.3	15	
USGP	501 089	409 851	-22.3	10	-2.23*
Total Wheat					
Montana	167 295	103 225	-62.1	13	
S. Dakota	39 520	81 579	51.6	13	
MW states	206 815	184 804	-11.9	9	
USNGP	617 109	517 922	-19.2	8	-2.40*
USGP	1 352 409	1 204 093	-12.3	5	-2.46*

The CV's in table 4-1 were computed by the methods described in appendix A (section A.3.3.2). For the major regions, a significance test was performed to determine if the LACIE estimate was significantly different from the USDA/SRS estimate. The test statistic is given in the last column of table 4-1 and the method is described in appendix A (section A.2).

Winter Wheat Production

Plots 1 through 4 in figure 4-1 show the production estimates for winter wheat. Plot 1 shows that the LACIE estimates for the USSGP region were lower than the USDA/SRS estimates for every month except June; they were lower than the USDA/SRS final estimate for every month including June. The LACIE estimate was particularly low in April, due mainly to low acreage estimates in Kansas, Oklahoma and Texas, which were affected by drought (see section 4.2.2.1). However, the LACIE estimate improved considerably in May and again in June. The June LACIE estimate was considerably better than the June USDA/SRS estimate relative to the final USDA/SRS estimate. The final LACIE estimate had a relative difference of -7.2 percent. The significance test showed that the LACIE estimate was not significantly different from the USDA/SRS estimate for any month except June. In this case it was the USDA/SRS estimate that was low (relative to the final USDA/SRS estimate).

The most serious problem in the USSGP region was in Oklahoma (plot 2), where the wheat production was consistently underestimated throughout the season due to underestimates of wheat acreage. Also, Montana was underestimated by a wide margin, primarily due to underestimation of acreage, and South Dakota was overestimated by a wide margin due to overestimation of both acreage and yield.

The production estimates for winter wheat in the two mixed wheat states are shown in plot 3. They were very low in June but increased throughout the season and had a relative difference of -7.4 percent for the final estimate.

Plot 4 shows the estimates for the total winter wheat in the USGP region. The relative difference for the final estimate was -7.2 percent. The LACIE estimate was not significantly different from the USDA/SRS estimate for any month or for the final estimate.

Spring Wheat Production

Plots 5 through 7 show the estimates for spring wheat production. The LACIE estimates were consistently low in the spring wheat states, the mixed wheat states, and the overall USNGP. The significance tests show that the LACIE estimates for the USNGP region were significantly different from the USDA/SRS estimate for every month and for the final estimate. These underestimates in production were due to underestimates of spring wheat acreage, since the yields were overestimated by LACIE except in September when they were slightly less than the USDA/SRS estimate. (See plot 7 in figure 4-2.) This tendency to underestimate spring wheat acreage is discussed further in section 4.2.2.2. Looking at the individual states, the largest underestimates occurred in Minnesota and Montana. In both cases the problem was primarily due to underestimates in acreage. In South Dakota there was a large overestimate due to overestimation of the yield.

Total Wheat Production

Plot 8 shows the total wheat in the four-state USNGP region. It was consistently underestimated and the LACIE estimate was significantly different from the USDA/SRS estimate for every month and for the final estimate.

The wheat production estimates for the nine-state USGP region are shown in plot 9. The LACIE estimate was consistently low. The final estimate had a relative difference of -12.3 percent due to an underestimate of 57×10^6 bushels (relative difference of -7.2 percent) in the winter wheat crop and an underestimate of 91×10^6 bushels (relative difference -22.3 percent) in the spring wheat crop. The LACIE estimate was significantly different from the USDA/SRS estimate for every month and for the final estimate.

4.1.3 FIRST-ORDER PRODUCTION ERROR COMPONENTS

The first-order production error components consist of yield prediction error and acreage estimation error. Acreage estimation error is further subdivided into sampling error and classification error. The effect of each error component on production is assessed by determining the reduction in the estimate for the CV of production when this error component is set equal to zero. Details of the method employed are given in appendix A (section A.3.3.5).

Table 4-2 shows the results for the CV's of the Phase II final estimates when acreage and yield errors are omitted. It will be seen that omitting the yield error leads to larger reductions in

TABLE 4-2.— REDUCTIONS IN THE PRODUCTION CV CAUSED BY OMITTING VARIOUS ERRORS

Region	Total CV, %	Yield error omitted		Acreage error omitted		Classification error omitted		Sampling error omitted	
		CV, %	Reduction, %	CV, %	Reduction, %	CV, %	Reduction, %	CV, %	Reduction, %
Winter Wheat USSGP	7.0	4.5	35.7	5.3	24.3	6.5	7.1	5.9	15.7
Spring Wheat USNGP	10.0	6.3	37.0	7.5	25.0				
Total Wheat USGP	5.2	3.7	28.8	4.4	15.4				

the CV for all three regions listed. This indicates that the yield error has a more dominant effect than the acreage error on the production CV.

Table 4-2 also shows the results when sampling and classification errors are omitted. The estimates of classification and sampling errors are presented in section 4.2.3. The spring wheat regions were not included due to the small number of blind sites available for estimating these errors. The results indicate that sampling contributes slightly more than classification to the production CV. However, it is reasonable to believe that the sampling and classification errors contribute about equally to the production CV, since the difference between the two fractional reduction rates is rather small and may well be statistically insignificant.

4.2 ASSESSMENT OF ACREAGE ESTIMATION

This section contains three major subsections: a comparison of LACIE and USDA/SRS wheat acreage estimates (section 4.2.1), a discussion of classification error (section 4.2.2), and a discussion of the variance of sampling and classification error (section 4.2.3).

4.2.1 COMPARISON OF LACIE AND USDA/SRS ACREAGE ESTIMATES

The USDA/SRS and LACIE acreage estimates are shown in figure 4-2 and table 4-3. These are in the same format as table 4-1 and figure 4-1 except that the estimates are for acreage rather than production.

Winter Wheat

Plots 1 through 4 in figure 4-2 show the acreage estimates for winter wheat.

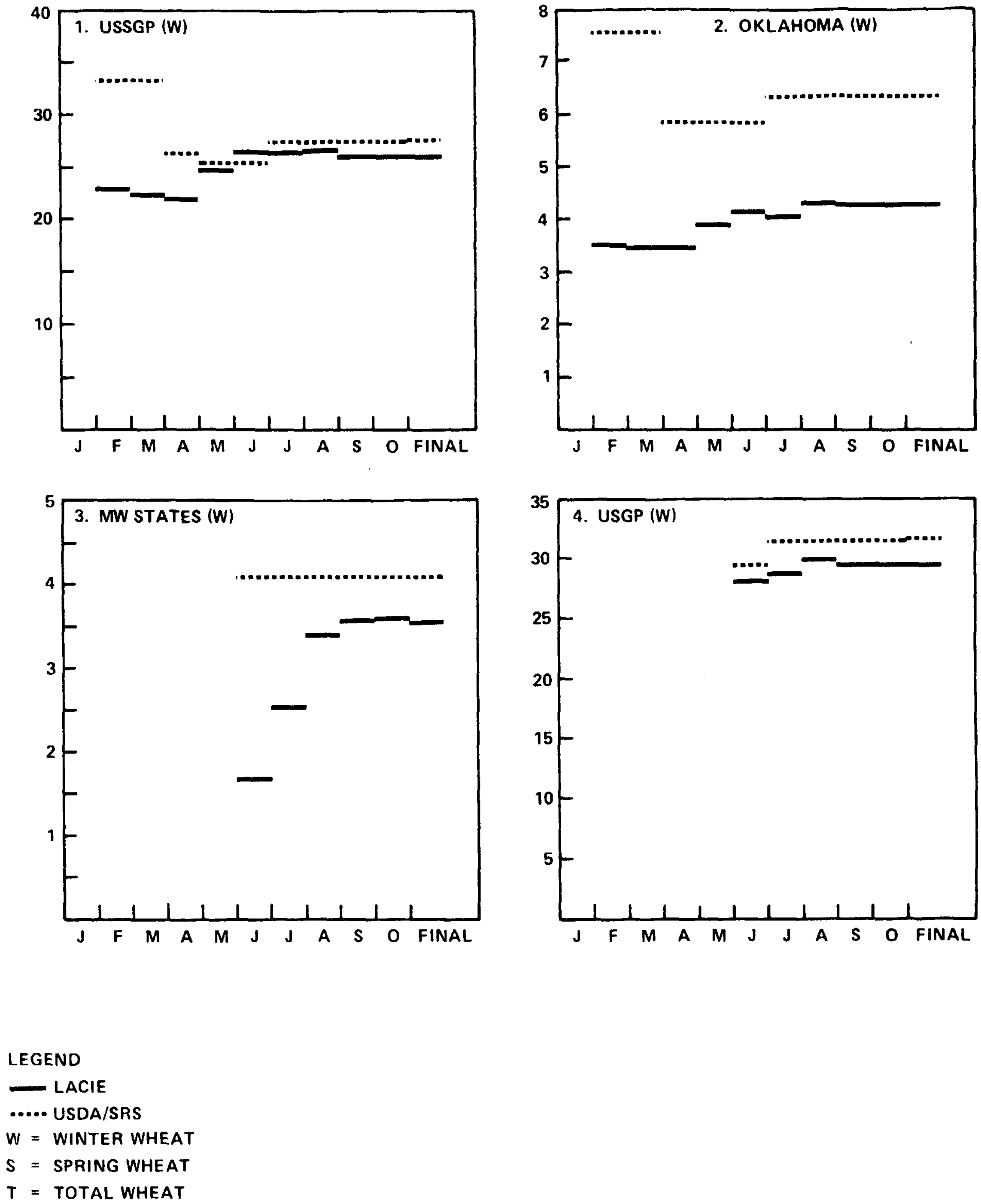


Figure 4-2.– LACIE and USDA/SRS acreage estimates [acres × 10^6].

