COSPAR D.2.3

THE LARGE AREA CROP INVENTORY EXPERIMENT
A MAJOR DEMONSTRATION OF SPACE REMOTE SENSING

(NASA-TM-74840) THE LARGE AREA CROP INVENTORY EXPERIMENT: A MAJOR DEMONSTRATION OF SPACE REMCTE SENSING (NASA) 49 p N77-30570
CSCL 02C Unclas
G3/43 42053

Presented By

R. B. MacDonald and F. G. Hall

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS 77058

To The

W. NORDBERG MEMORIAL SYMPOSIUM
ON
CONTRIBUTION OF SPACE OBSERVATIONS TO GLOBAL FOOD INFORMATION SYSTEMS

COSPAR/IAMP (IUGG) in Collaboration with FAO and WMO
Tel-Aviv, Israel

June 8-10, 1977

THE LARGE AREA CROP INVENTORY EXPERIMENT
A MAJOR DEMONSTRATION OF SPACE REMOTE SENSING

R. B. MacDonald and F. G. Hall

National Aeronautics and Space Administration
Lyndon B. Johnson Space Center
Houston, Texas 77058

ACKNOWLEDGMENT

The authors wish to acknowledge the many contributions to the planning and execution of this experiment by their colleagues not only in NASA but within the U.S. Department of Agriculture, the National Oceanic and Atmospheric Administration of the U.S. Department of Commerce, and the various participating universities and industrial research entities.

1. INTRODUCTION

Historically, agricultural production throughout the world has been subject to large and irregular fluctuations. These fluctuations are likely to become even more pronounced in the future due to potential changes in the global climatological trends, an increasing shortage of fuels and fertilizers, and increasing population pressures projected for the remainder of this century.

Long range climatic forecasts reported by U.S. government studies indicate a continuation of significant cooling trends in the northern hemisphere and major changes in rainfall patterns. These climatic changes could lead to reduced and shifting crop lands with a net reduction of available cropland in the more frigid northerly latitudes. Strategies in agricultural technology to increase the resistance of crops to a wider range of meteorological conditions in order to reduce year-to-year variations in crop production might well result in reduced average yields. Such uncertainties in agricultural production, together with the consumer demands of an increasing world population, have greatly intensified the need for early and accurate annual global crop production forecasts. These forecasts must predict fluctuations with an accuracy, timeliness and known reliability sufficient to permit necessary social and economic adjustments, with as much advance warning as possible.

2. THE NEED FOR GLOBAL AGRICULTURAL MONITORING INFORMATION SYSTEMS

Leaders of most industrial and developing nations acknowledge that global agricultural planning is a minimum requirement in assuring adequate food supplies at an equitable price. It follows that this planning requires timely and accurate global crop forecasts, and that such forecasts entail a global agricultural monitoring system.

Global agricultural planning is of particular importance to the United States. The United States is the largest food exporter in the world and accounts for approximately one-half of the global grain trade. As a result, from an economic as well as from a humanitarian standpoint, the United States endeavors to maintain, promote, and expand its foreign agricultural markets.

Importing and exporting countries must manage a delicate balance between supply and demand, anticipating the determining factors as far in advance of transactions as possible. In some years wheat reserves have declined to a fraction of the historic demand. In 1974, world reserves of wheat were reported to have dwindled to an amount equivalent to 27 days of consumption. In such a situation, timely information relevent to anticipated resupply from new harvests is crucial. Without reliable and timely crop demand and supply information, an exporting nation might impose a costly, but unncecessary, moratorium on its grain sales. On the other hand, importing countries with limited storage must have early forecasts of their own supply positions in order to make intelligent purchasing decisions. Arrangements involving transportation of foodstuff by shipping and rail can benefit greatly by advance planning when timely and accurate crop forecast and food supply information is available. It is the context of balancing worldwide supply and demand which has historically defined and, more recently, brought attention to the need for a global food and fiber monitoring system.

Accurate and timely crop production forecasts with known reliability must incorporate two types of assessment: first, a periodic within-season assessment of the crop hectarage and condition based on estimates of the areal extent of the standing crop and the seasonal growth conditions through the reporting period; second, an accurate forecast of the most likely range of future growth conditions and the range of probable effects on production at harvest. In addition, it is vitally important to predict the confidence or "odds" that the forecast will agree, to a specified tolerance, with the production actually harvested.

Wheat production estimates serve as an illustrative example of the fundamentals involved in the two types of assessments necessary for accurate crop forecasts. The quantity of wheat to be produced by a current crop will depend on the quantity of producing units (wheat plants) which are finally harvested (product of wheat hectarage and the average number of plants per hectare) and the average productivity per harvested plant (number of heads, grains per head, weight per grain). At each reporting period in the season, prior to harvest, the production forecast must consider the total hectarage of wheat currently standing as well as the current condition of the standing crop, determined by factors such as soil type, slope precipitation, temperature history, and other growth conditions to date. These conditions in turn are manifest through crop condition parameters such as stand density (number of tillers, plant population density) and root development which, along with future weather, will determine the final production. As an example, the yield of a season's wheat crop in regions of soils with high water-holding capacities and adequate soil moisture can often be predicted with a high reliability well before harvest, given an accurate assessment of the stand density and height. Thus, at each particular point in the season, observations of the plant, together with measurements of the past and present weather parameters, can be used to assess the present quantity and condition of the crop. A prediction of future events is then required to forecast the production at harvest. Within a season, both hectares of standing wheat and wheat yield per hectare are subject to a forecast. For example, in winter wheat regions during the late fall period, the existing hectares of wheat plants can be measured while the potential loss to winterkill must be forecast.

While there is no comprehensive economic theory which exactly quantifies the value of improvements resulting from improved forecast accuracies, an analysis of available global agricultural information and the systems which produce the input agricultural statistics provides some insight into the question of the required performance of an improved and acceptable global information system. (See footnote 1.) World supply estimates are a compilation of country supply estimates generated for the most part by the various national agricultural information systems. The quality of world estimates, therefore, is a direct function of the quality of the systems in the various countries. The estimates from this conglomerate range from timely and reliable to nonexistent. All too frequently estimates based on past trends, sometimes adjusted by judgment, are used in lieu of objective sources. The primary properties of an effective world agricultural information system are objectivity, reliability, timeliness, adequacy in terms of coverage, efficiency and effectiveness. "An ideal system would provide timely and unbiased interpretations of the current global situation and an outlook based upon estimates of known reliability for all commodities and countries through the use of the most cost effective procedures known to mankind".

Some comparisons can be made between the international crop production forecasts of the Food and Agriculture Organization (FAO) of the United Nations and the Foreign Agricultural Service (FAS) of the U.S. Department of Agriculture (USDA), the two organizations which currently operate world agricultural information systems. The USDA's world agricultural information system provides more timely information than does FAO's. Statistical reliability of supply estimates is essentially the same for both organizations since they rely in large measure on the data produced by national systems. The USDA releases estimates and assessments more frequently than FAO. The FAO is steadily working to improve the reliability of the estimates of supply through its work with member nations to improve techniques utilized by the nations. This qualitative analysis leads to the following summarization of the primary characteristics of currently available world agricultural supply estimates:

(1) The objectivity of estimates is largely a function of the objectivity of the estimates released by the host government.

(2) The reliability of the estimates is largely a function of the methods used by the nation to collect agricultural statistics and to assess them. This varies significantly from country to country.

(3) Most national systems rank poorly in terms of timeliness of estimates of supply.

(4) Adequacy is impaired by lack of uniformity of reporting both in terms of content and geographic coverage from nation to nation.

(5) The efficiency and effectiveness of most national systems requires significant improvements.

These factors are the main determinants of the forecast accuracies of the various USDA surveys. On the average, the most accurate and timely estimates are made for U.S. agriculture. The Statistical Reporting Service (SRS) of the USDA utilizes probability surveys of area planted, area harvested and the average productivity (yield) from area harvested. For example, winter wheat

---

[1]The following review of the state of the current world agricultural information system and required improvements is in large based on the work of John Schnittker Associates as reported by Howard W. Hjort while a Vice President and partner of the firm; H. W. Hjort currently serves in the U.S. Department of Agriculture as the Director of Agricultural Economics

production estimates are made in December and May and every month thereafter through harvest until the following December. A final estimate for that crop is then made the following December — 1 year later.

The SRS periodically utilizes questionnaires mailed and received from about 150 000 U.S. farmers and follows up with actual farm visits to survey over 16 000 randomly selected samples. While these estimates are quite accurate at the national level, the statistical design does not provide such high accuracies at state levels and below. An analysis of historic SRS survey forecasts for wheat in comparison to the SRS final survey estimates* indicates that at the national level early in the season, prior to June 1, roughly the heading to mature period for U.S. winter wheat, the SRS forecasts of wheat production were to within ± 10 percent of their final estimates in about 85 percent of the years from 1966 to 1975. The SRS estimates made after July 1 were always within ± 10 percent of the final figures, and in most years, were to within ± 2 percent of the final figures.

However, at the state levels and below, the figures are not as accurate. To obtain accurate figures at these levels using the same approach would greatly increase the expense of the existing SRS forecast system. The SRS is currently investigating the use of Landsat data as a cost effective aid to improve the precision of estimates at the lower geographic levels [1].

The most accurate of the estimates for countries other than the United States are those for Australia and Canada, the other two major wheat exporters, with the USDA at-harvest estimates of Canadian wheat production being to within ± 10 percent of the final Canadian figures in about 90 percent of the years from 1966-1975. For Australia, the USDA at-harvest estimates were to within about ±10 percent in only 80 percent of the same years. For both these countries, however, the pre-harvest estimates are much less accurate. USDA pre-harvest estimates of wheat were to within ±10 percent in roughly half the years in Canada and for about one-fourth the years in Australia. The estimates for the wheat crops of two major importers, the Union of Soviet Socialist Republics (U.S.S.R.) and India, were less reliable. The USDA at-harvest estimates of wheat were to within ±10 percent for only about one-third of the years in U.S.S.R. and somewhat over half the years for India. For most foreign countries the very early season estimates are to within ±10 percent in about 25 percent of the years.

The frequency and magnitude of these earlier-season-to-harvest differences can be explained in part by the fact that the early-season estimates assume that historic trends in weather and planting patterns will prevail. Generally, these estimates are based on reports of planted hectarage by national governments and the historic value for average yields. Because the weather patterns differ so widely from year to year, the chance in any one year for weather conditions being very near the average (or normal weather) is not very large. Because hectares planted, the fraction of hectares actually harvested and the resulting yields from the hectares harvested are so critically dependent on weather patterns, there is a correspondingly small chance that actual hectarage, actual yield or actual production will be very close to average or normal values.

This leads then to a discussion of the precise manner in which government policy, economics and variable weather patterns affect the hectarage planted, the hectarage harvested and the average productivity of the harvested hectarage, i.e., yield for harvested hectares. As a first step in this discussion, it is wise to review a bit of terminology, since often the term yield is used interchangeably with production; in addition, the term hectarage must be carefully defined as to whether planted hectarage or harvested hectarage is being discussed. In crops such as wheat, the quantity of interest is tonnage

*These final estimates are made several months after harvest and are considered the most accurate information concerning U.S. wheat production.

produced from hectarage actually harvested. Hectarage planted, but not harvested (abandoned) for one reason or another, will produce some grain, but of course will not contribute in the market place. Yield for harvested hectares is an average productivity for the hectares harvested. Yield for planted hectares is defined as the production from harvested hectares averaged over all planted hectares. In contrast, biological yield is the potential average production from all hectares planted including the lower yields from those hectares abandoned. In existing reporting systems the quantity of importance is production from harvested hectarage. However, reporting systems all make separate estimates of hectarage planted and hectarage harvested, as well as yield, and combine these estimates to estimate production. In other words, forecast production is inferred from individual estimates of hectarage and yield.

Weather is extremely variable over a geographic region; western Oklahoma may be relatively dry while the eastern half can be experiencing favorable conditions. To get an acceptably accurate forecast of production, it is critical to associate the right weather with the actual hectarage being affected. Where the effects are so severe as to remove hectarage from production, as in the case of winterkill or severe drought, then, this reduced hectarage must be accounted for. It can be shown that the estimate of yield at a country level is directly dependent upon the geographic distribution of hectarage actually planted and then harvested. Therefore, not only must a survey system monitor the total hectarage harvested, it must also monitor the geographic distribution of the hectares harvested as well as the associated geographic pattern of weather and other growing conditions. In addition, the economics of the region for which planting and harvesting decisions are being made can also affect the average yield since economic factors to a large extent determine the minimum value of yield for a field at which the farm operator makes a decision to abandon, i.e., not harvest the field. Thus, the poorer the weather, the more likely is planted acreage to be abandoned; however, the decision to abandon will be based to some extent on the cost of harvesting, in comparison to the benefit of doing so. Thus in addition to the factors discussed earlier, the average yield for harvested hectarage is dependent upon the degree of abandonment. In a region with a marginal crop, the average yield per harvested hectare would be lower if the farm operator decided to harvest all hectares than if a decision were made to simply abandon the fields with lower yields. It is these facts which forge an inseparable link between yield and acreage and require a monitoring system which considers both.

3. LACIE, ITS PURPOSE AND ACCOMPLISHMENTS

3.1 EXPERIMENT OVERVIEW

The Large Area Crop Inventory Experiment (LACIE) was initiated in 1974 as a "proof of concept" experiment to assimilate remote sensing technology developed over the previous decade, apply a resultant experimental system to the task of monitoring a singularly important agricultural commodity over the world, modify the approach as necessary and conceivable and, finally, demonstrate the technical and cost feasibility of global agricultural monitoring systems.

The roots for LACIE were carefully and intentionally established in 1960 by the Agricultural Board of the National Research Council. By late 1962 experiments had been designed to examine the feasibility of utilizing multispectral remote sensing for agricultural crop monitoring. An organized research program was established by the USDA and the National Aeronautics and Space Administration (NASA) in 1965. This program led in an orderly fashion from the first successful computer recognition of wheat using multispectral measurements collected with aircraft in 1966 to follow-on development and testing of a satellite capability by 1972. Successful feasibility investigations in 1972-1973 were conducted with the Earth Resources Technology Satellite (ERTS)* which led to the design and initiation of LACIE in 1973-1974.