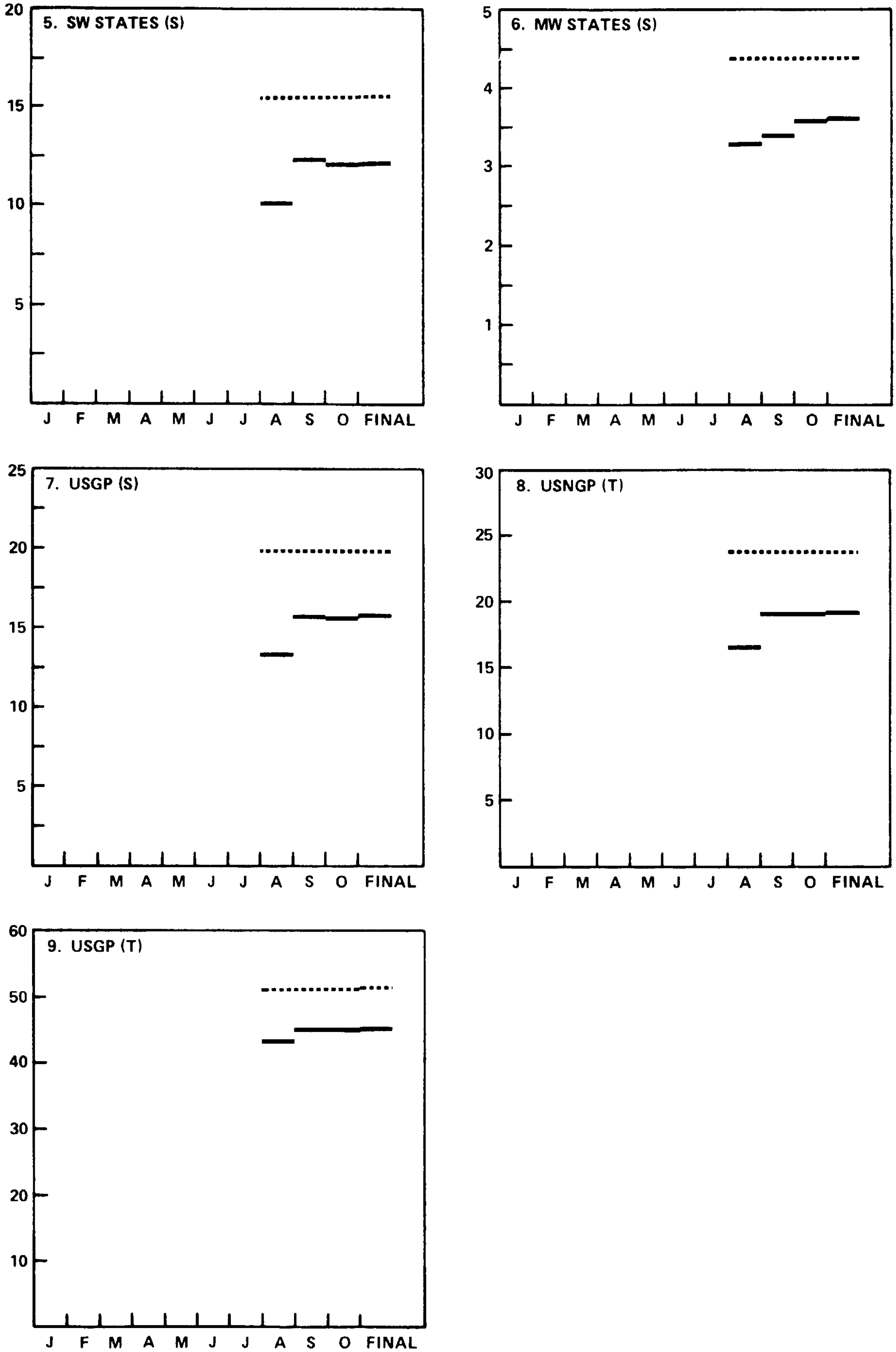


Figure 4-2.– Concluded.

TABLE 4-3.— COMPARISON OF USDA/SRS AND LACIE ACREAGE ESTIMATES

[Acres × 10^3]

Region	n/M (a)	USDA/ SRS (b)	LACIE	Relative difference (%)	CV (%)	Test statistic
February						
Winter Wheat						
Colorado	13/32	2 830	3 539	20.0	26	
Kansas	43/84	13 100	8 013	-63.5	12	
Nebraska	13/35	3 400	4 500	24.4	18	
Oklahoma	30/40	7 550	3 499	-90.0	24	
Texas	31/49	6 300	3 170	-98.7	25	
USSGP	130/240	33 180	22 721	-46.0	9	-5.11*
March						
Winter Wheat						
Colorado	25/32	2 830	2 768	-2.2	25	
Kansas	61/84	13 100	8 536	-53.5	8	
Nebraska	21/35	3 400	3 632	6.4	13	
Oklahoma	36/40	7 550	3 450	-118.8	18	
Texas	42/49	6 300	3 725	-69.1	30	
USSGP	185/240	33 180	22 111	-50.1	8	-6.26*

[a] n is the number of segments used; M is the number of segments allocated.

[b] The USDA/SRS estimates for February and March are the December, 1975, estimates of seeded acreage.

*The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-3.– Continued.

Region	n/M (a)	USDA/ SRS	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
April						
Winter Wheat						
Colorado	25/32	2 040	2 768	26.3	25	
Kansas	62/84	11 000	8 536	-28.9	8	
Nebraska	22/35	3 400	3 583	5.1	13	
Oklahoma	36/40	5 800	3 450	-68.1	18	
Texas	44/49	3 900	3 479	-12.1	20	
[c]USSGP	189/240	26 140	21 816	-19.8	7	-2.82*
May						
Winter Wheat						
Colorado	26/32	1 900	2 807	32.3	24	
Kansas	70/84	10 800	9 392	-15.0	6	
Nebraska	27/35	2 950	3 653	19.2	13	
Oklahoma	38/40	5 800	3 897	-48.8	16	
Texas	47/49	3 900	4 810	18.9	14	
[c]USSGP	208/240	25 350	24 559	-3.2	6	-.53[N]

[a]n is the number of segments used; M is the number of segments allocated.

[c]The five-state U.S. southern Great Plains region.

[N]The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

*The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-3.– Continued.

Region	n/M (a)	USDA/ SRS	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
June						
Winter Wheat						
Colorado	26/32	1 900	2 995	36.6	23	
Kansas	75/84	10 750	10 535	-2.0	6	
Nebraska	30/35	2 950	4 104	28.1	12	
Oklahoma	38/40	5 300	4 148	-39.8	14	
Texas	47/49	3 900	4 556	14.4	15	
USSGP	216/240	25 300	26 338	3.9	5	-.78[N]
Montana	10/38	3 020	488	-518.9	193	
S. Dakota	8/10	1 040	1 159	10.3	43	
[d]MW states	18/48	4 060	1 647	-146.5	65	
[e]USGP	234/288	29 360	27 985	-4.9	6	-.81[N]
July						
Winter Wheat						
Colorado	30/32	2 200	2 867	23.3	25	
Kansas	78/84	11 100	10 795	-2.8	6	
Nebraska	32/35	3 000	4 133	27.4	11	
Oklahoma	40/40	6 300	4 025	-56.5	15	
Texas	47/49	4 700	4 314	-8.9	15	
USSGP	227/240	27 300	26 134	-4.5	5	-.09[N]
Montana	21/38	3 020	1 044	-189.3	52	
S. Dakota	9/10	1 040	1 482	29.8	23	
MW states	30/48	4 060	2 526	-60.7	25	
USGP	257/288	31 360	28 660	-9.4	5	-1.88*

[a]n is the number of segments used; M is the number of segments allocated.

[d]The mixed wheat states, Montana and South Dakota.

[e]The nine-state U.S. Great Plains region.

[N]The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

*The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-3.– Continued.

Region	n/M (a)	USDA/ SRS	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
August						
Winter Wheat						
Colorado	31/32	2 200	2 830	22.3	24	
Kansas	78/84	11 100	10 932	-1.5	5	
Nebraska	32/35	3 000	4 086	26.6	11	
Oklahoma	40/40	6 300	4 305	-46.3	15	
Texas	47/49	4 700	4 310	-9.0	16	
USSGP	228/240	27 300	26 463	-3.2	5	-.64[N]
Montana	22/38	3 020	1 911	-58.0	35	
S. Dakota	9/10	1 040	1 482	29.8	23	
MW states	31/48	4 060	3 393	-19.7	22	
USGP	259/288	31 360	29 856	-5.0	5	-1.00[N]
Spring Wheat						
Minnesota	10/13	3 826	1 741	-119.8	40	
N. Dakota	31/85	11 540	8 161	-41.4	14	
[f] SW states	41/98	15 366	9 902	-55.2	13	
Montana	14/22	2 315	1 127	-105.4	28	
S. Dakota	14/23	2 050	2 169	5.5	12	
MW states	28/45	4 365	3 296	-32.4	12	
USGP	69/143	19 731	13 198	-49.5	10	-4.95*
[g]Total Wheat						
Montana	36/60	5 335	3 038	-75.6	19	
S. Dakota	23/33	3 090	3 651	15.4	13	
MW states	59/93	8 425	6 689	-26.0	11	
[h] USNGP	100/191	23 791	16 591	-43.4	9	-4.82*
USGP	328/431	51 091	43 054	-18.7	5	-3.74*

[a]n is the segment used; M is the number of segments allocated.

[f]The spring wheat states, Minnesota and North Dakota.

[g]Spring wheat plus winter wheat.

[h]The four-state U.S. northern Great Plains region.

[N]The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

*The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-3.– Continued.

Region	n/M (a)	USDA/ SRS	LACIE	Relative difference (%)	CV (%)	Test sta-tistic
September						
Winter Wheat						
Colorado	32/32	2 200	2 704	18.6	24	
Kansas	81/84	11 100	10 989	-1.0	5	
Nebraska	33/35	3 000	3 399	11.7	11	
Oklahoma	40/40	6 300	4 261	-47.9	14	
Texas	47/49	4 700	4 344	-8.2	16	
USSGP	233/240	27 300	25 697	-6.2	5	-.39[N]
Montana	35/38	3 020	2 103	-43.6	29	
S. Dakota	9/10	1 040	1 452	28.4	23	
MW states	44/48	4 060	3 555	-14.2	20	
USGP	277/288	31 360	29 252	-7.2	5	-1.44[N]
Spring Wheat						
Minnesota	10/13	3 826	2 551	-50.0	27	
N. Dakota	67/85	11 540	9 650	-19.6	5	
SW states	77/98	15 366	12 201	-25.9	7	
Montana	19/22	2 315	1 291	-79.3	23	
S. Dakota	18/23	2 050	2 095	2.1	13	
MW states	37/45	4 365	3 386	-28.9	12	
USGP	114/143	19 731	15 587	-26.6	6	-4.43*
Total Wheat						
Montana	54/60	5 335	3 394	-57.2	14	
S. Dakota	27/33	3 090	3 547	12.9	12	
MW states	81/93	8 425	6 941	-21.4	9	
USNGP	158/191	23 791	19 142	-24.3	6	-4.05*
USGP	391/431	51 091	44 839	-13.9	4	-3.48*

[a] n is the segment used; M is the number of segments allocated.

[N] The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

* The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-3.— Continued.

Region	n/M (a)	USDA/ SRS	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
October						
Winter Wheat						
Colorado	32/32	2 200	2 704	18.6	24	
Kansas	81/84	11 100	10 989	-1.0	5	
Nebraska	33/35	3 000	3 399	11.7	11	
Oklahoma	40/40	6 300	4 261	-47.9	14	
Texas	47/49	4 700	4 344	-8.2	16	
USSGP	233/240	27 300	25 697	-6.2	5	-1.24[N]
Montana	36/38	3 020	2 131	-41.7	28	
S. Dakota	9/10	1 040	1 452	28.4	23	
MW states	45/48	4 060	3 583	-13.3	19	
USGP	278/288	31 360	29 280	-7.1	5	-1.42[N]
Spring Wheat						
Minnesota	11/13	3 826	2 198	-74.1	30	
N. Dakota	79/85	11 540	9 735	-18.5	5	
SW states	90/98	15 366	11 933	-28.8	7	
Montana	20/22	2 315	1 487	-55.7	24	
S. Dakota	19/23	2 050	2 079	1.4	13	
MW states	39/45	4 365	3 566	-22.4	12	
USGP	129/143	19 731	15 499	-27.3	6	-4.55*
Total Wheat						
Montana	56/60	5 335	3 618	-47.5	12	
S. Dakota	28/33	3 090	3 531	12.5	12	
MW states	84/93	8 425	7 149	-17.8	8	
USNGP	174/191	23 791	19 082	-24.7	5	-4.94*
USGP	407/431	51 091	44 779	-14.1	4	-3.53*

[a] n is the segment used; M is the number of segments allocated.

[N] The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

* The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-3.– Concluded.

Region	n/M (a)	USDA/ SRS	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
			Final			
Winter Wheat						
Colorado	30/32	2 200	2 704	18.6	24	
Kansas	81/84	11 300	11 125	-1.6	5	
Nebraska	33/35	2 950	3 399	13.2	11	
Oklahoma	40/40	6 300	4 261	-47.9	14	
Texas	47/49	4 700	4 344	-8.2	16	
USSGP	233/240	27 450	25 833	-6.3	5	-1.26[N]
Montana	36/38	3 080	2 079	-48.1	28	
S. Dakota	9/10	970	1 452	33.2	23	
MW states	45/48	4 050	3 531	-14.7	19	
USGP	278/288	31 500	29 364	-7.3	5	-1.46[N]
Spring Wheat						
Minnesota	11/13	3 893	2 198	-77.1	30	
N. Dakota	79/85	11 520	9 856	-16.9	5	
SW states	90/98	15 413	12 054	-27.9	7	
Montana	20/22	2 335	1 516	-54.0	22	
S. Dakota	19/23	2 020	2 079	2.8	13	
MW states	39/45	4 355	3 595	-21.1	12	
USGP	129/143	19 768	15 649	-26.3	6	-4.38*
Total Wheat						
Montana	56/60	5 415	3 595	-50.6	12	
S. Dakota	28/33	2 990	3 531	15.3	12	
MW states	84/93	8 405	7 126	-17.9	8	
USNGP	174/191	23 818	19 180	-24.2	5	-4.84*
USGP	407/431	51 268	45 013	-13.9	4	-3.48*

[a] n is the segment used; M is the number of segments allocated.

[N] The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

* The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

Plot 1 shows that the LACIE estimates for the USSGP region were lower than the USDA/SRS estimates for every month except June.

The statistical tests showed that the LACIE estimates for February, March, and April were significantly different from the corresponding USDA/SRS estimates. These lower estimates are expected early in the season, because a significant number of wheat fields have not yet "greened up" enough to have a characteristic wheat signautre. In 1976 this effect was especially apparent in Kansas, Oklahoma, and Texas because these states were affected by drought. In May and June, the LACIE estimate for the USSGP improved and was not significantly different from the USDA/SRS estimate from May through the final estimate. In June, it was closer to the final USDA/SRS estimate (which held from July on) than the June USDA/SRS estimate. The final LACIE estimate had a relative difference of -6.3 percent and a CV of 5 percent.

The most serious problem in the USSGP region was the underestimates for Oklahoma, shown in plot 2. Blind site investigations (section 4.2.2) indicate that the major source of the underestimate in Oklahoma was due to analyst-mislabeled fields resulting from early dry conditions and an unusual wheat growth cycle following spring rains. In the latter case, the wheat was late in greening up and had signatures that were quite different from normal wheat. In fact, comparisons of LACIE blind site ground observations, aircraft photography and analyst labels on a field-by-field basis indicated that the analysts rarely misidentified nonwheat fields as wheat, but the underestimate resulted primarily from labeling wheat fields as nonwheat.

The winter wheat acreage estimates for the two mixed wheat states are shown in plot 3. These estimates were very low in June but increased throughout the season. The relative difference for the final estimate was -14.7 percent.

Plot 4 shows the total USGP winter wheat estimates. The final estimate had a relative difference of -7.3 percent. July was the only month for which the LACIE estimate was significantly different from the USDA/SRS estimate.

Spring Wheat

Plot 5 shows the spring wheat in the spring wheat states, Minnesota and North Dakota. There was consistent underestimation by LACIE but there was a considerable improvement in September. Part of this was due to a change in the ratios of wheat to small grains that were used to calculate the wheat acreage. For spring wheat, CAMS normally determines only small grains proportions, and the wheat proportions are then calculated by multiplying these by the historical wheat-to-small-grains ratios for the county in which the segment is located. A change in these ratios accounted for 48 percent of the improvement in North Dakota and 53 percent of the improvement in Minnesota. In North Dakota a further 36 percent of the improvement was due to the addition of 21 new segments. These new segments were added to North Dakota to correct a sampling problem identified during Phase I. It is also expected that there was a undersampling problem in Minnesota, since the acreage has increased from 829 000 acres in 1969 (the year that was used for the sampling allocation) to 2 844 000 acres in 1976. Blind site investigations (section 4.2.2.2) indicated a number of causes for the underestimate in North Dakota, including poor Landsat resolution of strip fallow areas, weak or missing signatures, and poor acquisition histories.

Plot 6 shows the spring wheat estimates for the two mixed wheat states, Montana and South Dakota. They show consistently low estimates in the total, but the estimates improved as the season progressed. The improvement was due partly to improved spring-wheat-to-small-grains ratios. The final spring wheat estimate for the mixed wheat states had a relative difference of

-21.1 percent. The results presented in table 4-3 show that there was an underestimation problem in Montana, where the relative difference for the final estimate was 54.0 percent. Investigations (section 4.2.2.2) indicated that this was due largely to underestimates of wheat proportions in strip fallow areas, which did not classify well because Landsat resolution is not fine enough to resolve the fields.

The monthly estimates for the total spring wheat in the USGP region are shown in plot 7. The LACIE estimates were consistently low and were significantly different from the USDA/SRS estimates for every month and for the final estimate. Of the four states contributing to the total spring wheat estimate, only for one, South Dakota, was the spring wheat acreage not consistently underestimated. This indicates a serious underestimation problem for spring wheat. In addition to the reasons given above, blind site studies discussed in section 4.2.2.2 indicate that this underestimation was also due to errors in the ratios of wheat to small grains that were used to calculate the wheat acreage.

Total Wheat

Plot 8 shows the total wheat in the four-state USNGP. It was consistently underestimated and was significantly different from the USDA/SRS estimate for every month and for the final estimate. The final estimate had a relative difference of -24.2 percent due to underestimates of spring wheat in Montana, Minnesota, and North Dakota, and of winter wheat in Montana.

Plot 9 shows the total wheat in the nine-state USGP region. The LACIE estimate was consistently low and was significantly different from the USDA/SRS estimate for every month and for the final estimate. The final estimate had a relative difference of -13.9 percent due to an underestimate of 2.2×10^6 acres (relative difference -7.3 percent) in the winter wheat acreage and an underestimate of 4.1×10^6 acres (relative difference of -26.3 percent) in the spring wheat acreage.

4.2.2 INVESTIGATIONS OF CLASSIFICATION ERROR

Blind site investigations for winter and spring wheat are discussed separately in this report. Refer to section 4.2.2.1 for discussion of winter wheat investigations and 4.2.2.2 for spring wheat investigations.

4.2.2.1 Winter Wheat Blind Site Investigations

The winter wheat blind site investigation consisted of two parts: (1) an early-season investigation for April, and (2) a late-season investigation for October. A different set of blind sites was used in each investigation and each is described separately in the following paragraphs.

Early Season Investigation

The LACIE Phase II examination of early season acreage estimation involved evaluations of acquisitions acquired after emergence and through February; these acquisitions were classified by the CAMS and passed to CAS. Forty blind sites were selected randomly from these acquisitions, and aircraft photography was obtained. Field overlays were prepared and then used by the USDA/ASCS to acquire ground truth land-use information. Classification and ground truth data were obtained for 29 of the 40 blind sites and for 6 intensive test sites. This was the basic data set used in the early season acreage estimation evaluations, the results of which are reported in table 4-4.