LACIE was designed to test and evaluate the use of remote sensing to estimate (on an experimental basis) wheat production over important producing regions of the world. Timeliness and accuracy goals were established in recognition of the essential requirements for global agricultural information. The experiment was designed to establish the feasibility of acquiring and analyzing Landsat data within a 15-day interval. Importantly, the at-harvest estimates were to be within 10 percent of the true estimate at the national level 90 percent of the time. An important additional performance goal was to determine how early in the crop year estimates could be produced and with what accuracy and repeatability. Also, the estimates were to be made with repeatable and objective procedures. Qualitative judgments were to be kept to a minimum. The experiment was scheduled to be conducted in three phases:

(1) In Phase I, the technology to estimate the proportion of regions planted to wheat would be implemented and tested and, similarly, the technique to estimate the yield from specific areas would be developed and tested.

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

(3) In Phase III, the modified technology would be tested and evaluated over a still wider range of geographic conditions.

The experiment was made of three major elements:

(1) A quasi-operational element to acquire and analyze Landsat and meteorological data to make experimental estimates of production.

(2) An off-line element to test and evaluate alternative approaches as required to meet the performance goals of the experiment, and

(3) An element to research and develop alternative approaches.

The experiment has been jointly conducted by personnel from NASA, USDA and NOAA (National Oceanic and Atmospheric Administration) of the DOC (Department of Commerce). They represent the many disciplines (including physics, plant

---

*The ERTS was renamed Land Satellite 1 (Landsat 1). A second earth resources satellite, Landsat 2, identical to Landsat 1, is now in orbit.

pathology, engineering, agronomy, statistics and mathematics, soil sciences, economics, and plant physiology) important to meeting the objectives of the experiment.

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

3.2 THE LACIE TECHNICAL APPROACH

The LACIE approach utilizes the direct observational capabilities afforded by Landsat together with estimates of weather variables to estimate production. This approach requires that each geographic subregion (selected to be relatively homogeneous with regard to wheat hectarage and yield) in a country be monitored to (1) forecast the quantity of wheat hectares available for harvest (both winter and spring individually in each subregion), and (2) to forecast the expected productivity for each subregion (yield) of the hectares available for harvest. The total wheat production for each subregion is then obtained by the product of available hectares for harvest and yield for harvested hectares. The production forecasts for all subregions are then summed to obtain the country-level forecast. In addition, the subregional forecasts of hectares for harvest are summed to obtain a forecast of national hectares for harvest. An average yield for all hectares harvested nationally is then obtained, which is by definition the hectarage-weighted average of subregion yields. This hectarage-weighted average yield is a desirable estimate to have since, when multiplied by the national hectarage, it will reproduce the national production estimate.

Within each of the subregions described in the opening paragraph, Landsat multispectral data is collected by Landsat each 18 days from 5 x 6 n.mi. segments randomly drawn from each stratum. Within each segment, wheat is distinguished from non-wheat by monitoring the temporal development of wheat from planting through harvest. The areal percentage of wheat in each segment in the stratum is then estimated and thereby an average percent for the stratum can be determined. The average areal percent wheat can then be multiplied by the total agricultural hectarage* in the stratum to estimate total wheat for the stratum.

As an example of the information contained in the sequential Landsat coverage consider the sequence of imagery shown in figures 1a, 1b, and 1c. This imagery was acquired over a LACIE 5 x 6 n.mi. segment located in Sherman County in the northwest corner of Kansas. Annotated on this sequence are selected wheat and non-wheat fields. Note that as the wheat begins to emerge, it appears on the color-IR imagery (computer generated from digital magnetic tapes) as a pink response indicating the increasing reflectance in the near infrared channel monitored by Landsat. Following winter dormancy, the February imagery indicates that all fields have survived the winter without loss to winterkill, although there is some bare soil spottiness in several of the fields. Note in the March, June, and July images the wheat is of varying degrees of vigor. Some fields are quite pink in the image indicating vigorous growth, while others are quite mottled indicating large areas with little or no vegetative cover. In the July 1 image, harvesting has begun in the outlined circular field (left center of image). In the final image acquired only 18 days later, the harvesting of wheat is almost complete as indicated by the bright signature and none of the fields appear to have been abandoned. In addition, the spring crops which had begun to emerge in mid-June are now quite vigorous on the image.

---

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

The areal proportion of the 5 x 6 n.mi. sample segment standing in wheat and harvested is estimated using a manually assisted machine processing technique. Through Phase II* the temporal and spectral information described above was used by an analyst to identify several blocks of data within the 5 x 6 n.mi. image as either wheat or non-wheat. These blocks of data, called training areas, were then submitted to the computer which read the digital magnetic tapes and estimated the statistical distributions of the radiometric (Landsat pixels) measurements acquired within the wheat and non-wheat training areas. The radiometric measurements recorded for the non-training area portion of the segment were automatically classified as wheat or non-wheat by a machine processing algorithm which computed the likelihood that the radiometric values of each pixel was more representative of the distribution determined for wheat than for non-wheat. These distributons are more commonly referred to as signatures. The end result of this manually aided machine analysis was an estimate of what fraction of the total Landsat pixels within a segment corresponded to wheat. This fraction was then assumed to represent the correct areal percentage of wheat contained by the 5 x 6 n.mi. segment. Figure 5 is a display of a classification map of the April 26 Saratov segment shown in figure 2b. The grey picture elements are those Landsat measurements which the computer classified as wheat. The infrequently occurring black elements represent Landsat measurements which were dissimilar from all of the measurements recorded over the training areas. All other crop classes are represented as white. The analysis process described was repeated for each 5 x 6 n.mi. sample segment in the LACIE survey region.

The errors associated with this technique derive from the fact that certain other crop types mimic wheat, both in its growth cycle and its appearance at each time in the growth cycle. Such crops are referred to as confusion crops. In addition, the Landsat spatial resolution of approximately 1.1 acre, introduces error in measurement on field boundaries, particularly in agricultural regions which have field dimensions on the order of Landsat's resolution. Results of LACIE to date have indicated that the major confusion crops to wheat are certain small grains, particularly spring barley and winter rye. In subregions where these confusion crops are in appreciable abundance, LACIE has identified total small grains and reduced these estimates to wheat estimates using historic relative abundance figures for these crops to wheat. These will be discussed in more detail in the results section.

From the Sherman, Kansas, acquisitions, the following characteristics of the Landsat estimates of harvested wheat hectarage can be noted: (1) It is both the spectral differences over time and at any one time between wheat and other crops which permit wheat to be identified and its hectarage estimated. The estimates are made for wheat that is emerged and detectable. (2) Wheat areas subjected to weather conditions so harsh as to result in disappearing hectarage, such as represented by areas of bare soil or extremely sparse vegetation in Landsat data, will not contribute to the LACIE estimate of wheat hectarage and thus will reduce the LACIE estimate of production. In this way Landsat data accounts for severe conditions in the production estimates. (3) Landsat data can be used to monitor abandonment. For example, if a field identified in the early winter (November-December) timeframe does not re-emerge following dormancy in January-February, hectarage loss to this factor can be identified. (4) In early season, LACIE estimates only the detectable (pink) wheat hectarage as opposed to planted wheat hectarage. Generally, a minimum of 20 percent ground cover is required before wheat is detectable. As the season progresses, the wheat hectarage detectable by Landsat will increase and converge in early season, following complete emergence, as will the LACIE wheat hectarage forecast, to the total standing** hectarage potential for harvest.

---

*A significantly modified procedure is being implemented in Phase III. See Section 3.6 for discussion.

**By way of contrast most existing systems measure total field hectarage including bare spots. Generally total field hectarage will exceed standing hectares of wheat.

As a remote sensing system such as LACIE begins to develop an image data base with several years of data, the estimates of area will become increasingly accurate, and additional information can be gathered regarding potential yield. Consider, for example, the two-year sequence of Landsat imagery acquired in the Saratov region of the U.S.S.R. as shown in figures 2a and 2b. Figure 2a displays the sequence of imagery acquired in the 1974-1975 crop season and the identities of the outlined fields. In figure 2b is the sequence for 1975-1976 crop season. From this multi-year sequence we can see the effect of the crop rotation patterns. The fields labeled 1 and 2 were fallowed in 1974-1975 and were planted to spring (field 1) and winter (field 2) grains in the 1975-1976 crop year. Field 3 which had the characteristic winter grain development cycle in the 1974-1975 sequence has been planted to a row crop (possibly corn or sunflowers) for the 1975-1976 crop season. In addition fields 4, 5, and 6 show rotation patterns, respectively, of winter grains ('75) to winter grains ('76), spring grains ('75) to winter grains ('76) and winter grains ('75) to spring grains ('76). Since rotation patterns are known to affect productivity of the fields, it can be easily seen how multi-year spectral data contains information related to crop yield.

There is, in addition, within any one year, potential information in the spectral data related to crop condition and thus yield. Note in the 1974-1975 sequence of the Saratov segment that the apparent vigor of the fields declines drastically from the May 21 image to the June 8 image. In this year, this particular area was undergoing a significant drought. The impact of no moisture is clearly evident on this detailed image, as it is in the two full-frame images of figures 3 and 4 acquired in this same area in the U.S.S.R. Note the image of figure 3 which was acquired in the 1975-1976 crop year with adequate soil moisture and then the image acquired over the same area 1 year earlier. The lack of vigorous response in the 1974-1975 image is indicative of the serious moisture shortage.

As of this date, the Landsat multispectral data cannot be used to completely quantify the reduction on yield of soil moisture deficiencies and other such episodic events which affect the spectral reflectance. Of course, if such events are severe enough to cause abandonment of hectarage, this would be detected in the Landsat data and the resulting decrease in the hectarage estimate would decrease the estimate of total production. At the present time, the spectral data is used only to monitor the geographic extent of the episodic events and the regular LACIE analyses are used to quantify the impact of these events on yield and production [2,3]. Research efforts are underway to utilize the spectral data directly to estimate yield.

Another very important way in which the spectral data contributes directly to the estimation of production is by monitoring year-to-year fluctuations in total hectarage. Apparent from the multi-year image sequence of figures 2a and 2b is the fact that there are significant changes in hectarge from year to year within a segment. As discussed in the first section of this paper, such changes affect yield also.

The yield for harvested hectares is forecasted in LACIE through the use of regression models which utilize weather-related variables obtained from the ground-based stations of the WMO network. These models [4, 5, 6] are referred to as agrometeorological models. The first-generation models currently used in LACIE are developed around monthly averages of temperature and precipitation. In the United States Great Plains yardstick area there are both winter and spring wheat models, covering the 12 areas designated in figure 6. The yield and climatic data base used to derive the U.S. models is approximately 45 years in length. The yield data is obtained by aggregating the USDA/SRS estimates of harvested acreage and production to obtain yield in bushels per harvested hectares, individually, for both winter and spring wheat, in each of the 12 subregions. The climatic data consists of monthly climatic division averages of precipitation and temperature. These averages are weighted using hectares harvested to obtain the monthly average temperature and total precipitation for a given region. A piecewise linear trend is used to model the technology trend.

Yield models must cover a wide range of climate found in the U.S.S.R. whose wheat growing region spans over a thousand miles from north to south; covering the 33 crop regions of figure 7 are 15 winter wheat and 16 spring wheat yield models. Winter wheat is grown primarily in European U.S.S.R. Since 1949, both spring and winter wheat have shown an upward yield trend. Factors contributing to improved yields include improved varieties, increased mechanization, greater fertilizer use, increased irrigation and applications of pesticides. Winterkill and moisture stress are two major weather hazards that reduce both harvested acreage and harvest yields.

3.3 PHASE I SYSTEMS PERFORMANCE

In Phase I, the experiment was successful in piecing together from existing hardware and technology a system capable of processing the required volume of Landsat data. Approximately 2604 acquisitions from 693 segments were analyzed. A total of 411 of these segments form a selected sample population representing the Great Plains yardstick region.

During Phase I, an average of about 12 hours of analyst "contact" time was required for analysis to derive a wheat proportion estimate for a 5- by 6-n.mi. sample segment. In the experimental one-shift-per-day, 5-day environment, an average of 30 days was required to move a segment from its moment of acquisition by Landsat through to a final wheat proportion estimate. In an operational environment, the required time could be reduced to less than 15 days.

Significant experience was acquired during Phase I in the analysis of Landsat data to estimate wheat area. A portion of that experience was gained locating and eliminating system and analysis problems. One analysis problem dramatically affected the early season LACIE area estimates until it was located during the latter half of Phase I. Bare soil was correctly classified as such but was erroneously aggregated as wheat acreage in early estimates. This led to high overestimates of wheat acreages in both the spring and winter wheat area reports until early acquisitions were replaced by later season data. Near the end of Phase I, this and other less significant problems were corrected, and a final analysis of all the Landsat acquisitions was completed with the result that the at-harvest estimates marginally satisfied the 90/90 criterion.

COMPARISON WITH SRS - RELATIVE DIFFERENCE + C.V.

Region		Area	Yield	Production
Large	Total yardstick region	-10.7% ± 5.7%	4.3% ± 2.3%	-5.6% ± 5.9%
	Southern portion of yardstick region	-0.13% ± 7.0%	4.2% ± 2.6%	4.95% ± 7.04%
Blind sites	• Insignificant relative difference • C.V. at segment level = 50% compared to 80% allowable			

PHASE 1 RESULTS ACHIEVED AT END OF PHASE WITH MODIFIED APPROACH
(These results are for a relatively "normal" agricultural year.)

As a result of an evaluation of the Phase I experience, significant changes were made for Phase II.

a. A requirement was instituted to have the complete analysis of a segment conducted by a single analyst or analyst team as opposed to having a series of different analysts individually perform the image analysis, the machine processing and the evaluation functions required to develop and check a proportion estimate for a 5- by 6-n.mi. sample segment. The team approach afforded analysts an opportunity to develop an understanding of the interactions of the various analysis procedures, thus leading to a more accurate final estimate.

b. Every cloud-free acquisition of each sample segment was to be analyzed as opposed to utilizing one acquisition in each of four different biowindows. This change was required because of the uncertainty of estimating the biowindow of wheat at a specific time as well as a lack of understanding of the best times to differentiate wheat from other confusion vegetation.

c. Modifications to the sample strategy were made to compensate for a larger-than-desired sample error detected by accuracy assessment analyses. Full-frame Landsat imagery was used to refine the sample frame by deleting segments with no agriculture and reallocating them. In North Dakota where sample error was deemed excessive, 20 additional segments were allocated for Phase II.

d. Blind site ground truth proved to be an invaluable aid in problem diagnosis during Phase I. In Phase II, the Phase I blind site complement of 20 segments was increased to 140. Each blind site is visited twice a year by personnel of the USDA Agricultural Stabilization and Conservation Service (ASCS) and the identity of each field is established.