A review of table 4-4 shows that the average $\bar{\bar{X}}$ of the LACIE estimates over the 35 sites in the five states of the USSGP was less by -9.17 percent than the average $\bar{X}$ of ground-observed proportions in these states. More detailed investigations were then

TABLE 4-4.— ESTIMATES OF EARLY SEASON SMALL-GRAIN PERCENTAGES FOR 29 BLIND SITES AND 6 INTENSIVE TEST SITES IN THE USSGP

Region	Number of segments	$\bar{\hat{X}}$, %	$\bar{X}$, %	$\bar{\hat{X}} - \bar{X}$, %
Colorado	2	2.30	10.15	-7.85
Kansas	14	22.50	29.80	-7.30
Texas	10	9.80	19.58	-9.78
Nebraska	3	13.43	21.76	-8.33
Oklahoma	6	21.48	35.06	-13.58
Overall 5-state	35	16.50	25.97	-9.17

conducted over a subset (20) of the blind sites, where comparisons of analyzed Landsat and aircraft imagery could be made. These assessments showed:

a. Visual interpretations of Landsat and aircraft color infrared signatures were very similar when acquisition dates were within 10 days of each other.

b. Overall, many wheat fields had little if any wheat signatures (pink) on either the aircraft or Landsat color infrared products, indicating that thin stands of wheat were not being detected.

c. Many reasons for thin (undetectable) wheat stands were identified — most stemming from drought effects; e.g.,

- Eight of the twenty segments showed drought effects.
- Six of the twenty segments were damaged by mosaic virus, army worms, or greenbugs.
- Heavy grazing of cattle was also identified as a cause, inasmuch as it is a common practice in some areas (e.g., Oklahoma) until mid-March, regardless of drought conditions.

The drought effects were studied further over a representative intensive test site (ITS) in the fall drought area (Rice County, Kansas). Acquisitions and classifications over this site showed no significant change until after favorable weather occurred in the spring (March). At that time, a significant improvement in detectable wheat signatures was noted, and the LACIE estimate (47 percent wheat) was fairly close to the ground-truth proportion (50 percent wheat).

Late Season Investigation

The early investigation was conducted with only 29 blind sites, because when those studies were begun, ground truth data were available for only a limited number of blind sites. However, by October, the data had been obtained for many more blind sites in the five-state winter wheat region. As a result, a new investigation was performed using 103 blind sites and the CAMS classification results for these blind sites corresponding to the October LACIE estimates. The results are shown in figure 4-3 and tables 4-5 and 4-6.

Figure 4-3 shows plots of the proportion error $\hat{X} - X$ as a function of X where $\hat{X}$ is the CAMS wheat proportion estimate and X is the ground truth wheat proportion. These plots are for the five individual states and the total USSGP five-state region. Points lying above the horizontal line $\hat{X} - X = 0$ correspond to overestimation of wheat proportions by CAMS, and points lying below the line correspond to underestimation.

The plots in figure 4-3 indicate that there is an overall trend toward negative values of $\hat{X} - X$ as X increases for the five-state region and for each of the individual states except Colorado. In other words, for these regions, CAMS tends to underestimate the true wheat proportion when the true wheat proportion is large. In fact, for X > 28 percent, there is only

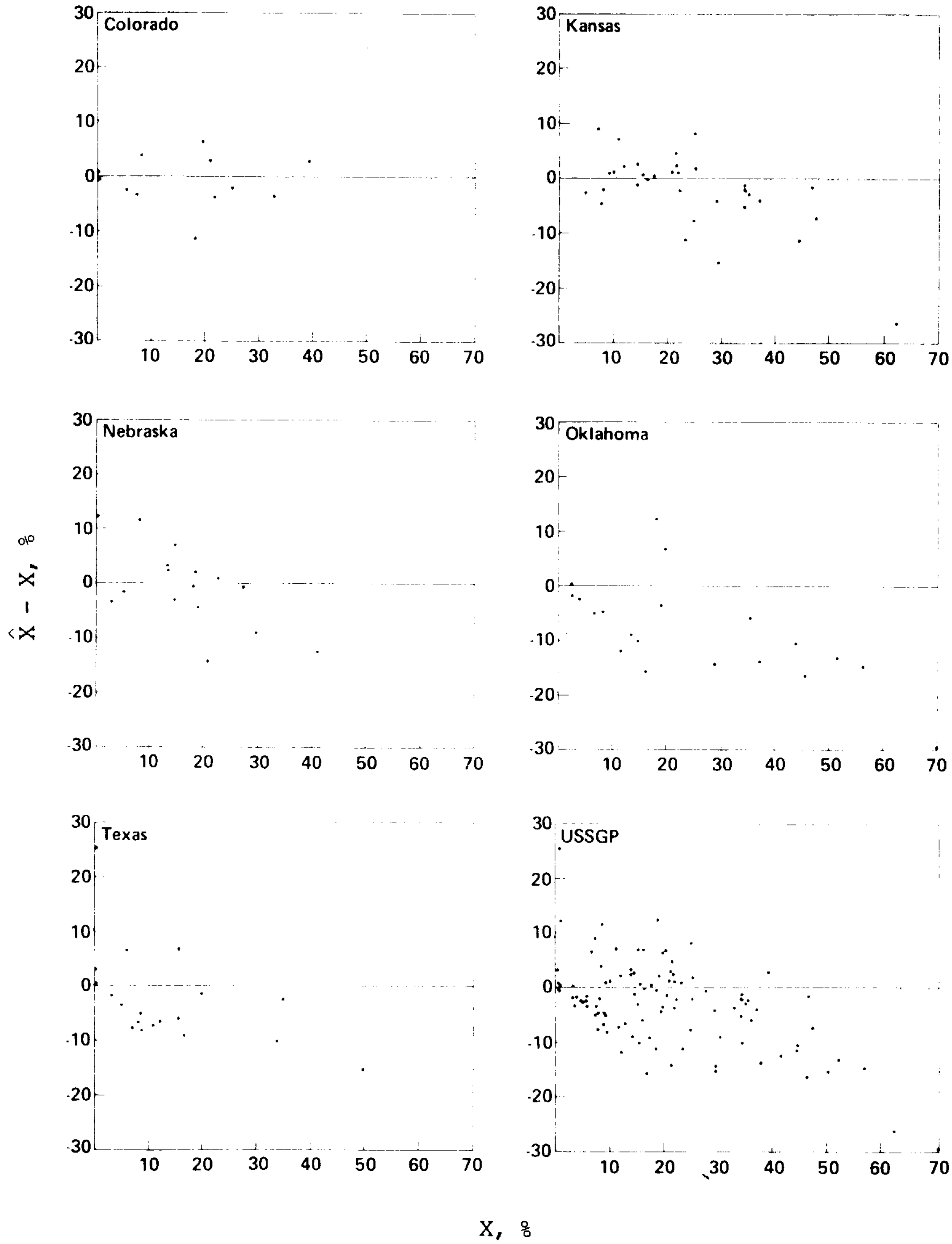


Figure 4-3.— Plot of winter wheat proportion estimation errors versus ground truth winter wheat proportions for blind sites in the USSGP.

one blind site out of 26 in the five-state region for which the CAMS result is not an underestimate relative to ground truth. Also, figure 4-3 indicates that underestimates occur in Oklahoma and Texas for all values of X. In Oklahoma, 17 of 20 (85 percent) of the blind sites were underestimated, as were 15 of 19 (79 percent) in Texas. A statistical analysis of these data follows.

A statistical analysis of the data shown in figure 4-3 was performed using the technique described in appendix A (section A.3.1.1). The results are shown in table 4-5. It lists the following factors: (1) the number of blind sites for which data were available for each state or region, (2) the number of segments allocated to each state or region, (3) the average ground truth wheat proportion, $\overline{X}$, (4) the average CAMS wheat proportion estimate $\overline{\hat{X}}$, (5) the average difference $\overline{D} = \overline{\hat{X}} - \overline{X}$, (6) the standard error $S_{\overline{D}}$ of $\overline{D}$, and (7) 90-percent confidence limits for the average error μ_D.

In order to determine if the average difference for a particular region is significantly different from zero, we need only observe whether the corresponding confidence interval contains zero. If it does, the average difference is not significantly different from zero, i.e., there is insufficient evidence to conclude that there is a bias due to classification error. If it does not contain zero, then the hypothesis of no bias is rejected at the 10-percent level of significance.

In the following paragraphs the results presented in table 4-5 are discussed separately for each state and for the USSGP. The discussion also includes preliminary results from an investigation by CAMS to determine the causes of classification error. At the end of the 1976 crop year, the data for one-half of the blind sites in the USGP were released to CAMS for evaluation of the accuracy and sources of error in the operational analysis

during Phase II. These evaluations were carried out in most cases by the analyst that conducted the original interpretation and classification. In the following paragraphs these studies will be referred to as the "CAMS investigation."

Oklahoma

The results for Oklahoma (table 4-5) show that the 90-percent confidence interval for μ_D is given by (-9.19, -3.97). This interval does not contain zero. Hence, we conclude that there is a negative bias in the CAMS estimates for the segments allocated to Oklahoma. The CAMS investigation showed that underestimates were due to atypical, weak, and missing signatures, small fields, and spotty stands. Some of these effects were attributed to drought conditions. Only one of the segments checked in the CAMS investigation was overestimated; hail damage of wheat at harvest was the cause of the overestimate.

Kansas

In table 4-5 it is also observed that a "significant" bias occurs for the state of Kansas. However, inspection of the data plotted in figure 4-3 reveals one outlier, a difference of -25.56 percent, corresponding to a ground truth of 61.56 percent wheat. Omitting this one outlier yields an estimate of the bias that is not significantly different from zero. From the CAMS investigation it was concluded that in Kansas, overestimates were due to pasture, fallow, and sorghum being included as wheat. Underestimates were usually caused by missed wheat signatures; i.e., wheat signatures that were not included in the training data.

Texas

For Texas, 79 percent of the blind sites were underestimated. However, the $S_{\bar{D}}$ was so large that there was insufficient evidence to conclude that a bias existed. Inspection of the data plotted in figure 4-3 for Texas reveals an outlier, a difference of +25.31 percent, corresponding to a ground truth of 0.69 percent; i.e., an extreme overestimate of a trace of wheat. If this outlier is omitted the results do indicate a negative bias. The CAMS investigation showed that the overestimate for this outlier was due to fallow fields and pasture fields which appeared red and tan, respectively, on the PFC and which were classified as wheat. No explanation was found for the red fallow signatures. The underestimates that occurred for most of the segments were generally due to atypical signatures. Some stands of wheat were spotty.

Colorado and Nebraska

Neither of the average differences for the other two states, Colorado and Nebraska, were significantly different from zero, nor were any apparent outliers observed. The analysts in CAMS were apparently having some success in identifying wheat for these two states. The CAMS investigation showed that in Colorado overestimates were caused by confusion crops such as spring wheat and winter rye being classified as winter wheat; underestimates were caused by missed signatures in drought areas and by strip crop areas not being resolvable by the Landsat system. In the latter case the wheat pixels were all essentially border pixels and therefore many were misclassified as nonwheat.

In Nebraska overestimates were caused by atypical wheat signatures and small fields. Underestimates in Nebraska were due to missed signatures, the absence of key acquisitions such as biowindow 2, some narrow fields that were missed, and some wheat fields that were never picked up on the imagery.

USSGP

At the USSGP five-state level, there was sufficient evidence to conclude that the CAMS wheat proportion estimates were significantly different from the ground wheat proportions at the

10-percent level. The average difference at this level was -1.93 percent with a standard error of 0.58 percent.

Variation of Proportion Error Throughout the Season

Table 4-6 presents the results of a blind site investigation to study the variation of classification error throughout the season.

At the time this investigation was performed (December 1976), all the blind site data were available, but all of the segments could not be used since CAMS estimates for the whole season were not available for all of them. It is, of course, desirable that the same number of segments be used for each month. It was found that 95 segments had data for March through the end of the season, but only 71 segments had data for February.

In table 4-6 four quantities relating to the classification error are given: the mean square error (MSE), the mean difference ($\bar{D}$), the relative mean difference (RMD) and the percentage of the segments in which the LACIE underestimated the at-harvest wheat proportions. There was a declining trend in the MSE throughout the season. The final figure represents a 55-percent reduction from the February estimate.

The $\bar{D}$ and the RMD showed the same behavior; i.e., a general reduction in the size of the error as the season progressed. These errors were all negative, indicating underestimates by LACIE. From February through the final estimate there was a 58-percent reduction in the magnitude of the $\bar{D}$ and a 57-percent reduction in the magnitude of the RMD.

The percentage of segments underestimated by LACIE also decreased throughout the season, falling from 83 percent in February to 68 percent for the final estimate.

TABLE 4-6.— COMPARISON OF LACIE ESTIMATES TO GROUND-OBSERVED PROPORTIONS OVER WINTER WHEAT BLIND SITES IN THE USGP

Month	No. of Segments	MSE (a)	$\bar{D}$, % (b)	RMD, % (c)	Percent underestimated (d)
February	71	157.5	-6.46	-30.6	83
March	95	112.8	-5.43	-26.2	79
April	95	112.8	-5.43	-26.2	79
May	95	102.5	-4.44	-21.4	75
June	95	89.5	-3.25	-15.7	72
July	95	90.4	-3.35	-16.2	70
August	95	75.0	-3.16	-15.2	71
September	95	65.3	-2.76	-13.3	68
October	95	69.6	-2.84	-13.7	68
Final	95	70.8	-2.74	-13.2	68

[a] $MSE = \frac{\sum\left(\hat{X}_i - X_i\right)^2}{n}$ where $\hat{X}_i$ is the wheat proportion estimate for the *ith* segment, X_i is the ground-observed, harvested wheat proportion for the *ith* segment, and n is the number of segments.

[b] $\bar{D} = \frac{\sum\left(\hat{X}_i - X_i\right)}{n} = \bar{\hat{X}} - \bar{X}$.

[c] $RMD = \bar{D}/\bar{X}$.

[d] This column contains the percentage of blind site segments in which LACIE underestimated the wheat proportions.

All these estimates thus indicate a general improvement in the CAMS estimates as the season progressed.

4.2.2.2 Spring Wheat Blind Site Investigations

The spring wheat blind site investigation was conducted in 33 segments in the four USNGP states of Minnesota, Montana, North Dakota, and South Dakota. Figure 4-4 shows plots of the proportion error $\hat{X} - X$ as a function of X, where $\hat{X}$ is the CAMS wheat proportion estimate and X is the ground truth wheat proportion estimate. The plots are for each of the four USNGP states and for the USNGP total spring wheat. Points lying above the horizontal line $\hat{X} - X = 0$ correspond to overestimation of wheat proportions by CAMS, and points lying below the line correspond to underestimation by CAMS.

The plots in figure 4-4 show a tendency toward underestimation in every state except South Dakota. Twenty-eight of the thirty-three sites in the USNGP were underestimated by CAMS. In the plot for the USNGP there appeared to be a slight dependence on the value of X (i.e., the underestimates seem to be greater for larger values of X), but this trend was less pronounced than that shown in figure 4-3 for the USSGP.

The statistical analysis of these data is presented in table 4-7. The quantities listed are the same as those in table 4-5.

Table 4-7 shows that the LACIE acreage estimates were low for all of the states; however, the only state in which the underestimate is statistically significant at the 10-percent level of significance is North Dakota. The CAMS investigation* found many factors

*See section 4.2.2.1.

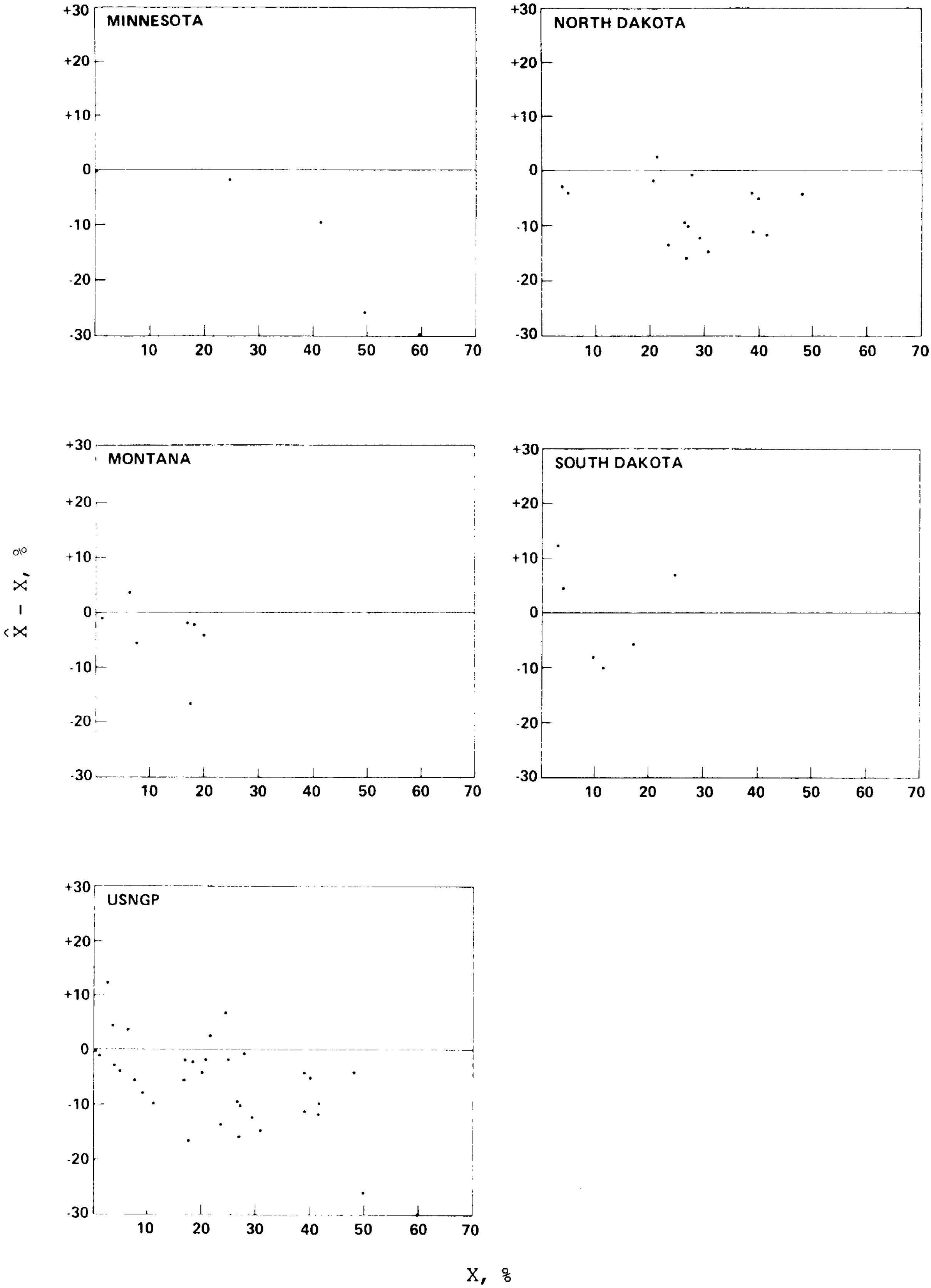


Figure 4-4.– Plots of spring wheat proportions estimation errors versus ground truth values for blind sites in the USNGP.

TABLE 4-7.— SPRING WHEAT BLIND SITE RESULTS FOR THE USNGP

Region	n	N	$\bar{X}$	$\bar{\hat{X}}$ (a)	$\bar{D}$	$S_{\bar{D}}$	90% Confidence Limits for μ_D
Minnesota	5	13	35.43	22.60	-12.82	5.11	(-23.71, 1.93)
North Dakota	17	85	26.64	20.82	-5.82	1.95	(-9.22, -2.42)*
Montana	7	22	12.71	8.57	-4.13	1.95	(-7.92, 0.34)
South Dakota	6	23	11.34	11.17	-0.17	3.20	(-6.62, 6.28)
USNGP	35	143	22.48	16.97	-5.51	1.44	(-7.95, -3.07)*

[a]Final estimates from the CAS annual report for the 1976 crop year.
*μ_D significantly different from zero at the 10-percent level of significance.

which contributed to the underestimate in North Dakota. Among these were:

a. Strip fallow areas unresolvable by the Landsat system

b. Weak or missing signatures

c. Poor color balance on Landsat images due to the transformation that is applied to the Landsat data before the images are made

d. The absence of early biowindow acquisitions

e. The omission of some late-planted spring wheat because its signature was behind the adjustable crop calendar for jointing

f. Problems in choosing training fields caused by small fields or the absence of identifiable field patterns

For Minnesota, Montana, and South Dakota, the analysis did not indicate that there was a bias in the CAMS estimates. However, for these states the number of data points was small. Therefore, the inference of "no bias" should not be regarded as reliable.

Minnesota

In Minnesota underestimation generally occurred in segments with very high wheat density and was caused by unusual wheat signatures, e.g., red-green, light green and dark green, on the PFC products. There is some evidence that these unusual signatures were the result of color distortions in the Landsat imagery.