3.4 PHASE II

In Phase II, 9276 acquisitions over 1720 segments were collected and analyzed. The system was augmented with a computer parallel processor to support the increased processing loads. In addition, the average contact time required for the manual portion of the analysis of a sample segment was reduced from 12 hours of Phase I to 6 hours, as a result of more efficient analysis procedures.

In order to handle the increased data load, incurred by examining each acquisition, an additional type of analysis routine was used. A "no change" analysis routine required an analyst to overlay a computer classification map from a previous acquisition over a color-infrared image created from the new acquisition and manually determine if a change of more than 2.5 percent in wheat area had taken place. Given such a change, the segment would be reprocessed. The average time required for this was approximately 1 hour. As a result of inadequate amount of wheat training data in low-hectarage segments, a third type of routine required an analyst to manually interpret a color-infrared image made from the Landsat multispectral data and handcount wheat pixels where less than five percent of the sample segment was in wheat. The average time for this type of analysis was approximately 2-1/2 hours.

In Phase II, the LACIE system was successful in acquiring and processing the meteorological data from the WMO stations through the yield and crop growth models programmed on digital computers. Thirty-day average values of precipitation and temperature were utilized in the yield models in Phase II. Daily maximum and minimum temperatures were collected as inputs for the wheat growth-stage model.

LACIE experimenters were particularly interested in the repeatability of Phase I at-harvest results in the Phase II crop year. Also, the question of how early and how accurate wheat production estimates could be made prior to harvest was of primary interest. In addition, critical attention was placed on an evaluation of how well the yield models would perform in foreign regions where historic data was thought to be of lower quality than that of the United

States. Also in question was an issue of how well the models might perform under abnormal weather conditons that might occur in some part of the United States, Canada, or the U.S.S.R. The 1976 crop year did provide a departure from normal weather patterns in the U.S. Great Plains yardstick region. Except for the months of November 1975 and April 1976 the crop year was very dry. It should also be noted that much of the above-average November precipitation occurred at a time when the crop was entering dormancy.

Because of the severity of the drought conditions during Phase II, LACIE established an Episodal Events Team (EET) to monitor the development and intensification of the drought in selected U.S. regions that were initially identified by the use of meteorological data.

The objectives of monitoring the drought episodal event were: (1) to determine the extent of the 1975-1976 drought in the selected regions, (2) to determine the effects of this drought upon area, yield, and production of wheat, and (3) to develop procedures for monitoring drought using remote-sensing-based criteria. The data developed by the EET investigation were not to be used directly in Phase II analysis and production estimates.

The technical approach of the EET involved the use of Landsat images of 5 x 6 n.mi. LACIE sample segments, and full-frame (100 x 100 n.mi.) imagery at 9-day intervals to identify the drought area and quantify the effects on the wheat area. Yield model simulations were run to extrapolate the effects of the drought on yield estimates at harvest, assuming 10 to 90 percent of normal rainfall for subsequent months and 30-day forecast. A survey of Landsat data for improvement of distribution of rainfall patterns in the drought area was done for April and yield models were run for drought-affected crop reporting disricts (CRDs). While these procedures were not utilized in Phase II operations, special aggregations were performed for the drought area CRDs by the LACIE Crop Assessment Subsystem to evaluate the utility of remote sensing for monitoring the effect of the drought on wheat area, yield and production [2, 3].

3.5 ACCURACY OF SURVEY ESTIMATES

Results of LACIE to date are particularly encouraging in the winter wheat regions of the world where, in Phase I and II, the LACIE survey estimates have greatly exceeded expectations. The LACIE technology has produced encouraging early-and excellent mid-season estimates. In addition, the winter wheat estimates at harvest were more than adequate to support the 90/90 criterion. In fact, for the U.S. winter wheat yardstick region, the 90/90 criterion was exceeded for the June and later estimates (figure 8 and table 1). The June estimates were based on Landsat data acquired through the first week in May. Therefore, an operational system with a 14-day turnaround could have produced quite an accurate estimate in mid-May some 1-1/2 to 2 months prior to harvest. The LACIE estimates of area for harvest in the LACIE May 7 report, based on Landsat acquisitions acquired through early April, were to within four percent of the SRS May estimates for harvest — in addition, the coefficient of variation of the LACIE area estimate was supportive of a 90/90 production estimate even at a five-state level.

In the U.S.S.R. winter wheat indicator region (fig. 9 and table II), all indications point to survey estimate accuracies comparable to those in the United States. While the excellent yardstick estimates are not available* for comparison at the U.S.S.R. indicator region level, the computed confidence of the LACIE acreage survey estimates indicate accuracies supportive of the 90/90 criterion.

*The FAS estimates shown in figure 9 are derived from country level estimates, assuming a fixed hectarage ratio between the country and indicator region level. Analysis of hese ratios for the past 17 years indicates a year-to-year variation in this tio of about five percent. The differences noted in figure 9 are statistica. y non-significant.

Only one significant problem has been encountered to date in the winter wheat survey regions. During Phase II, Oklahoma and other states of the southern Great Plains, experienced generally dry conditions through April 1976. These conditions created poor wheat stands and subsequent acreage underestimates. In some cases, sparsely vegetated fields were not detected as "emerged" acreage in the Landsat or even the aircraft-ground-truth color-infrared imagery. The April rains greatly improved the wheat stands. However, the drought-altered growth cycle misled the analysts in late season to believe the late-recovering wheat to be a spring-planted crop. A tendency to underestimate wheat area in Oklahoma was not observed in Phase I, LACIE estimates being to within three percent of the SRS. Episodal events such as the drought-altered growth cycle in Oklahoma just described are a part of the learning process. As more of these situations are encountered, the technology will adapt to accurately estimate their impact on acreage, yield, and production. Phase III will see a greatly enhanced episode monitoring effort.

The results of 2 years in the U.S. northern Great Plains and 1 year in Canada (figs. 10, 11 and tables III, IV), indicate a greater tendency to underestimate spring wheat acreage in the western hemisphere than is seen for winter wheat. However, such a tendency is not observed in the U.S.S.R. for either spring (fig. 12*) or winter wheat. As was identified at the end of LACIE Phase I, some spring small grains cannot yet be reliably differentiated from spring wheat using Landsat data alone. Spectrally these crops are similar, as are their growth cycles. Therefore, until procedures could be developed and tested in Phase II for use in Phase III to improve discriminability of these crops, historic ratios of these acreages were used to reduce the Landsat estimates of total small grains to an estimate of wheat acreage. The use of these historic ratios introduced additional error into the spring wheat acreage estimates, paricularly in the Phase II crop year for which the planting of wheat in preference to non-wheat small grains had greatly increased from previous years. In many instances, the current ratios were as much as 60 percent greater than the historic ones used in LACIE. This was responsible for a significant amount of the underestimates of wheat acreage in Canada. There is, however, in addition to the ratio factor, a residual tendency to underestimate spring small grains acreage in the United States and Canada. This is verified by the comparisons of the Landsat estimates to ground-observed small grains acreage in the LACIE blind sites. The cause is thought to be partially a result of the greatly increased tendency toward strip-fallow practice in the spring wheat regions. Strip-fallow fields, small compared to the Landsat resolution, are difficult to detect and measure in the Landsat imagery (see figure 14). The absence of the U.S.S.R. spring wheat hectarage underestimation problem may be indicative of more stable year-to-year ratios of spring wheat to other small grains ratios (resulting from governmental controls) and a decrease in strip-fallow practice.

An additional dimension to the accuracy of the LACIE survey estimates is the period in the growth stage of wheat when the Landsat data is acquired. Generally, three distinct regimes emerge in this regard: (1) An <u>early season regime</u> when a majority of the Landsat data used in area estimation was acquired in the emergence-to-jointing period of wheat development, (LACIE Biowindow 1), (2) a <u>mid-season regime</u> when a majority of the data was acquired in the jointing-to-mature (green-to-senesence) period of wheat development, (LACIE Biowindows 2 and 3), and (3) an <u>at-harvest regime</u> when most of the data has been acquired through harvest (Biowindow 4). These periods are indicated on the abcissa of figure 8 and figures 9 - 12. Note that for each country, the area estimates steadily increase through the growing season. In the case of U.S. southern Great Plains winter wheat, the <u>early-season</u> area estimates are substantially below the final estimates. In fact, they are about as much below

---

*The FAS estimates shown in figure 12 are derived from country level estimates assuming a fixed ratio between the country and indicator region level. Analysis of these ratios for the past 17 years indicates a year-to-year variation in this ratio of about 45 percent.

the final estimate as the initial SRS area estimates (dashed line) are above it. The mid-season estimates increase substantially and are not significantly different than the final estimates. The at-harvest estimates increase just slightly and are somewhat more accurate, i.e., more in agreement with SRS/FAS estimates, than the midseason ones.

An analysis of ground truth and other data shows that this phenomenon is purely physical in nature and not merely a statistical artifact. In the early-season reports, when a majority of the Landsat data is acquired in Biowindow 1, the wheat plant sizes vary from about an inch to over a foot in height with percentages of field area in vegetative ground cover varying from almost none to somewhat less than 40 percent. Observations from ground truth indicate that fields with less than 20-percent vegetative ground cover do not provide a sufficiently "pink" response on color-infrared Landsat imagery. That is, sparsely vegetated fields are not discernable as vegetation by the analyst. Since the analyst procedures call for the identification of detectable wheat (as opposed to an estimate of wheat planted) these early-season estimates are low, a result of incomplete emergence of all wheat.

By mid-season, the wheat has completely emerged and the LACIE acreage estimates agree quite well with ground truth.

Results of comparisons of LACIE estimates to the 103 blind site ground derived estimates of wheat proportion indicate that there is a moderately large variation between these estimates at a segment level; however, this variation is sufficiently small to be more than adequate to support 90/90 estimates at the national level. Primarily this is because the variation of the aggregated estimate decreases in proportion to the square root of the number of segments used in the aggregation: a result of the statistical independence of the segment estimates.

There is, however, a tendency to underestimate the wheat area in the region as observed from ground truth. Of the 103 blind sites investigated in the southern Great Plains and the 33 in the northern Great Plains, a majority of the segments are underestimated to some extent. For segments with larger proportions of wheat there is a stronger tendency to underestimate as can be seen from figure 13a and 13b. As the growing season progresses toward harvest the tendency to underestimate decreases as a result of increasing wheat emergence (shown in table VI). In this table the average relative mean difference (RMD) between the LACIE/Landsat estimates and the ground based estimates of wheat proportions has been computed for the blind site acquisitions on which the LACIE wheat area estimates were based on the U.S. LACIE crop reports released monthly beginning February 1976 through the final estimate for 1976. The final column of table VI indicates the percent of the segments for which an underestimate was observed.

A review of the LACIE Blind Site Data on a field-by-field basis indicates that the majority of the segment wheat-proportion underestimation results from wheat signatures labeled as non-wheat in the manual analysis process. There are two major classes of wheat signatures which most frequently are mislabeled: (1) The first major class includes wheat signatures which were outside the range of the wheat signatures usually observed. Generally, these signatures were associated with very thin stands of wheat (in some cases drought affected or incompletely emerged) which appeared only faintly pink on the color IR image. Also in this first class were signatures for wheat fields developing either significantly ahead or behind their nominal development calendar, and highly variable signatures acquired from strip-fallow areas with field widths small compared to the Landsat spatial resolution of about 80 meters (see figure 14). (2) The second major class of mislabeled wheat occured for those wheat signatures which (for a particular combination of Landsat acquisitions) were also characteristic of non-wheat signatures. Much of the discrimination between wheat and non-wheat vegetation is based on the temporal differences observed between the wheat and non-wheat signature cycles over a complete growing season. With Landsat there is at lease an 18-day interval between observations: even greater periods elapse if cloud cover obscures the target on a particular overpass. Thus a given collection of Landsat cloud-free acquisitions may be inadequate to permit all of the wheat signatures to be uniquely associated with wheat.

The misidentification of abnormally developing wheat signautres should decrease as more experience is gained with the variety of growing conditions to which wheat is subjected from year to year. Another significant reduction in signature confusion between wheat and other crops will result from improved sensors. But regardless of how much experience is gained, or how good the sensors become, there will always remain a problem with labeling "confusion crop signatures", i.e., signatures which, for a variety of reasons, are not unique to a given crop. The labeling procedure utilized to date in LACIE can be described as a "wheat conservative" procedure. That is, a particular signature is labeled wheat only in case there is a high degree of confidence that the signature is uniquely associated with wheat. If the signature is not typical of signatures normally observed for wheat, or in a significant number of cases is also observed as a non-wheat signature, the signature will be labeled as non-wheat. This "wheat conservative" tendency is verified by examining the analyst labeling errors in the blind site data. The analyst labeled wheat fields are, in almost all cases, called wheat fields by ground observers; very rarely does an analyst label a non-wheat field as wheat. However the analyst labels a significant number of ground-observed wheat fields as non-wheat. The "wheat conservative" procedure obviously has a built-in negative bias. However, a "wheat liberal" alternative of labeling a signature as wheat if there was a reasonable chance that it might be wheat would lead to an overestimate of wheat. Therefore, the problem in dealing with non-unique or unusual signatures boils down to the following: How can such signatures be labeled in a manner which produces a minimally biased wheat proportion estimate?

The LACIE research, test and evaluation program is investigating a procedure which has two features: First the procedure includes a means for the analyst to specify quantitatively the certainty with which each signuature is uniquely associated with wheat or with non-wheat. Second, a method is being developed which permits this "figure of certainty" to be utilized in the proportion estimation process in such a way as to minimize the estimation bias resulting from non-unique signatures; however, for the near term, the LACIE design effort has focused on the development of (1) products and ancillary information which will increase the ability of the analyst to correctly identify wheat signatures (2) more automated machine processing procedures which will free the analyst from all non-essential manual functions so he may concentrate on signature labeling (3) more optimum machine processing procedures from the point of view of producing minimally biased proportion estimates given correct signature labels. This approach will be described in the next section.