Montana

In Montana underestimation was usually due to strip fallow areas which were not classified well. Some overestimates were due to hay being classified as wheat even though the two were not confused in the training fields.

South Dakota

In South Dakota both overestimates and underestimates were caused by drought conditions. There was noticeable difference between the Landsat data for this area and for the USSGP. In the spring, wheat and small grains appeared very similar to pasture, alfalfa, and corn on the PFC products due to stress caused by drought. At harvest time, some corn was grazed or cut for silage and some alfalfa was cut and, because of drought, never reappeared. In both cases it was difficult to distinguish these crops from harvested small grains. Many small grains were not harvested, but were fall plowed and could not be distinguished from harvested small grains by CAMS; therefore, wheat was overestimated. Underestimates were due to missing signatures from poor stands of small grains and poor acquisition histories.

USNGP

For the blind sites in the USNGP, the analysis indicated a bias in the CAMS wheat proportion estimates. The average difference was -5.51 percent with a standard error of 1.44 percent.

Contribution of the Classification and Ratio Errors to the Ratioed Wheat Proportion Estimation Errors at the Segment Level

CAMS makes estimates of the small-grains proportion $\hat{X}_i$ for each segment i and, subsequently, CAS obtains wheat proportion estimates by multiplying the $\hat{X}_i$ by the ratios $\hat{r}_i$ of the wheat-to-small-grains proportions for the counties in which the segments are located as determined from the 1975 SRS estimates. In this section, the blind site data are used to compare the error incurred by using these ratios to the error incurred by misclassification of small grains.

Let n be the number of blind sites, r_i be the ground-observed ratios of wheat-to-small-grains proportions, and X_i be the ground-observed small-grains proportions. The bias (B) and the mean-squared error (MSE) of the wheat proportion estimate for a segment may be estimated by

$$\hat{B} = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{r}_i \hat{X}_i - r_i X_i\right)$$

and

$$\hat{MSE} = \frac{1}{n} \sum_{i=1}^{n} \left(\hat{r}_i \hat{X}_i - r_i X_i\right)^2$$

respectively. It is clear that these errors are both caused by two factors: the CAMS classification of small grains and the estimated ratio of wheat to small grains. The contribution of a particular error factor may be measured by the reduction in the bias or mean-squared error which would be achieved if that error factor were omitted. Specifically, the following formulas are used in this study.

a. Proportion bias estimate without ratio error:

$$\hat{B}' = \frac{1}{n} \sum_{i=1}^{n} \left(r_i \hat{X}_i - r_i X_i\right)$$

b. Proportion bias estimate without classification error:

$$\hat{B}'' = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{r}_i X_i - r_i X_i\right)$$

c. Proportion mean squared error without ratio error:

$$\hat{MSE}' = \frac{1}{n}\sum_{i=1}^{n}\left(r_i \hat{X}_i - r_i X_i\right)^2$$

d. Proportion mean-squared error without classification error:

$$\hat{MSE}'' = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{r}_i X_i - r_i X_i\right)^2$$

Table 4-8 presents the numerical results obtained for 37 spring wheat blind sites for Phase II in Minnesota, Montana, North Dakota, and South Dakota using the final estimates for $\hat{X}_i$.

TABLE 4-8.— PHASE II FINAL RESULTS FOR SPRING WHEAT BLIND SITES IN USNGP

Category	Estimate of bias, %	Standard dev. of bias	Reduction in bias, %	90% Confidence limits for bias	Mean squared error	Reduction in mean squared error, %
Phase II final result	-4.89	9.70	—	(-7.58, -2.19)	115.36	—
No ratioing error	-2.45	8.54	49.9	(-4.82, -0.07)	76.91	33.3
No classification error	-3.12	4.03	36.2	(-4.23, -2.00)	25.50	77.9

From table 4-8 it can be seen that the reduction in bias is not much larger when there is no ratioing error than when there is no small grain classification error. On the other hand, a much larger reduction in mean-squared error is obtained when there is no small grain classification error than when there is no ratioing error. This indicates that the major problem is the classification of small grains. If the classification problem is solved, or at least reduced, then a bias still exists due to ratioing. Hence, both problems need to be attacked, with more emphasis on the classification problem.

Variation of Proportion Error Throughout the Season

Table 4-9 shows the results of a blind site investigation to study the variation of classification error throughout the season. All 33 segments were used. The definitions of the quantities listed are the same as those given in section 4.2.2.1 in connection with table 4-6.

TABLE 4-9.— MEASUREMENTS OF CLASSIFICATION ERROR (LACIE ESTIMATES VERSUS GROUND-OBSERVED PROPORTIONS) OVER ALL AVAILABLE BLIND SITES IN THE USGP

SPRING WHEAT					
Month	No. of segments	MSE	$\bar{D}$, %	RMD, %	[a]% under-estimated
August	33	158.5	-9.29	-41.6	88
September	33	120.1	-5.72	-25.6	82
October	33	115.3	-5.38	-24.1	79
Final	33	110.1	-5.05	-22.6	79

[a]This column contains the percentage of blind site segments in which LACIE underestimated the wheat proportion.

The mean-squared classification error dropped from 158.5 in August to 110.1 at the end of the season — a decrease of 30 percent.

The average difference $\bar{D}$ was negative for all months, indicating that the wheat proportions were consistently underestimated throughout the year. The magnitude of the errors declined 45 percent in the period from August to the final estimate. In spite of these reductions there was still substantial underestimation at the end of the season. At that time the wheat proportion in 79 percent of the sites was still being underestimated by LACIE.

4.2.2.3 Bias Due to Classification Error

Ground truth information from blind site data obtained at harvest was used to estimate bias due to classification. The procedure is described in appendix A, section A.3.1.4. In addition to the assumption of normality for $\hat{X}$, it is based on the following assumptions:

a. The blind sites within a state are representative of the sample segments allocated to the state.

b. The estimates of classification bias at the segment level are assumed to be independently and identically distributed for each allocated segment within a state.

c. The acreage estimates are uncorrelated at the state level and any bias in a state acreage estimate is due to classification.

d. The derived state level yield estimates are uncorrelated and are unbiased.

e. The state level acreage and yield estimates are uncorrelated.

f. The bias due to the Group III ratio estimates is negligible.

Under these assumptions, the segment level classification bias for each state is estimated by the average difference between the CAMS wheat proportion estimates and the ground truth wheat proportions as determined from the blind sites within each state. The state level acreage bias is then estimated by aggregating this segment level classification bias estimate for each segment acquired in the state in Phase II. The results are given in table 4-10. The estimated acreage bias is significantly less than zero for the USGP region, the four-state spring wheat region

of the USNGP, the seven-state winter wheat region of the USGP, and the five-state winter wheat region of the USSGP. However, if Oklahoma is excluded from the five-state winter wheat region of the USSGP, no bias is indicated for this region.

4.2.3 ESTIMATION OF THE WITHIN-COUNTY ACREAGE VARIANCES DUE TO CLASSIFICATION AND SAMPLING ERRORS

In order to estimate the within-county acreage variances due to sampling and classification errors, one first constructs the following three basic regression models: (1) true segment proportion versus historical county proportion, (2) LACIE segment proportion versus ground truth segment proportion, and (3) LACIE segment proportion versus historical county proportion. Then, the regression equations are used to obtain the estimates for $\sigma_s^2 + \sigma_H^2$, σ_c^2, and $\lambda^2\sigma_s^2 + \sigma_c^2$, where σ_c^2, $\lambda^2\sigma_s^2$ and σ_H^2 represent, respectively, the contribution due to classification, the contribution due to sampling, and the variance of the residuals resulting from the regression of the current county proportion onto the historical county proportion. Assuming that σ_H^2 is much smaller than σ_s^2, σ_H^2 can be ignored in practice. Finally, the maximum likelihood estimation technique, assuming normality, is used to obtain the optimal estimates for sampling and classification variances. The detailed description of this method is presented in appendix A.

Table 4-11 provides the estimates of the acreage variances (within county) due to classification and sampling errors. These estimates were obtained using the CAMS proportion estimates given in the CAS Final Report, the ground truth proportions for the winter wheat blind sites, and the country proportions from the 1974 census.

As indicated in table 4-11, sampling contributes more error than classification does to the estimates of within-county acreage