A detailed analysis of the ground and meteorological information in the 1976 crop year indicated that the primary agrometeorological conditions responsible for acreage underestimation in Phase II were: (a) For the winter wheat region, the early drought in 1976 followed by late April rains created atypical growth conditions in which wheat signatures were not visible early in the year and then "greened" up later than expected. Many such fields were misidentified by analysts. The primary region affected by this problem was the state of Oklahoma and a portion of the Texas Panhandle. The Landsat estimates of wheat proportion agree favorably in the other southern Great Plains winter wheat states. (b) There is an increased tendency to underestimate in the Northern Great Plains spring wheat region (table VII). A more detailed investigation of these blind sites indicates strip-fallow fields, whose width is small compared to the Landsat resolution, to be a major source of the observed underestimation (see figure 14). These fields were difficult to classify with the Phase II procedures. In addition to the strip fallow problem, some of the same problems observed in the U.S. southern Great Plains winter wheat region were also observed in the northern Great Plains spring wheat.

As a result of these blind site investigations improved machine processing procedures have been developed and implemented for Phase III and will be discussed in the next section.

Regarding the performance of the first-generation yield models employed in LACIE, 2 years experience with the models and tests of them over 10 years of historic data indicate adequate performance in estimating wheat yields at the national levels of those countries for which adequate historic and current meteorological data are available. At levels below the national level, investigations have shown a need to improve the LACIE yield model's response to extreme weather conditions. In South Dakota, for example, 1975-76 was an extremely dry year with wheat yields estimated by SRS to be only 11 bushels per acre. The LACIE South Dakota yield model estimated 17 bushels per acre and would have estimated 13 bushels per acre even if zero values for precipitation had been entered into the model throughout the year. The tendency to over or underestimate yields in areas and in years for which there are large deviations from the average yield is common to overly simple crop-yield-model forms which cannot adequately reflect the total dynamic range of the response of the plant to its environment. A second-generation approach to yield modeling is being evaluated for selected regions in Phase III. The second-generation models employ improvements such as a versatile soil moisture budget (as opposed to precipitation input alone), response to moisture and temperature tied to actual development state and use of daily (as opposed to monthly) weather variables. In spite of difficulties inherent with the first-generation models, they have served LACIE well. In fact, from figure 8, it can be seen that the LACIE yields were quite variable from month to month. They reflected the early dry season by a reduction in the yields followed by a corresponding increase through harvest.

3.6 PHASE III TECHNOLOGY MODIFICATIONS

As discussed in earlier paragraphs of this section, substantial improvements in remote sensing crop surveys can be expected in the future. For Phase III, the highest priority lies with technology improvements for identifying spring wheat directly from the Landsat data. Procedures, utilizing improved analyst aids such as interpretation keys and displays of quantitative spectral data are being developed. In addition, econometric models for the prediction of wheat-to-small-grains ratios will be developed and tested in Phase III. These models will predict the current ratios of wheat to small grains resulting from influential factors such as historical crop and livestock patterns, current year growing conditions (available soil moisture, etc.), economic conditions, and prevailing government farm programs. In Phase III and the transition years beyond, LACIE will implement improved partitioning of the survey region into subregions which are climatologically and agriculturally homogeneous. Such partitioning will render sampling strategies more efficient and thus more cost-effective. In addition, the agrometeorological data compiled to effect partitioning will improve the understanding of the agrometeorological properties of the survey regions and thus improve the ability to correctly classify crop acreage and estimate yield.

3.6.1 Improved Machine Processing Procedure

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

As a result of this evolution, a nearly optimum automatic processing procedure has been developed and will be implemented by mid-Phase III of LACIE. The procedure can be described as nearly optimum in the sense that (a) the need for manual intervention is almost eliminated from the machine processing sequence, (b) every measurement in the scene, as well as the full dimensionality of the spectral data, is utilized in statistics computation prior to maximum likelihood classification, (c) with correct analyst determinations of crop identity for a very small sample of the segment, the machine processing procedure will provide an unbiased estimate of the segment crop proportion.

This Phase III procedure has automated many of the manual functions performed previously and incorporates many new features. Specifically, the important features are: (1) As shown in figure 14a, pixels (white dots) are randomly selected within the segment and presented to the analyst for labeling as wheat or non-wheat using image interpretation techniques. The analyst submits these labels to the machine which without further intervention by the analyst executes the remaining functions. (2) Machine clustering is performed to delineate the spectrally homogeneous modes within the multispectral/multidate segment data, and a color map is generated displaying the cluster groups (fig. 14b). (3) The spectral properties of these homogeneous groups are then automatically compared by the machine to the spectral properties of the randomly selected pixels which have been labeled with analyst-determined crop identifications. Based on its "closeness" or "similarity" to the labeled pixels each cluster is labeled wheat or non-wheat. In addition "conditional" clusters whose properties are significantly different from any signatures labeled by the analyst are automatically flagged for more intense examination. A color map is generated to display these conditional clusters. The unconditionally labeled wheat clusters are all displayed in a single color — the non-wheat clusters in different color, as shown in figure 14c. If later examination by the analyst of the spectral and spatial properties of these conditional clusters produces a non-concurrence with the label assigned by the automatic labeling logic, the analyst may then change the label. If the cluster comprises only a small part of the scene, as in figure 14c, he may assume that the automatic bias correction will account for any significant error introduced. Only in cases where significant numbers of conditional clusters occur would the analyst be required to resubmit the segment data for additional analysis.

Following the machine clustering and automatic labeling logic, the labeled clusters of all 22 500 scene pixels are characterized parametrically by the machine as multivariate normal distributions. Means and covariances are computed utilizing all measurements in each cluster. Each pixel is then machine classified as wheat or non-wheat utilizing a maximum likelihood decision rule. This machine processing algorithm sequence processes up to four temporal acquisitions of four-channel Landsat multispectral data. The four-channel, four-date Landsat data is treated by the machine as a 16-dimensional measurement vector. In case a fifth acquisition is obtained, a feature selection algorithm automatically selects the "best" three of the four acquisitions resident in the data base and replaces the "worst" acquisition by the incoming acquisition. Upon completion of classification, the frequency of agreement between the machine-assigned labels and the analyst-assigned labels is automatically computed from a comparison over a sample of analyst-labeled dots, independent of the dots utilized in automatic cluster labeling. This frequency is used by the machine to correct its wheat proportion estimate for bias resulting from causes such as automatic cluster labeling errors, etc. The frequency of agreement is also used as a performance measure, i.e., an indication of a need for possible rework.

The bias correction capability allows an incoming Landsat acquisition to be automatically processed utilizing analyst labels from an earlier acquisition. If the analyst reviews the labels and decides there has been no significant change in them, then an automatic estimate has been obtained utilizing more recent Landsat data with potentially improved spectral separability. Even should the analyst review indicate the need for a modest number of label changes, the estimate can be updated without reprocessing simply by utilizing the bias correction procedure to account for shifts in crop identities.

In summary, once the analyst assigns labels to each spectral class, the bias corrected wheat proportion estimate is obtained without further need for intervention on the part of the analyst. The analyst, in addition, receives many products which permit a quantitative assessment of the quality of the segment estimate. In many cases where problems are encountered, several diagnostic products are provided to the analyst to facilitate rework.

From an operational viewpoint, these procedures will be much less labor intensive than the first generation ones. Analyst "contact" time for segment analysis has been steadily declining from about 12 hours in Phase I, to 6 hours in Phase II and a projected 3 hours in Phase III with the new procedures, an efficiency increase of a factor of four from Phase I performance. In addition, the Phase III procedures should provide the analyst with improved and more repeatable decision-making procedures. The spectral differences between wheat and non-wheat, small grains and non-small grains as observable on multiple Landsat acquisitions have proven invaluable to LACIE analysts to manually identify wheat or small grains in order to train the classifier. Because of technical difficulties, however, not much use was made of multitemporal spectral data in the machine-processed estimates during Phase II.

3.6.2 Improved Sensors, Yield Models, and Sampling

Landsat C, to be launched in the near future, will have improved spectral range and spatial resolution in comparison to Landsat 2. This should significantly improve classification and area estimation accuracies. Improved yield models will also be implemented. These models include agronomic variables not now included but which are known to affect yield. In addition, the importance of these variables will be made a function of crop growth stage to reflect the changing importance of these different variables throughout the growing season. LACIE will also be monitoring episodic events more intensely to assess their impact on yield. Phase III will include an evaluation of a second-generation sample strategy. In addition, the first-generation strategy is being modified for Phase III. Two hundred U.S. segments have been added to the 400 existing Phase II segments. The Landsat full-frame data acquired in LACIE was also utilized to improve the sample frame by deleting segments which fell into areas with no agriculture and randomly reallocating them to agricultural areas. Over 700 such segments were relocated in the U.S.S.R. The first-generation sampling strategy is a stratified random strategy where the strata and sample allocations are based on historic data only. These strata are necessarily confined to political reporting boundaries. The second-generation approach utilizes Landsat full-frame imagery, along with climatological and soil information to develop the strata and to determine the optimal segment allocations to the strata. Such an approach was known from the outset of LACIE to be an improvement over the use of historic data, particularly in countries whose historic data is sparse. However, this approach was not possible to implement until only very recently because of the unavailability of Landsat imagery for foreign countries and the lack of techniques for discerning small grains on the imagery. A year and one-half of data collection by Landsat and a similar amount of image analysis experience in LACIE now makes implementation of such techniques possible.

4. A LOOK TO THE FUTURE

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

Early in LACIE Phase II, an effort was initiated to accomplish the transfer of technology to the USDA for further evaluation. This effort is now an approved follow-on to LACIE and is officially designated, LACIE Transition. The objective of LACIE Transition is the orderly transfer of proven technology to USDA facilities and personnel. In LACIE Transition, USDA will construct and operationally test a first-generation global information system capable of producing timely, reliable, and objective estimates of the global wheat supply. LACIE Transition will begin with the start of the 1977-1978 crop year and will conclude in 1981. As USDA begins an orderly, country-by-country expansion of its operational test system through 1981, the experimental system will be utilized to refine the wheat-inventory technology in important wheat producing regions and to validate the technology prior to transfer to the USDA.

In addition to the transition efforts, the technology developed in LACIE will be adapted to inventory production of other food and fiber crops. These include corn, rice, soybeans, and inventories of non-food crops such as forest products. It will also be adapted to monitor forage conditions within the world's important rangeland. This increased capability could conceivably be developed and incorporated in the mid-to-late 1980's in a second-generation global food and fiber monitoring system.

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

To be more specific, the LACIE approach to date has utilized primarily Landsat data to estimate wheat area for harvest and primarily meteorological data to estimate the average productivity, or yield, for each hectare harvested. In a sense, this separation is artificial; there is much information in the spectral data relating not only to total acreage but also to the plant population density within the acreage. There is, in addition, information relating to plant condition and, thus, average yield. Plant characteristics which can be measured well in advance of harvest are known to be correlated with final yield as well as the environment of the plant. Therefore, a model which includes the effects on yield of not only the plant's environment but also its physical characteristics (height and stand density) from which early yield estimates based on soil moisture may be made [7] will be a significant improvement over models utilizing only meteorological data. Potential quantitative connections through modeling, involve efforts which relate leaf area index to evapotranspiration [8], leaf area duration to yield and leaf area index to Landsat spectal response [9]. With the advent of thermal sensing on Landsat C, additional information will be available which is a potential predictor variable for crop yields [10].

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

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

Another arena for development within the near future is improved sensing and measurement of the basic predictor variables themselves. To date, LACIE has utilized first-generation earth-resources satellites along with meteorological data obtained from the WMO ground stations. With the advent of the second-generation earth-resources satellite, Landsat C, and the development of a capability to utilize environmental satellite data to obtain more complete coverage for temperature and precipitation estimates, the survey estimates should significantly improve. The LACIE analysis experience has indicated that the Landsat data itself contains information regarding temperature and moisture, as these factors are manifest in crop condition and loss of vigor resulting from drought (see figures 3 and 4). Parameters such as soil moisture or, alternatively, precipitation and temperature can probably be more reliably and accurately estimated from a combination of Landsat-type and meteorlogical satellites.

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

The NASA, together with the USDA and the DOC, is already beginning to look ahead and to plan a technology development program required to support the future implementation of operational global food and fiber monitoring systems. A methodology to best insure a suitable technology base, together with an adequate understanding of its use, needs to be developed over the next year or two and vigorously implemented, if its output is to be available for the mid-to-late 1980's.

5. REFERENCES

1. Pay, R. M. and Huddleston, H. F.: Illinois Crop-Acreage Estimation Experiment. Proc. Symp. on Machine Processing of Remotely Sensed Data, Purdue Univ. (W. Lafayette, Ind.), June 29-July 2, 1976; IEEE cat. 76, CH1103-1-MPRSD.

2. Thompson, D. R.: Results of LACIE Integrated Drought Analysis (Southern Great Plains Drought, 1975-76). LACIE-00424/JSC-11336, NASA/JSC (Houston), July 1976.

3. Thompson, D. R.,: Results of Drought Analysis (South Dakota Drought, 1975-76). LACIE-00437/JSC-11666. NASA/JSC (Houston), Sept. 1976.

4. Hill, J. D.: Wheat Yield Models for the United States. LACIE-00431/JSC-11656, NASA/JSC (Houston), June 1975; (addendum published Jan. 1977).

5. Hill, J. D.: Wheat Yield Models for the U.S.S.R. LACIE-00430/JSC-11343, NASA/JSC (Houston), Jan. 1976.

6. Hill, J. D.: Yield-Weather Regression Models for the Canadian Prairies, LACIE-00433/JSC-11658, NASA/JSC (Houston), Feb. 1976.

7. Ulanova, S. M.: Utilization of Air Inspection Data on the State of Winter Wheat over Large Areas for Harvest Forecasting. Meterorologiya, i Gidrologiya #8, Aug. 1970, pp. 64-72.

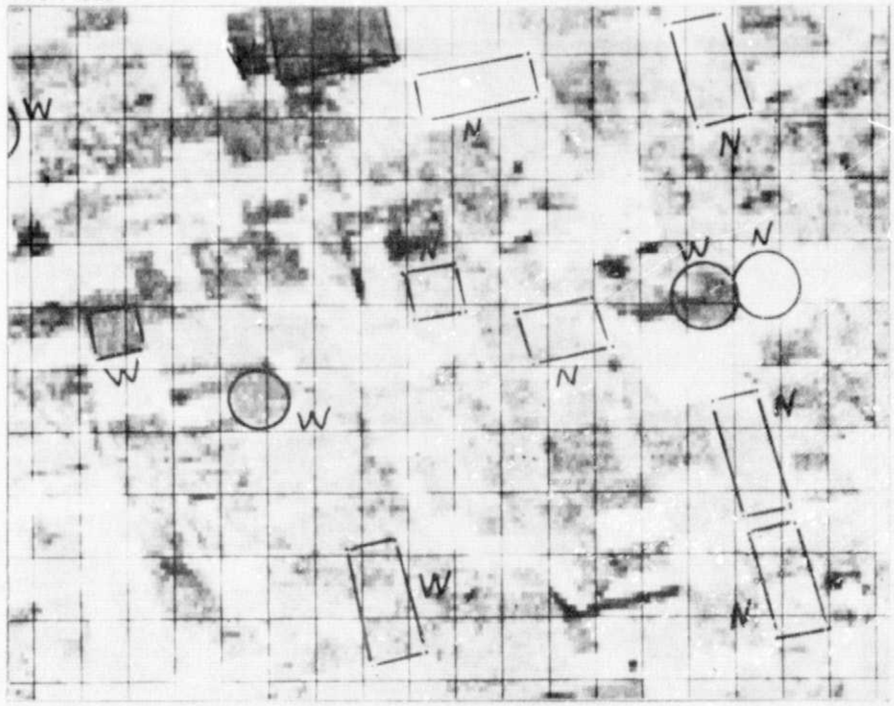
8. Haun, J. R.: Prediction of Spring Wheat Yields from Temperature and Precipitation Data, Agronomy Journal, vol. 66, May-June 1974, pp. 405-409.

9. Feyerherm, A. M., Kanemasu, E. T., and Paulsen, G. M.: Response of Winter and Spring Grain Yields to Meteorological Variation Final Report, Contract NAS9-14282, Feb. 1977.

10. Sherwood, B. I., Jackson, R. D., and Reginato, R. J.: Remote Sensing of Crop Yields. Science, vol. 196, April 1, 1977, pp. 19-25.

FIGURES AND TABLES

S-76-11502



10 NOV. 1975 EMERGENCE

S-76-11503

25 FEB. 1976 REGROWTH

25 FEB. 1976 REGROWTH

Figure 1a. Sherman County, Kansas. Segment 1021, 1975-76 winter wheat, Landsat computer imagery (W-Wheat, N-Nonwheat).

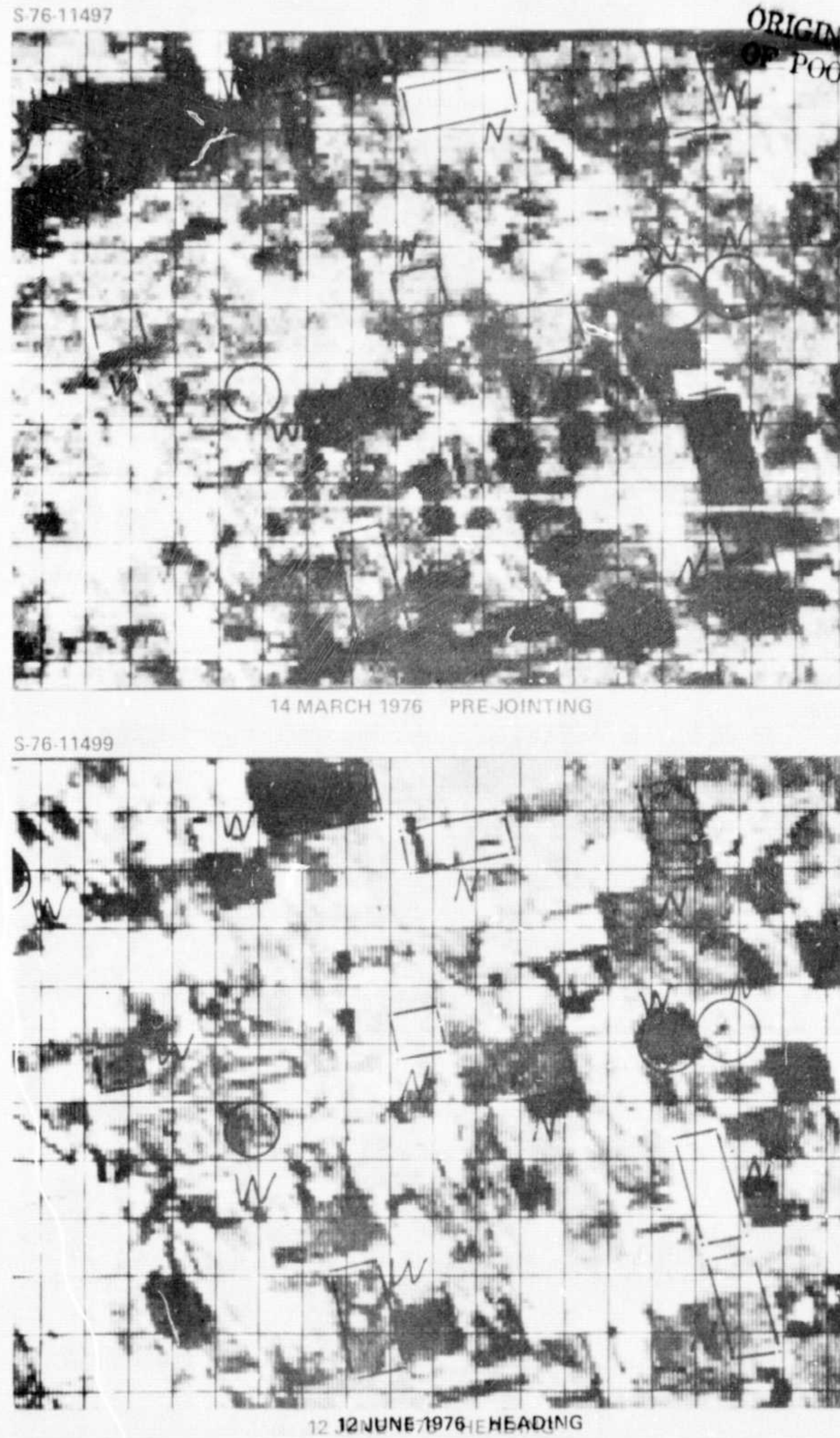
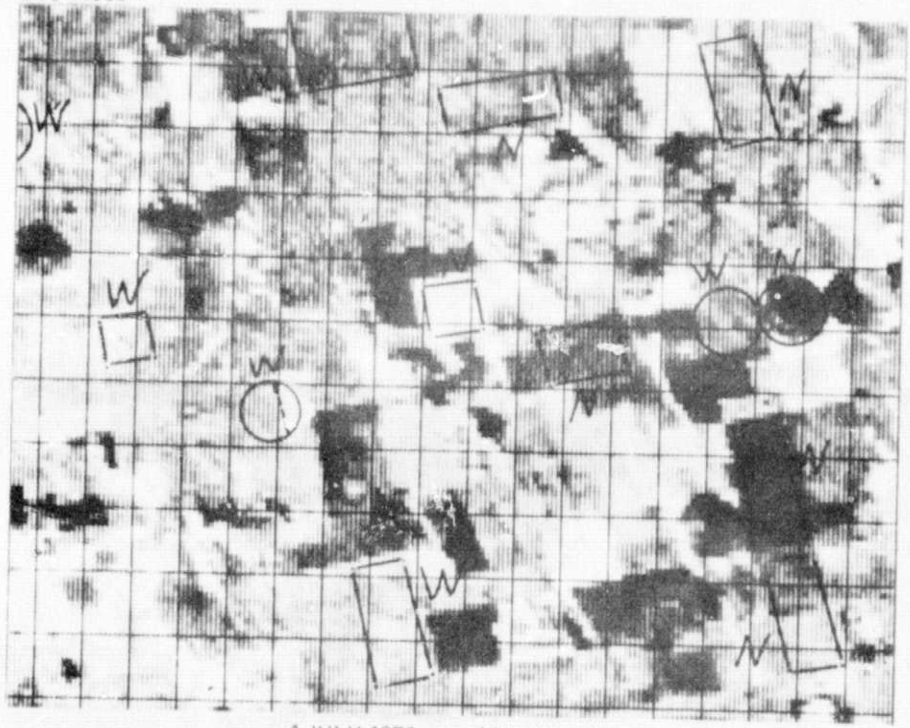


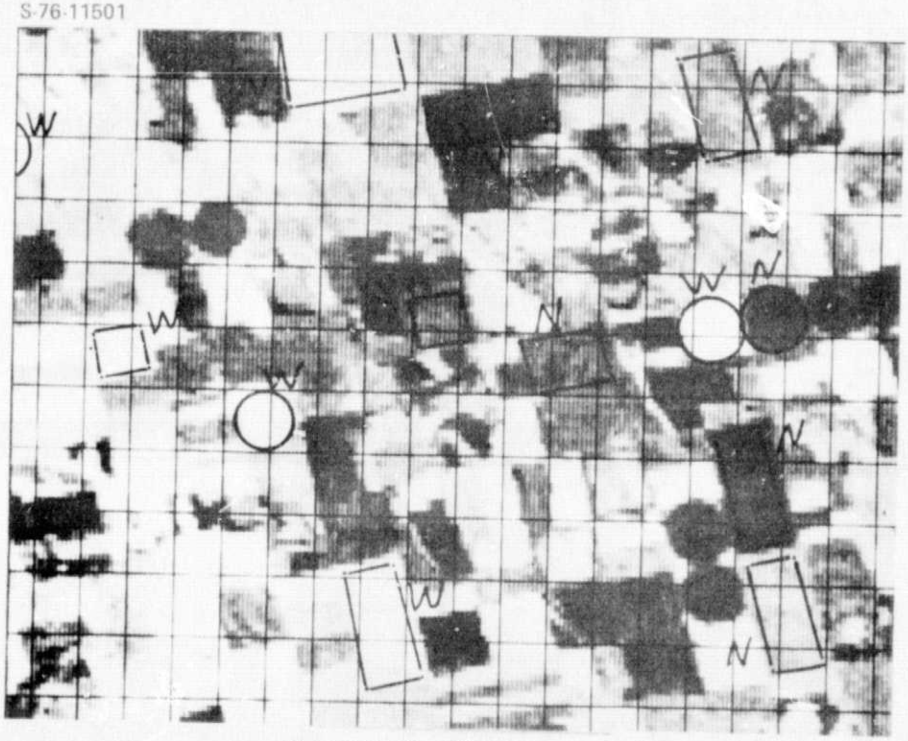
Figure 1b. Sherman County, Kansas. Segment 1021, 1975-76 winter wheat, Landsat computer imagery (W-Wheat, N-Nonwheat).

S-76-11500



1 JULY 1976 SOFT DOUGH

S-76-11501



19 JULY 1976 RIPE HARVEST

19 JULY 1976 RIPE HARVEST

Figure 1c. Sherman County, Kansas. Segment 1021, 1975-1976 winter wheat, Landsat computer imagery (W-Wheat, N-Nonwheat).

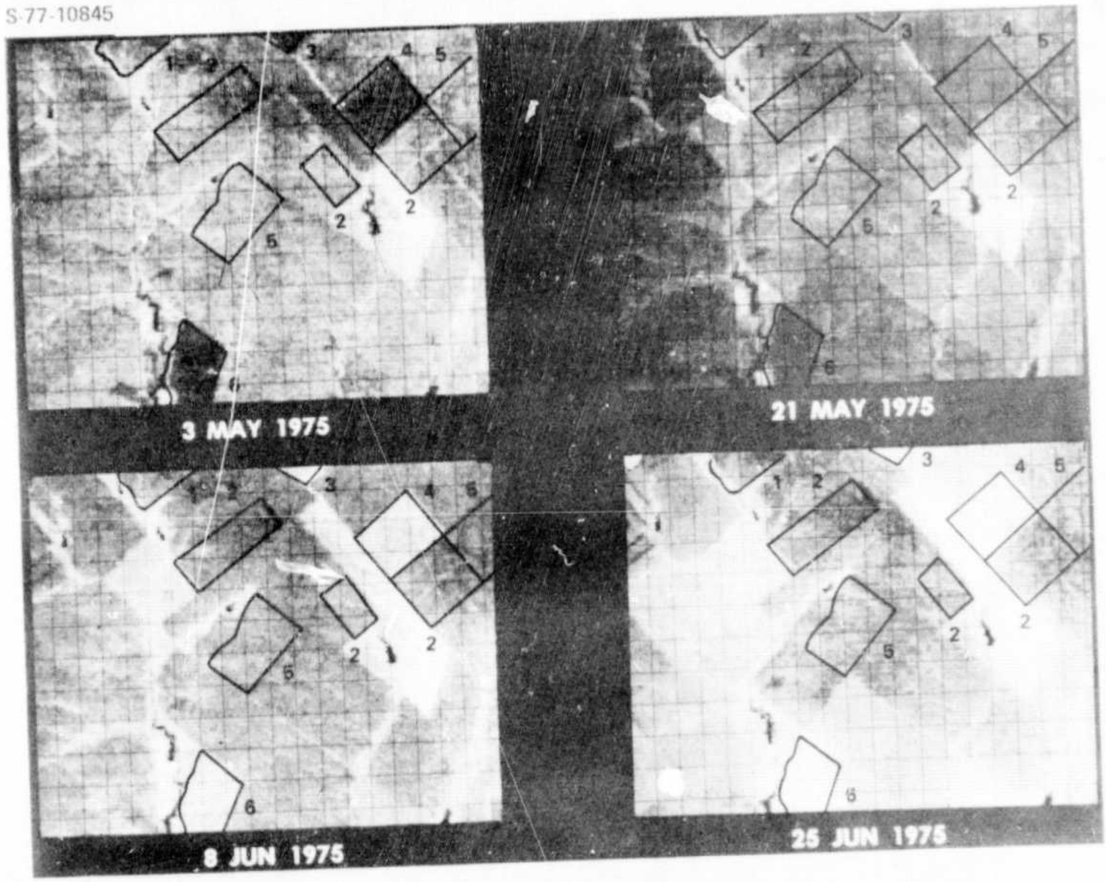


Figure 2a. Saratov, U.S.S.R., under drought conditions. Segment 7735, 1974-75 crop, Landsat computer imagery.