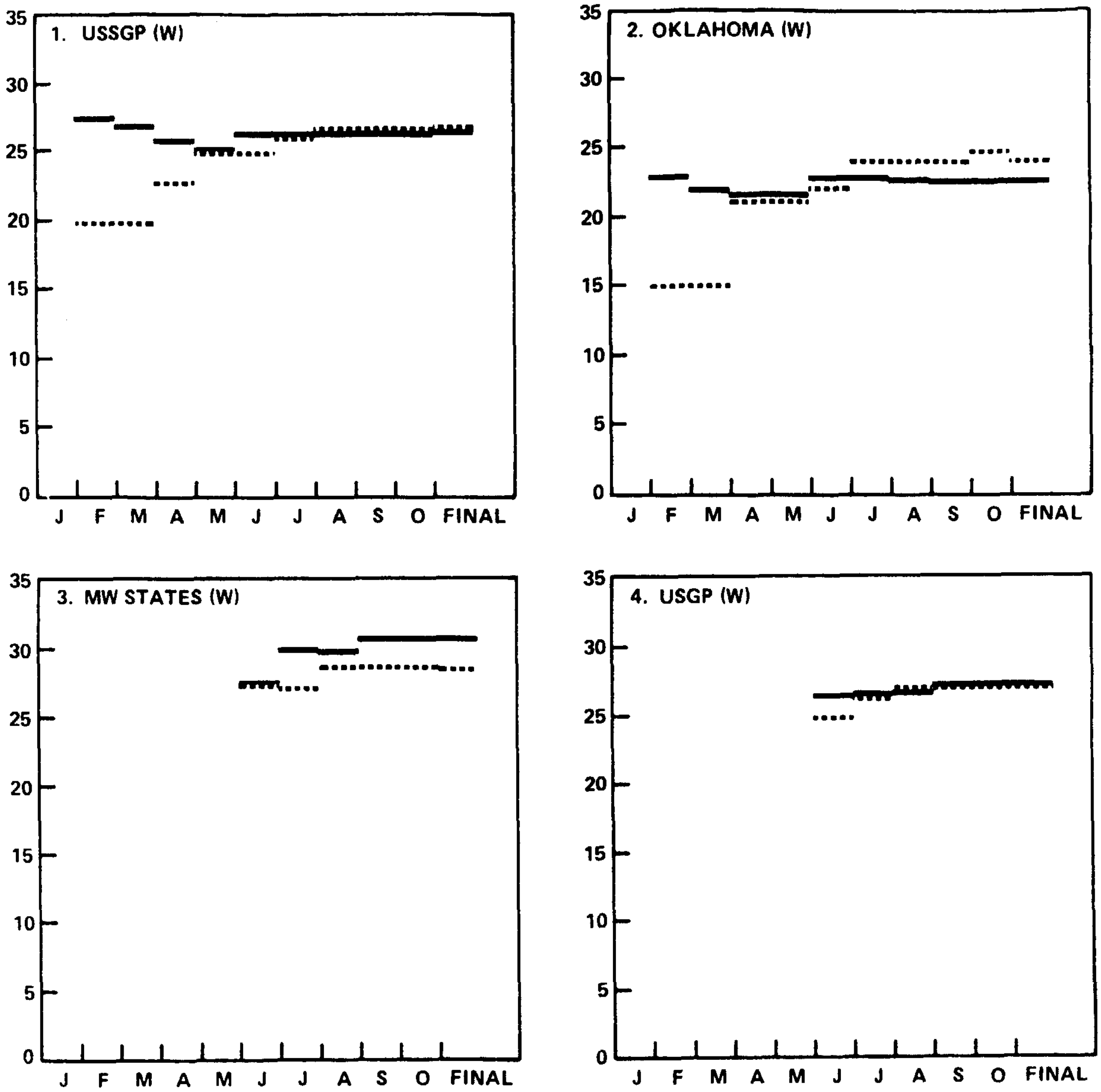
variances for the winter wheat states. No interpretation can be made for the spring wheat states due to (1) the lack of consistency of the results among those states, and (2) the limited number of blind sites used for the error estimation.

4.3 COMPARISON OF LACIE AND USDA/SRS YIELD ESTIMATES

Winter Wheat

The LACIE and USDA/SRS monthly winter wheat yield estimates for the USSGP, the state of Oklahoma, the mixed wheat states of Montana and South Dakota, and the USGP are displayed in plots 1 through 4 of figure 4-5. The estimates and their corresponding relative differences and CV's are presented in table 4-12. Also presented in the table is the test statistic used for determining whether the LACIE estimate is significantly different from the corresponding USDA/SRS estimate. This test statistic was calculated only at regional or higher levels, not at state levels. At the USSGP level, the LACIE estimates were significantly different from the USDA/SRS estimates only for the early season months of February, March, and April. The February and March estimates of yield for USDA/SRS were actually estimates derived by dividing the USDA/SRS production forecast for these months by estimates of seeded (or planted) acres. Therefore, the SRS estimates for these two months were yield per planted acre, rather than yield per harvested acre, which is forecast by LACIE. Hence, it is not surprising that these two estimates were significantly different for February and March. However, none of the monthly LACIE estimates were significantly different from the USDA/SRS final estimate at this level.

The monthly winter wheat yield estimates by LACIE and USDA/SRS for Oklahoma are displayed in plot 2 of figure 4-5 and the corresponding relative differences are given in table 4-12. Plot 2 indicates that the large underestimate of wheat production by



LEGEND
LACIE
USDA/SRS
W = WINTER WHEAT
S = SPRING WHEAT
T = TOTAL WHEAT

Figure 4-5.– LACIE and USDA/SRS yield estimates [bushels/acre].

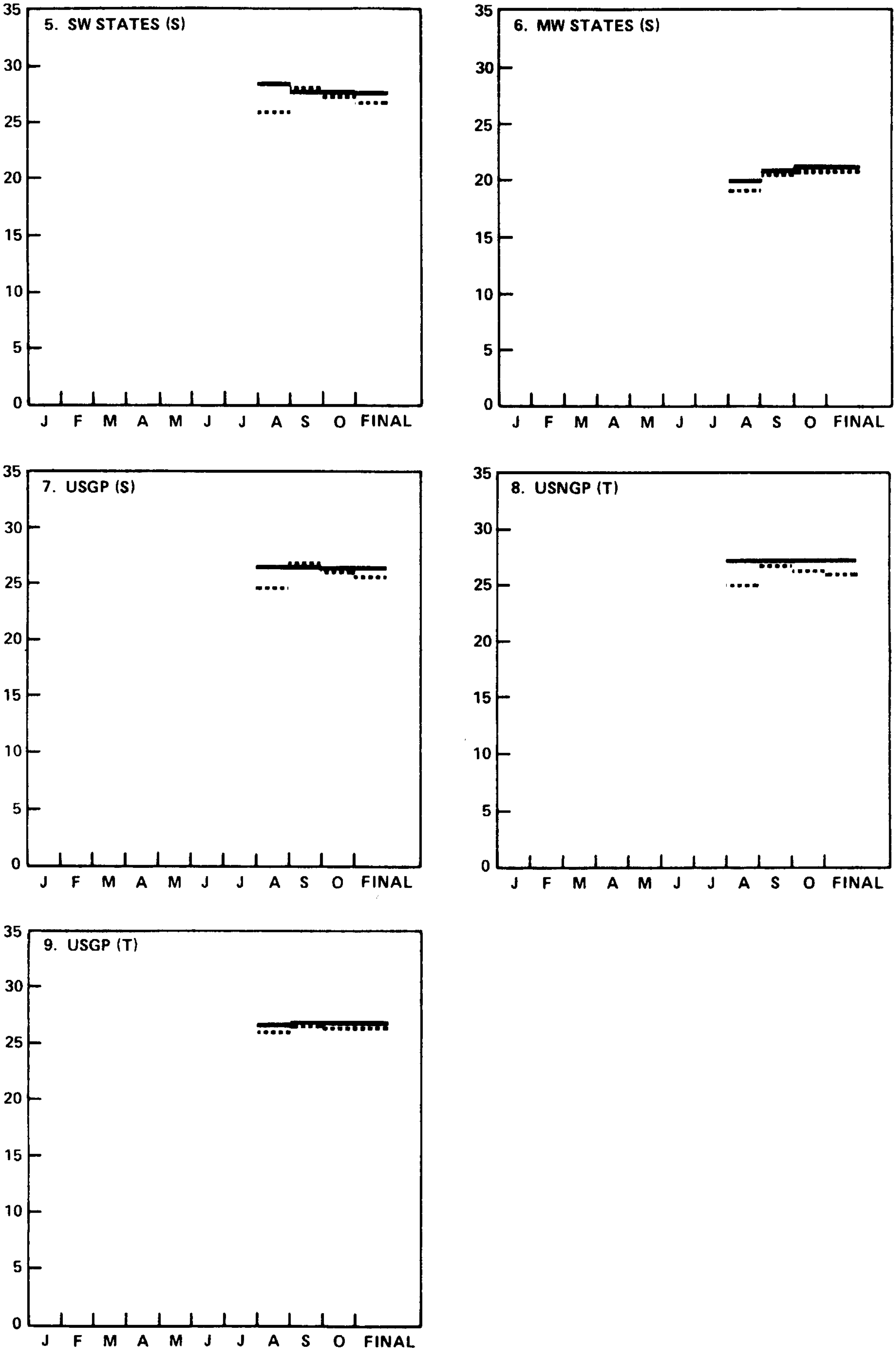


Figure 4-5.— Concluded.

TABLE 4-12.– COMPARISON OF USDA/SRS AND LACIE YIELD ESTIMATES

[Bushels/acre]

Region	USDA/SRS (a)	LACIE	Relative difference (%)	CV (%)	Test statistic
February					
Winter Wheat					
Colorado	17.0	21.6	21.3	21	
Kansas	25.0	32.2	22.4	12	
Nebraska	27.1	33.7	19.6	14	
Oklahoma	15.0	22.9	34.5	17	
Texas	12.0	18.8	36.2	19	
USSGP	19.8	27.6	28.3	7	4.04*
March					
Winter Wheat					
Colorado	17.0	22.0	22.7	21	
Kansas	25.0	31.6	20.9	12	
Nebraska	27.1	34.2	20.8	14	
Oklahoma	15.0	22.0	31.8	17	
Texas	12.0	17.9	33.0	18	
USSGP	19.8	27.0	26.7	7	3.81*

[a]The USDA/SRS yield estimates for February and March were obtained by dividing the production estimates by the corresponding acreage estimates.

*The LACIE estimate is significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-12.— Continued.

Region	USDA/SRS (a)	LACIE	Relative difference (%)	CV (%)	Test sta- tistic
April					
Winter Wheat					
Colorado	21.0	20.3	-3.4	21	
Kansas	26.0	29.9	13.0	10	
Nebraska	28.0	33.1	15.4	14	
Oklahoma	21.0	21.7	3.2	14	
Texas	17.0	17.1	0.6	14	
[b]USSGP	22.7	25.9	12.4	6	2.06*
May					
Winter Wheat					
Colorado	22.0	19.7	-11.7	20	
Kansas	28.0	30.1	7.0	10	
Nebraska	32.0	30.2	-6.0	14	
Oklahoma	21.0	21.7	3.2	14	
Texas	18.0	18.1	0.6	13	
[b]USSGP	24.9	25.3	1.6	6	.27[N]

[b]The five-state United States southern Great Plains region.

[N]The LACIE estimate is not significantly different from the USDA/SRS estimate at the 10-percent level.

TABLE 4-12.– Continued.

Region	USDA/SRS (a)	LACIE	Relative difference (%)	CV (%)	Test sta-tistic
June					
Winter Wheat					
Colorado	22.0	20.4	-7.8	17	
Kansas	26.0	31.0	16.1	9	
Nebraska	33.0	31.4	-5.1	13	
Oklahoma	22.0	22.9	3.9	10	
Texas	18.0	18.5	2.7	12	
USSGP	24.4	26.4	7.6	5	1.52^N
Montana	30.0	27.7	-8.3	12	
S. Dakota	20.0	27.2	26.5	15	
[c]MW states	27.4	27.4	0	9	
[d]USGP	24.8	26.5	6.4	5	1.28^N
July					
Winter Wheat					
Colorado	22.0	18.0	-22.2	17	
Kansas	29.0	30.9	6.1	9	
Nebraska	32.0	32.0	0	12	
Oklahoma	24.0	22.9	-4.8	10	
Texas	21.0	18.7	-12.3	12	
USSGP	26.2	26.4	0.8	5	0.16^N
Montana	31.0	28.8	-7.6	9	
S. Dakota	16.0	30.4	47.4	15	
MW states	27.2	29.8	8.7	9	
USGP	26.4	26.7	1.1	5	0.22^N

[c]The mixed wheat states, Montana and South Dakota.