S-77-10847

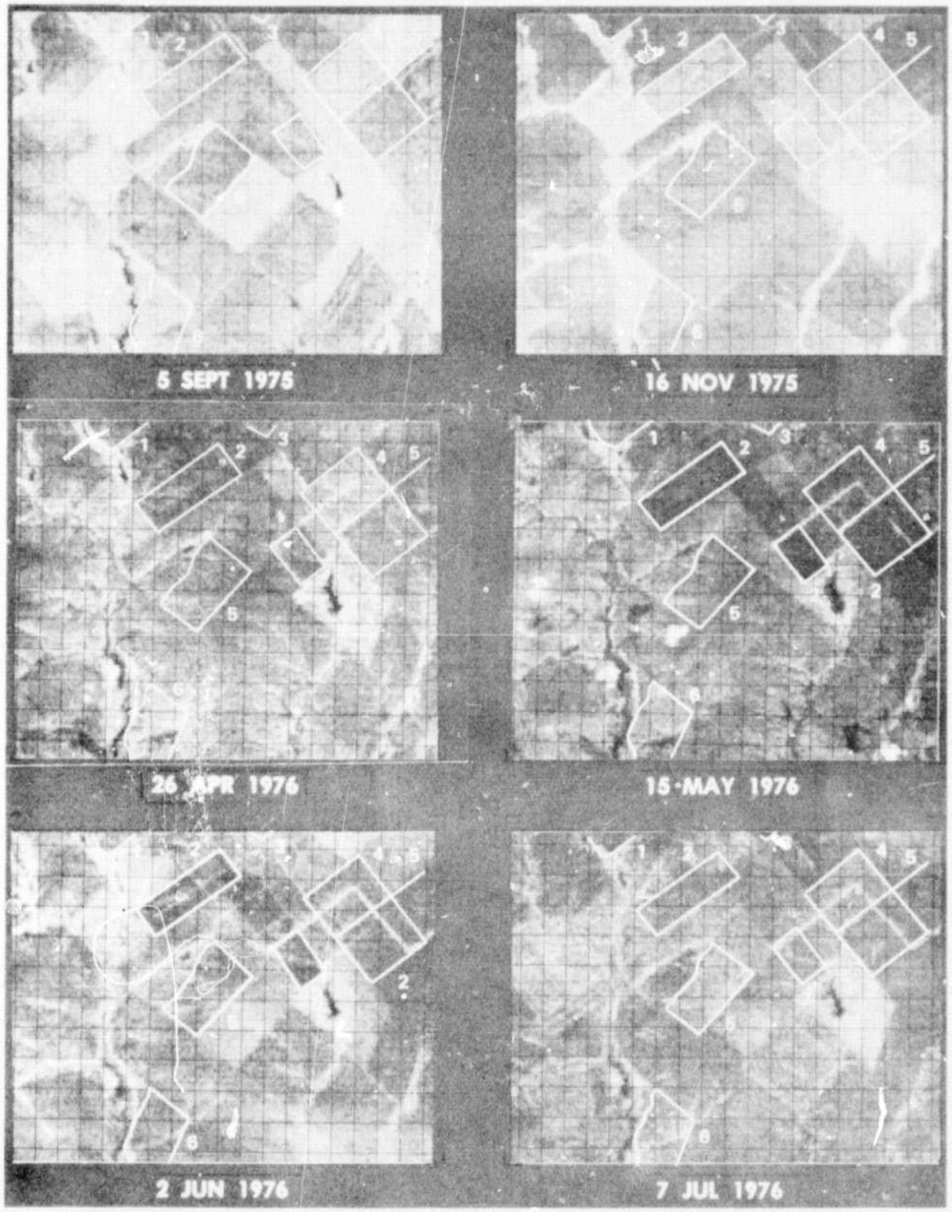


Figure 2b. Saratov, U.S.S.R., under normal moisture conditions. Segment 7735, 1975-76 crop, Landsat computer imagery.

S-77-10842

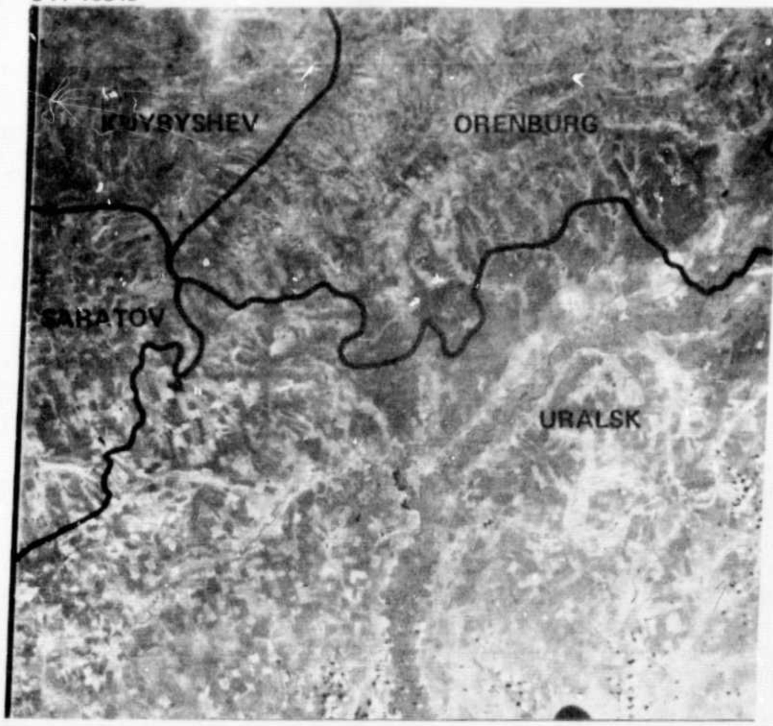


Figure 3. Saratov, U.S.S.R., region under normal moisture conditions. Full-frame Landsat image, 17 June 1976.

S-77-10841

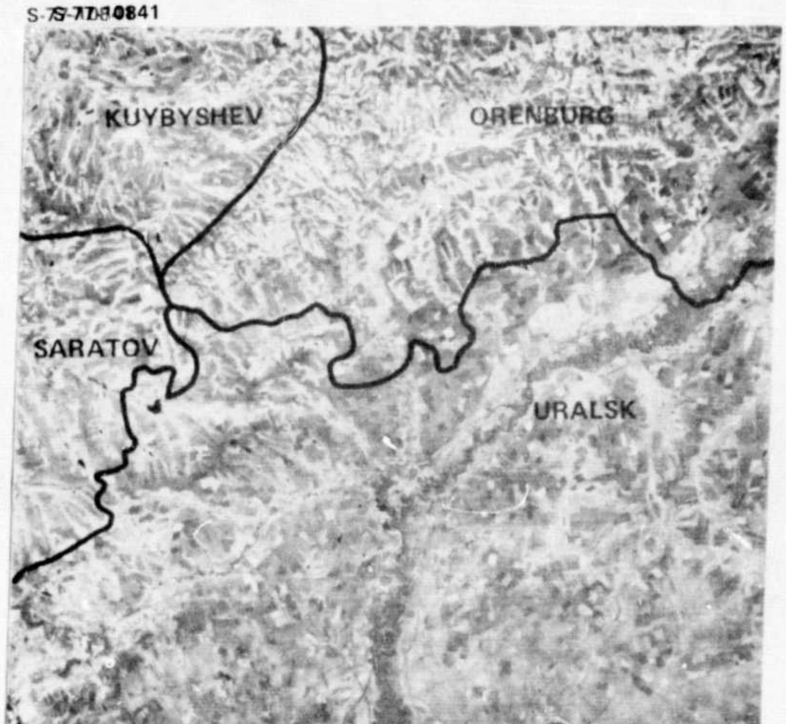


Figure 4. Saratov, U.S.S.R., region under drought conditions. Full-frame Landsat image, 23 June 1975.

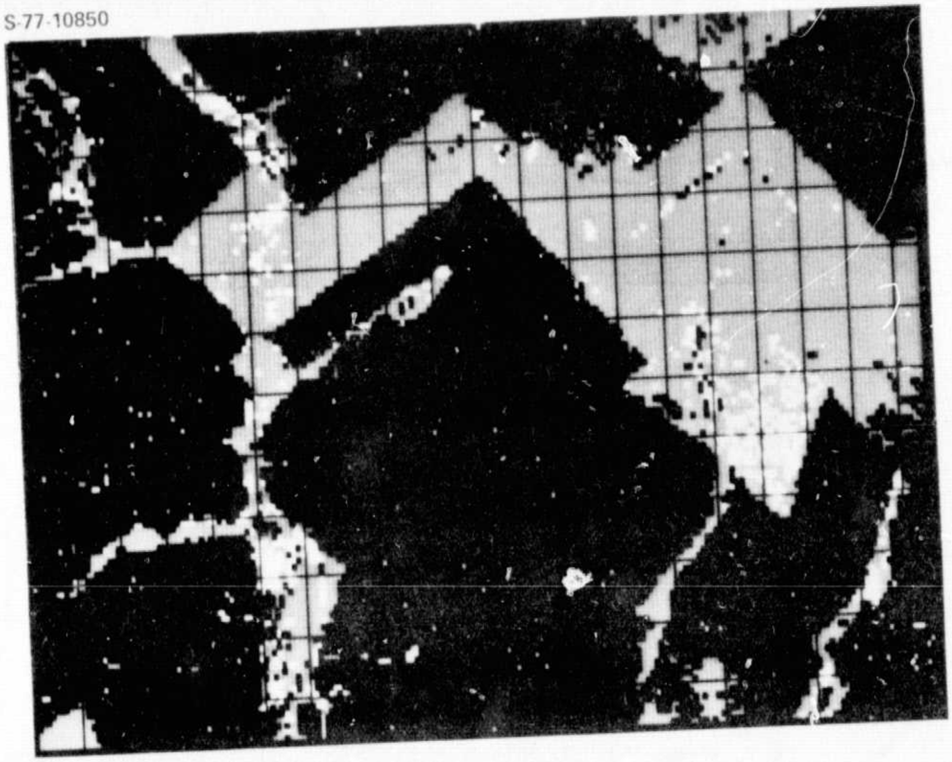


Figure 5. Computer-generated classification map for wheat and non-wheat. Corresponds to Landsat acquisition of 26 April 1976 over Saratov, U.S.S.R.

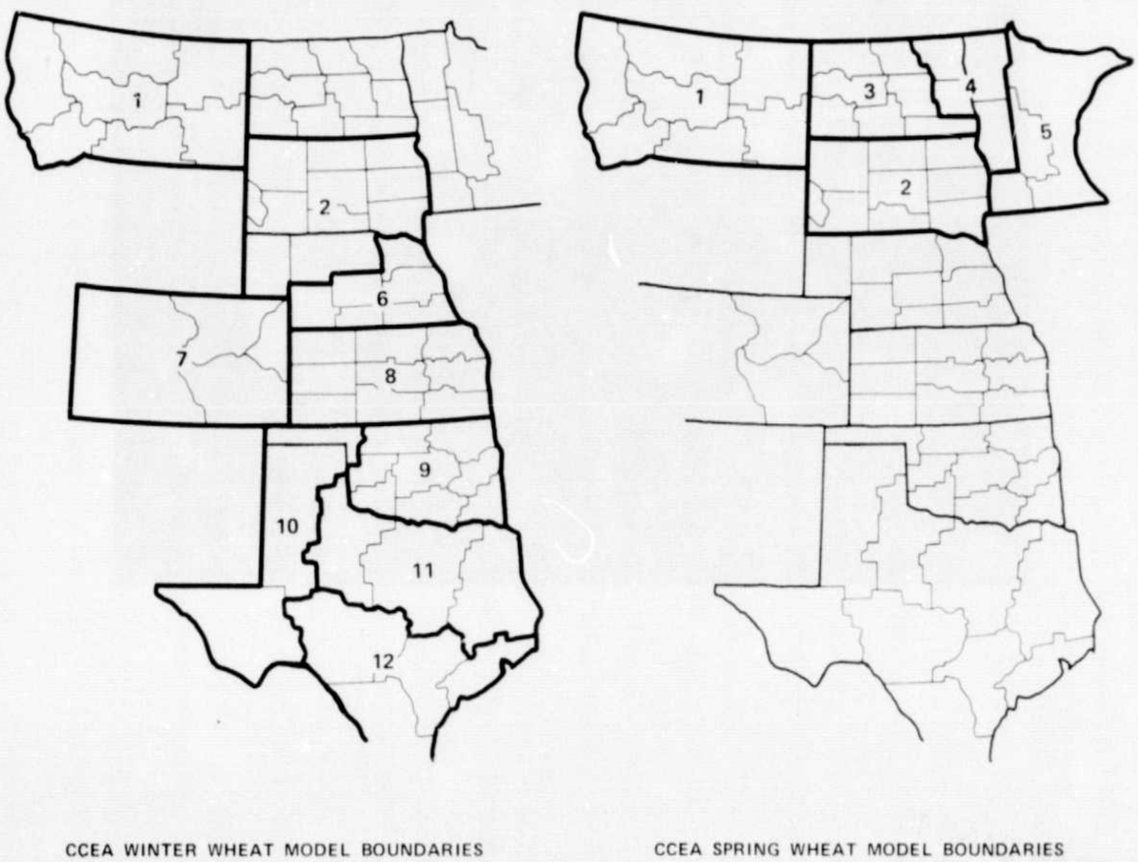


Figure 6. U.S. Great Plains wheat yield weather regression model coverage.

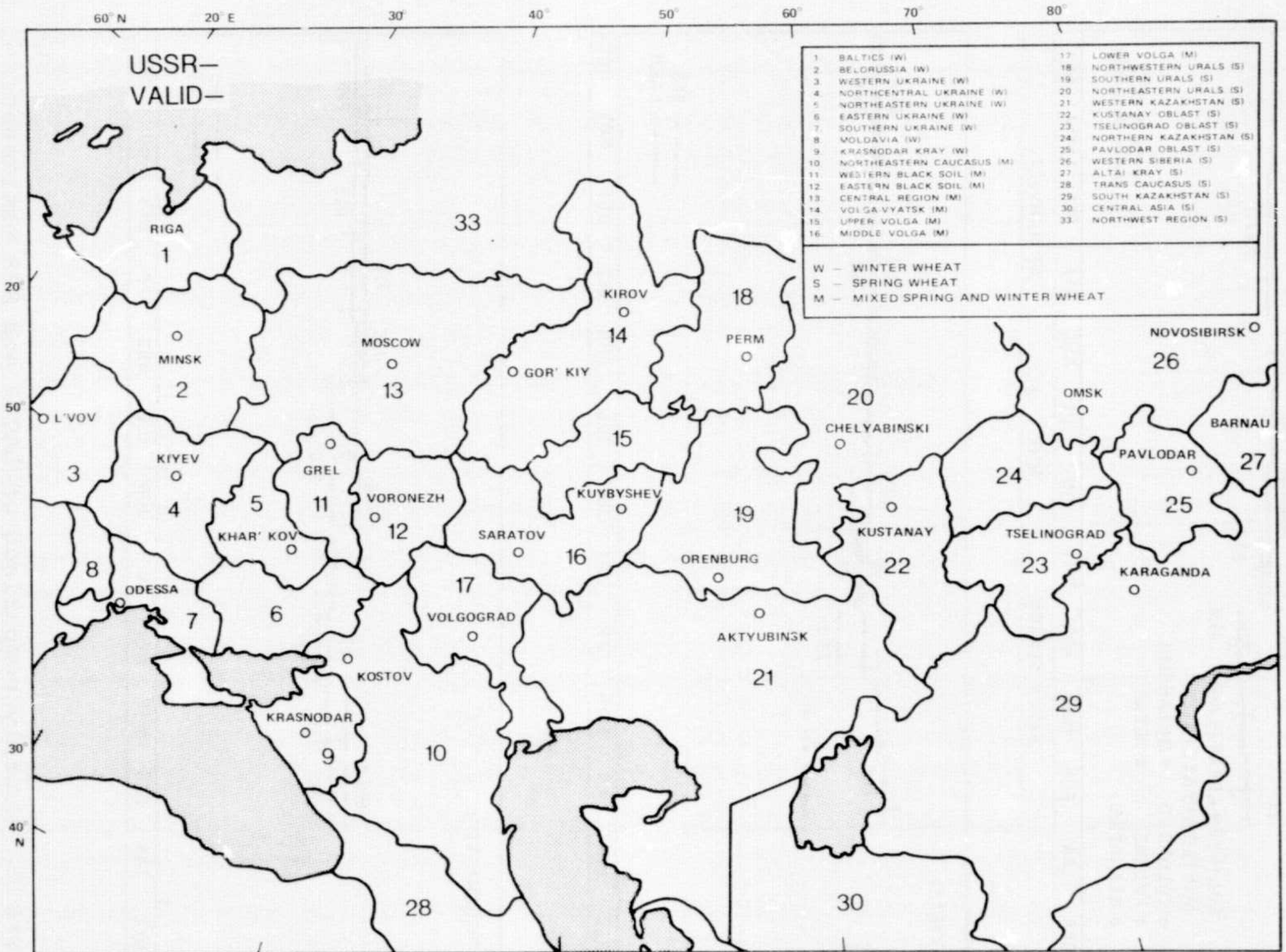


Figure 7. U.S.S.R. crop regions covered by spring and winter wheat regression models. Regions 31 and 32 not shown on map.

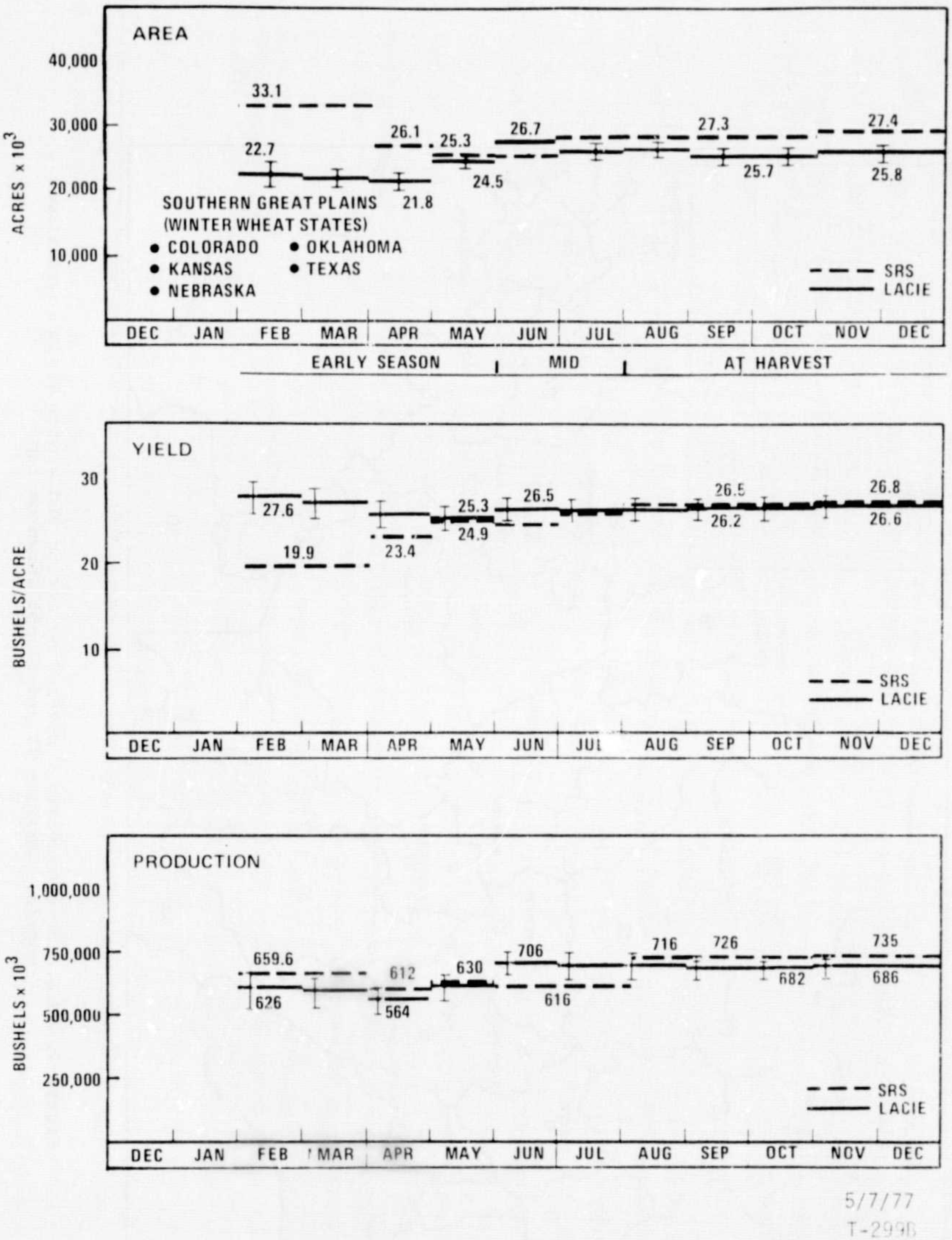


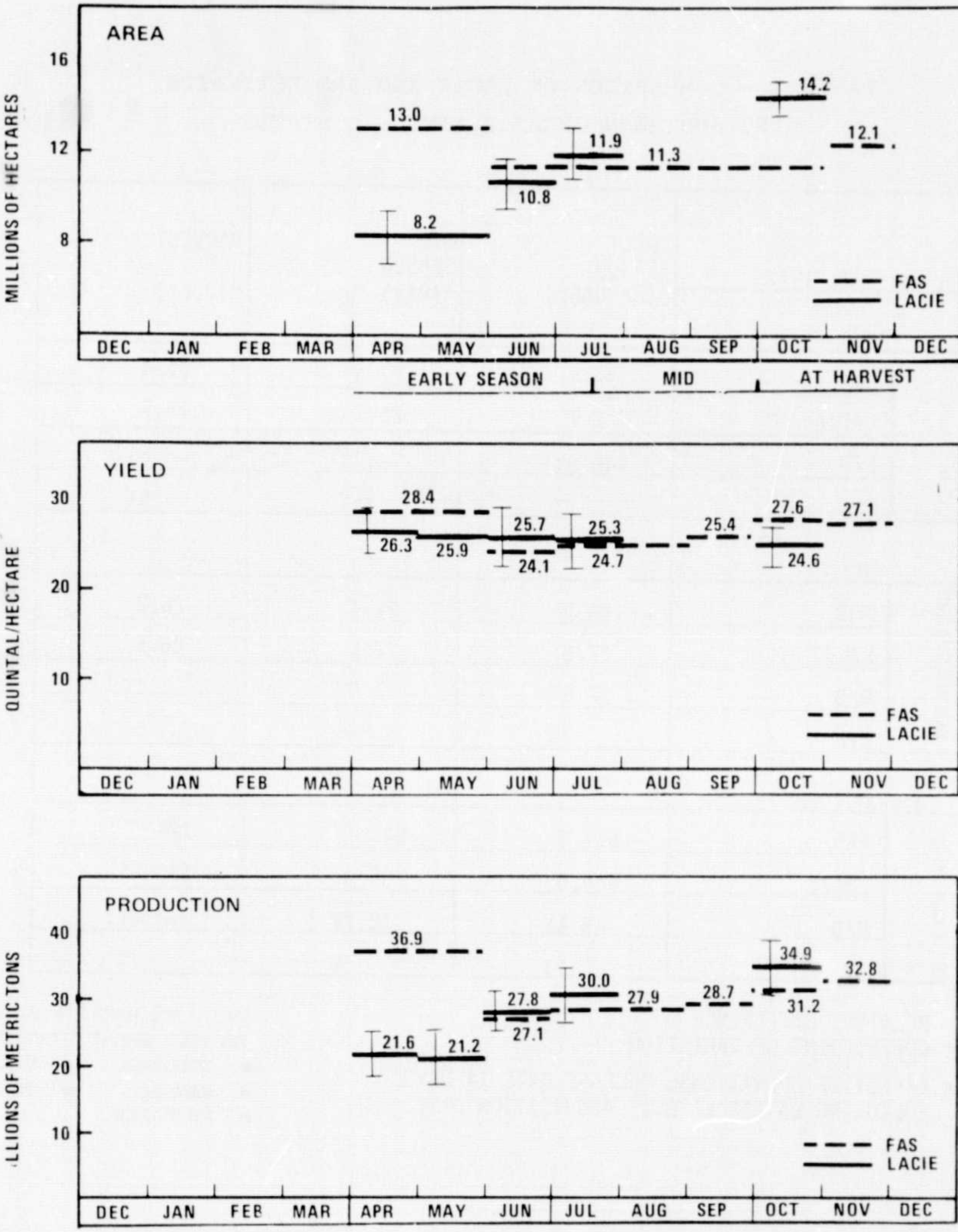
Figure 8. Monthly comparison of LACIE and SRS estimates. (Southern Great Plains)

TABLE I.– COMPARISON OF LACIE AND SRS ESTIMATES
U.S. SOUTHERN GREAT PLAINS – 5 STATES

May 27, 1977		EARLY SEASON *(JANUARY)	MID SEASON *(MAY)	HARVEST *(JULY)
AREA				
ACRES x 10^6	SRS	33.1	27.3	27.4
	LACIE	22.7	26.7	25.7
	R/D	-45.8%	2.2%	-6.6%
	CV	9%	5%	5%
YIELD				
BUSHELS/ACRE	SRS	19.9	24.4	26.2
	LACIE	27.6	26.5	26.5
	R/D	27.9%	7.9%	1.1%
	CV	7%	5%	5%
PRODUCTION				
BUSHELS x 10^6	SRS	659.6	616	726
	LACIE	626.0	706	682
	R/D	-5.4%	12.7%	-6.4%
	CV	11%	7%	7%

R/D = RELATIVE DIFFERENCE
CV = COEFFICIENT OF VARIATION
* = EFFECTIVE OPERATIONAL RELEASE DATE 14 DAYS FOLLOWING LATEST LANDSAT ACQUISITION DATE.

SOUTHERN GREAT PLAINS
(WINTER WHEAT STATES)
- COLORADO
- KANSAS
- NEBRASKA
- OKLAHOMA
- TEXAS



5/7/77
T-297B

Figure 9. Monthly comparisons of LACIE and FAS estimates, U.S.S.R. winter wheat indicator region.

TABLE II. - COMPARISON OF LACIE AND FAS ESTIMATES
U.S.S.R. WINTER WHEAT - INDICATOR REGION

May 27, 1977		EARLY SEASON *(JUNE)	MID SEASON *(JULY)	HARVEST *(OCTOBER)
AREA				
MILLIONS OF HECTARES	FAS	11.3	11.3	11.3
	LACIE	10.8	11.9	14.2
	R/D	-4.1%	5.2%	20.3%
	CV	7%	6%	6%
YIELD				
QUANTAL/ HECTARE	FAS	24.0	24.7	27.6
	LACIE	25.7	25.3	24.6
	R/D	6.6%	2.4%	-12.3%
	CV	4%	6%	5%
PRODUCTION				
MILLIONS OF METRIC TONS	FAS	27.1	27.9	31.2
	LACIE	27.8	30.0	34.9
	R/D	2.5%	7.1%	10.6%
	CV	7%	8%	7%

R/D = RELATIVE DIFFERENCE
CV = COEFFICIENT OF VARIATION

* = EFFECTIVE OPERATIONAL RELEASE DATE 14 DAYS FOLLOWING LATEST LANDSAT ACQUISITION DATE.