[d]The nine-state United States Great Plains region.

TABLE 4-12.– Continued.

Region	USDA/SRS (a)	LACIE	Relative difference (%)	CV (%)	Test sta-tistic
August					
Winter Wheat					
Colorado	22.0	17.7	-24.3	17	
Kansas	29.5	30.9	4.5	9	
Nebraska	32.0	32.0	0	12	
Oklahoma	24.0	22.8	-5.3	10	
Texas	22.0	18.7	-17.6	20	
USSGP	26.6	26.4	-0.8	5	$-.16^N$
Montana	32.0	29.2	-9.6	9	
S. Dakota	19.0	30.4	37.5	14	
MW states	28.7	29.7	3.4	8	
USGP	26.9	26.7	-0.7	5	$-.14^N$
Spring Wheat					
Minnesota	32.0	31.9	-0.3	11	
N. Dakota	23.6	27.7	14.8	11	
[e]SW states	25.7	28.4	9.5	9	
Montana	27.3	25.9	-5.4	9	
S. Dakota	9.9	16.9	41.4	14	
MW states	19.1	20.0	4.5	9	
USGP	24.3	26.3	7.6	7	1.08^N
[f]Total Wheat					
Montana	29.9	28.0	-6.8	4	
S. Dakota	13.0	22.4	42.0	5	
MW states	23.7	24.9	4.8	4	
[g]USNGP	25.0	27.0	7.4	6	1.23^N
USGP	25.9	26.6	2.6	4	$.65^N$

[e]The spring wheat states, Minnesota and North Dakota.

[f]Spring wheat plus winter wheat.

[g]The four-state United States northern Great Plains region.

TABLE 4-12.— Continued.

Region	USDA/SRS (a)	LACIE	Relative difference (%)	CV (%)	Test sta-tistic
September					
Winter Wheat					
Colorado	22.0	19.6	-12.2	17	
Kansas	29.5	30.9	4.5	9	
Nebraska	32.0	32.7	2.1	12	
Oklahoma	24.0	22.6	-6.2	10	
Texas	22.0	18.7	-17.6	5	
USSGP	26.6	26.5	-0.4	5	$-.08^N$
Montana	32.0	29.9	-7.0	9	
S. Dakota	19.0	31.6	39.9	14	
MW states	28.7	30.6	6.2	8	
USGP	26.9	27.0	0.4	5	$.08^N$
Spring Wheat					
Minnesota	34.1	30.3	-12.5	11	
N. Dakota	26.0	27.1	4.1	11	
SW states	28.0	27.7	-1.1	9	
Montana	28.3	27.2	-4.0	9	
S. Dakota	11.9	17.1	30.4	13	
MW states	20.6	21.0	1.9	8	
USGP	26.4	26.3	-0.4	7	$-.05^N$
Total Wheat					
Montana	30.4	28.9	-5.2	5	
S. Dakota	14.3	23.1	38.1	5	
MW states	24.5	25.9	5.4	4	
USNGP	26.7	27.1	1.5	7	$.21^N$
USGP	26.7	26.8	0.4	4	$.10^N$

[illegible]

LACIE for this state was not due to the yield predictions. The LACIE estimates of yield were only slightly lower than the corresponding USDA/SRS estimates from July to the final estimate.

The winter wheat yield estimates by LACIE and USDA/SRS for the two-state mixed wheat region of Montana and South Dakota are exhibited in plot 3. The LACIE yield estimates were consistently lower than the USDA/SRS yield estimates in Montana and consistently higher in South Dakota. Combining the two resulted in a consistent overestimation by LACIE over USDA/SRS for the two-state total. The overestimation in South Dakota was due to the incapability of the LACIE yield model for this state to forecast the impact of the unusually dry weather conditions for this crop year. This indicates the need for improved yield models at the zone level for predictions in extreme weather conditions.

The monthly total winter wheat yield estimates for the seven states in the USGP region are given in plot 4. At this level, the LACIE estimates were not significantly different from the USDA/SRS estimates for any of the months reported. In fact, the two final estimates were identical.

Spring Wheat

The LACIE and USDA/SRS spring wheat yield estimates for the two-state spring wheat region of Minnesota and North Dakota are given in plot 5 and the corresponding relative differences are reported in table 4-12. The monthly LACIE estimates of yield for Minnesota were consistently lower than the USDA/SRS estimates. On the other hand, the LACIE estimates of yield for North Dakota were consistently higher than the USDA/SRS estimates. As a result, the LACIE two-state total estimates were very close to the USDA/SRS estimates except for the month of August.

Plot 6 displays the monthly estimates of spring wheat yield by LACIE and USDA/SRS for the two-state mixed wheat region. Table 4-12 contains the corresponding relative differences for these plots. The LACIE estimates of yield for South Dakota were considerably higher than the USDA/SRS estimates. Recall that the same situation occurred for the winter wheat yield estimates for this state. The LACIE yield estimates for Montana, however, were lower but much closer to the corresponding USDA/SRS estimates, except for August when the LACIE estimate was slightly higher. The two-state total spring wheat estimates by LACIE were, as a result, higher but very comparable to the USDA/SRS estimates.

The total spring wheat yield estimates for the four states in the USNGP are given in plot 7. Table 4-12 shows the corresponding relative differences and CV's. The LACIE estimates were not significantly different from the corresponding USDA/SRS estimates for any month reported.

Total Wheat

The LACIE and USDA/SRS monthly total wheat yield estimates for the USNGP are displayed in plot 8 and the relative differences and CV's corresponding to this plot are shown in table 4-12.

The LACIE estimates were consistently higher than the USDA/SRS estimates for all four months, but were not significantly different from them.

The monthly total wheat yield estimates obtained by LACIE and USDA/SRS for all nine states in the USGP are displayed in plot 9 and the corresponding relative differences and CV's are given in table 4-12. The two estimates were not significantly different for any month reported. Hence, the LACIE yield estimates at this level were considerably more accurate (as compared to USDA/SRS estimates) than the LACIE acreage estimates for Phase II.

5. PHASE I SPECIAL STUDIES

A number of special studies that were carried out in Phase I are discussed in this section. With the exception of the crop calendar study described in section 5.5, they are all concerned with the effects of various factors on classification accuracy.

5.1 A STUDY OF THE EFFECTS OF SITE, BIOPHASE, AND AI

5.1.1 INTRODUCTION

A study was conducted to investigate the effects of three major factors – site, biophase, and analyst interpreter (AI) – on errors in the estimation of segment small grains proportions. All 14 AI's operating within CAMS for the LACIE Phase I operations participated in this experiment. The test was run on two intensive test sites (ITS's): segment 1969, Toole County, Montana, and segment 1976, Franklin County, Idaho. These segments were selected because MSS data were available for all four biophases. (Classifications for at least one biophase were missing for all the other ITS's.) Each AI was required to interpret each biophase acquisition for each segment using the Phase I operational procedure. This resulted in a total of 56 small grains proportion estimates for each segment. The data are given in table 5-1. Table 5-2 lists some general observations made regarding these two sites.

The analysis of variance (ANOVA) approach was used to analyze the data. Let $\hat{X}$ be the CAMS proportions expressed as a fraction rather than a percentage as in table 5-1 and let X be the ground truth proportion. The transformed data T obtained from the standard equation

$$T = \sin^{-1}\sqrt{\hat{X}} \qquad (5\text{-}1)$$

TABLE 5-1.— CAMS PROPORTION ESTIMATE, PERCENTAGE OF SMALL GRAINS

AI	ITS 1969, biophase				ITS 1976, biophase			
	1	2	3	4	1	2	3	4
A	18.8	46.7	50.3	46.6	29.4	29.2	36.7	50.4
B	51.3	36.0	53.6	56.4	49.1	25.2	12.1	30.5
C	16.8	37.4	60.2	31.0	41.0	10.9	17.2	25.7
D	31.4	13.8	53.0	39.3	8.6	15.7	5.6	16.4
E	12.8	47.2	54.6	57.6	23.5	22.6	19.6	32.4
F	35.5	46.6	56.8	57.6	0.0	9.8	0.0	0.0
G	67.5	48.0	52.0	37.0	37.0	25.7	30.5	36.0
H	17.2	41.6	49.0	48.4	22.6	17.8	26.3	26.2
J	25.0	39.7	48.6	38.1	22.6	21.9	30.9	17.4
K	32.1	68.2	32.8	32.1	48.7	10.3	39.4	28.7
L	7.5	44.9	57.4	46.7	42.4	19.6	27.8	35.8
M	25.0	42.5	66.2	47.2	44.2	30.5	35.1	2.9
N	55.2	42.3	38.1	48.3	26.8	21.7	20.2	20.1
O	89.2	36.8	36.1	36.7	49.0	38.3	25.4	48.9
Average per biophase	34.7	42.2	50.6	44.5	31.8	21.4	23.4	26.5
Ground truth	38.3	38.3	38.3	38.3	26.0	26.0	26.0	26.0

TABLE 5-2.— DIFFERENCES IN PHYSICAL CHARACTERISTICS OF INTENSIVE TEST SITES

Factor	Segment	
	1969	1976
Location	Toole County, Montana	Franklin County, Idaho
Size	3.7 by 11 km (2 by 6 n. mi.)	5.6 by 5.6 km (3 by 3 n. mi.)
Small-grain proportion	37.7%	26%
CAMS results	Overestimated in biophases 2, 3, and 4; underestimated in biophase 1	Underestimated in biophases 2 and 3; overestimated in biophases 1 and 4.
Imagery	10% to 15% cloud cover for biophases 2 and 3	Good
AI	More consistent	Higher variability
Ancillary data	More small grains; less winter wheat; strip cropping	Less small grains; more winter wheat; random field contour; irrigated fields in biophase 1