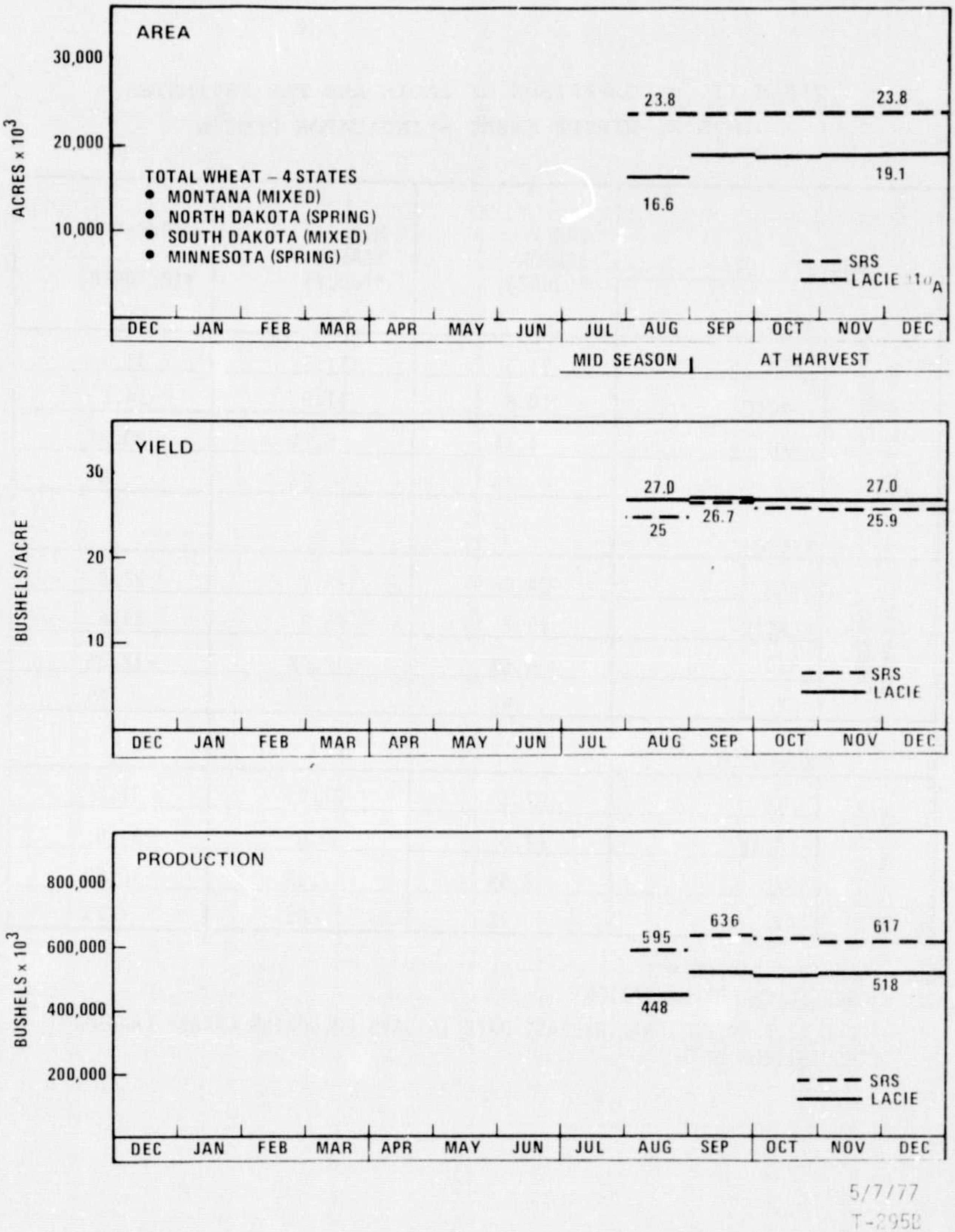


Figure 10. Monthly comparison of LACIE and SRS estimates (total wheat – 4 states)

TABLE III.— COMPARISON OF LACIE AND SRS ESTIMATES

U.S. 4 STATE — TOTAL WHEAT

May 27, 1977		EARLY SEASON *(JULY)	MID SEASON *(AUGUST)	HARVEST *(SEPTEMBER)
AREA				
ACRES x 10^6	SRS	23.8	23.8	23.8
	LACIE	16.6	19.0	19.1
	R/D	-43.3%	-25.26%	-24.6%
	CV	9.4%	6.2%	6.7&
YIELD				
BUSHELS/ACRE	SRS	25	26.7	25.9
	LACIE	27	27.1	27.0
	R/D	7.4%	1.5%	4.0%
	CV	29.6%	27.6%	27.7%
PRODUCTION				
BUSHELS x 10^6	SRS	595	636	617
	LACIE	448	518	518
	R/D	-32.8%	-22.7%	-19.1%
	CV	11.6%	8.9%	8.7%

R/D = RELATIVE DIFFERENCE
CV = COEFFICIENT OF VARIATION

* = EFFECTIVE OPERATIONAL RELEASE DATE 14 DAYS FOLLOWING LATEST LANDSAT ACQUISITION DATE.

TOTAL WHEAT - 4 STATES

- MONTANA (MIXED)
- NORTH DAKOTA (SPRING)
- SOUTH DAKOTA (MIXED)
- MINNESOTA (SPRING)

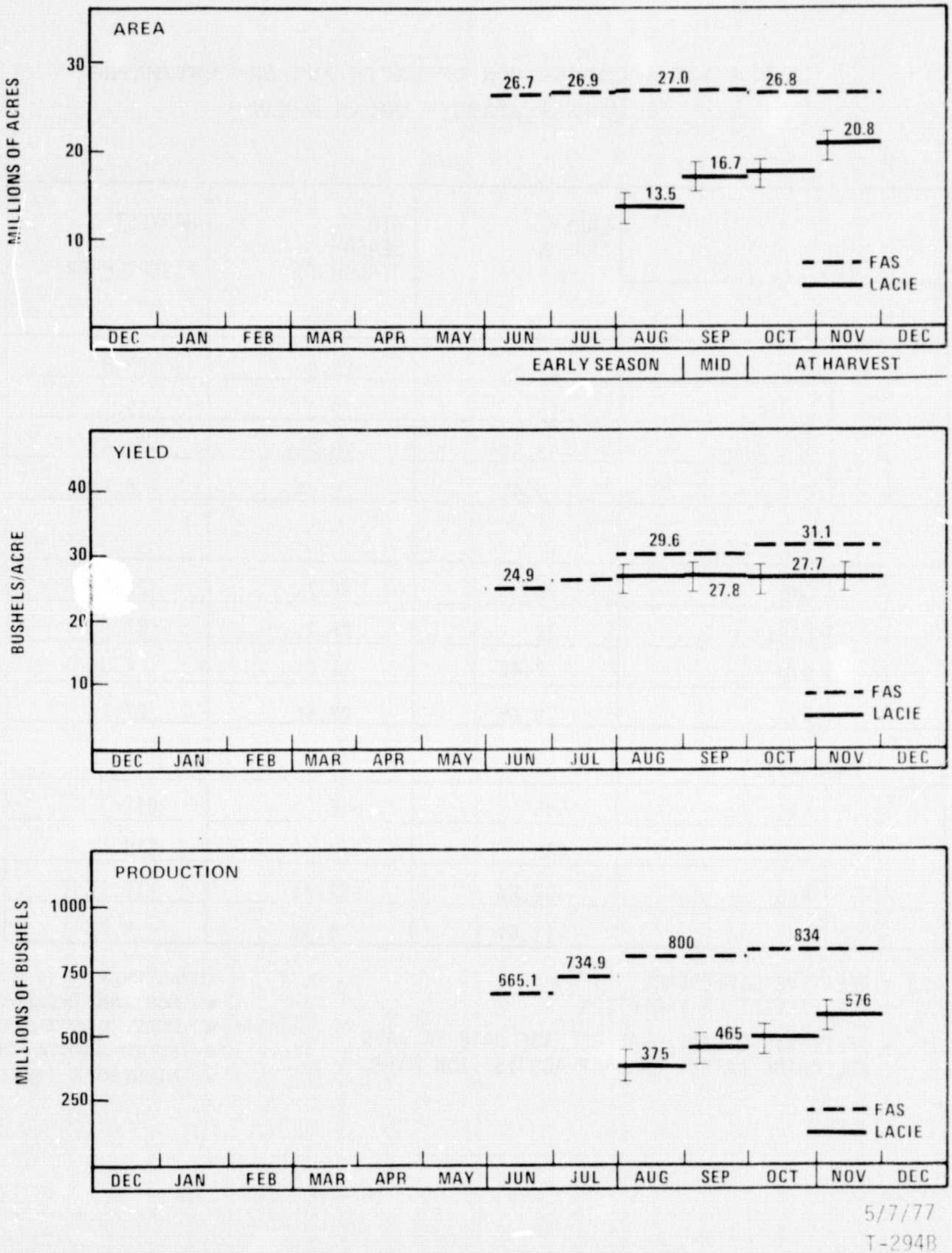


Figure 11. Monthly comparisons of LACIE and FAS estimates, Canada Spring wheat region.

TABLE IV. – COMPARISON OF LACIE AND FAS ESTIMATES
CANADA SPRING WHEAT

May 27, 1977		EARLY SEASON *(JULY)	MID SEASON *(AUGUST)	HARVEST *(SEPTEMBER)
AREA				
ACRES x 10^6	FAS	27	26.8	26.8
	LACIE	13.5	17.3	20.8
	R/D	-100%	-55%	-29%
	CV	4%	3%	3%
YIELD				
BUSHELS/ACRE	FAS	29.6	29.6	31.1
	LACIE	27.7	27.8	27.7
	R/D	-6%	-17.4%	-12%
	CV	4%	4%	3%
PRODUCTION				
BUSHELS x 10^6	FAS	800	800	834
	LACIE	375	481	576
	R/D	-113.3%	-83%	-57%
	CV	5%	5%	5%

R/D = RELATIVE DIFFERENCE
CV = COEFFICIENT OF VARIATION

* = EFFECTIVE OPERATIONAL RELEASE DATE 14 DAYS FOLLOWING LATEST LANDSAT ACQUISITION DATE.

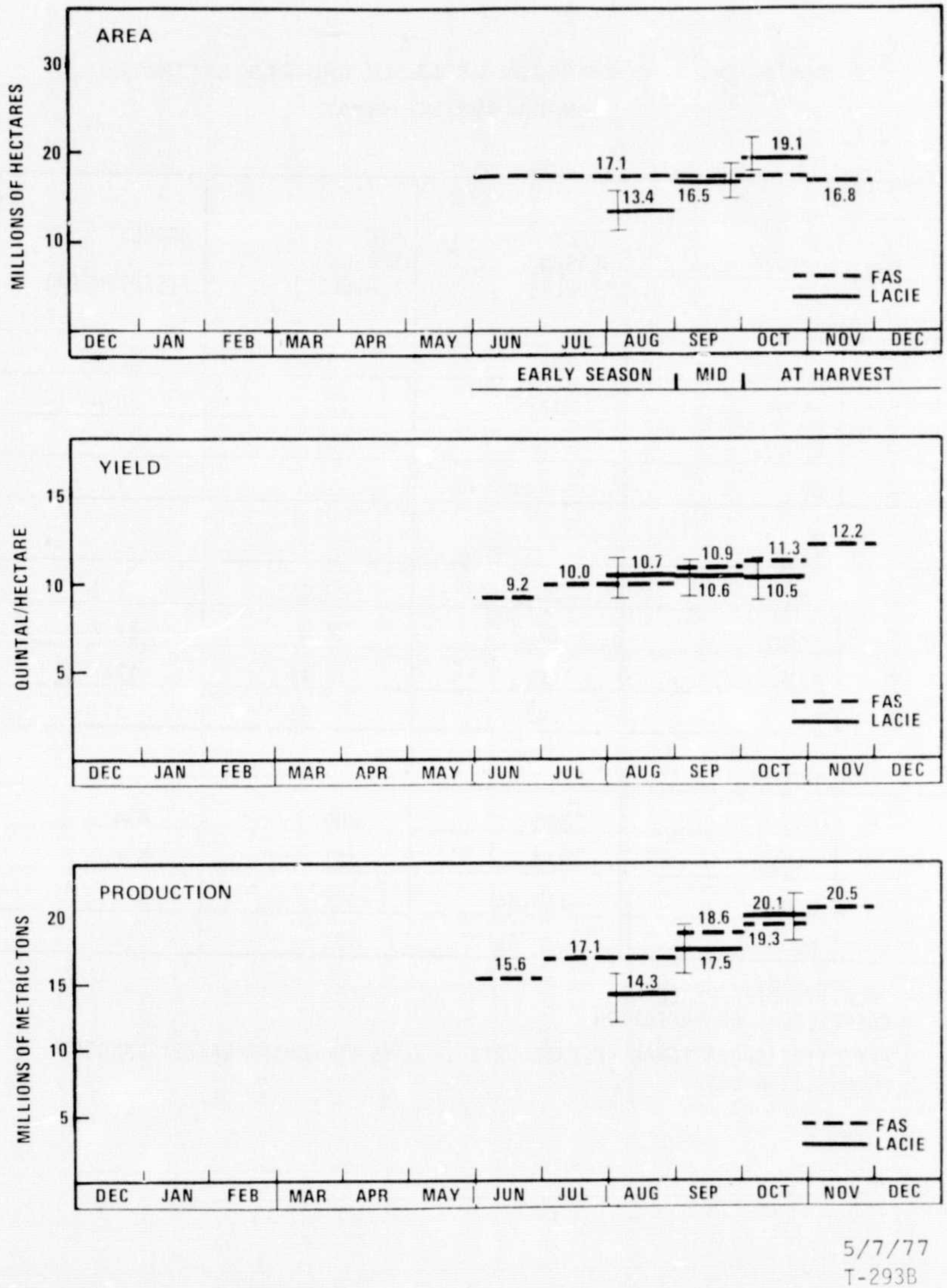


Figure 12. Monthly comparisons of LACIE and FAS estimates, U.S.S.R. Spring wheat indicator region.

TABLE V.— COMPARISON OF LACIE AND FAS ESTIMATES
U.S.S.R. SPRING WHEAT — INDICATOR REGION

May 27, 1977		EARLY SEASON *(AUGUST)	MID SEASON *(SEPTEMBER)	HARVEST *(OCTOBER)
AREA				
MILLIONS OF HECTARES	FAS	17.1	17.1	17.1
	LACIE	13.4	16.5	19.1
	R/D	-27.2%	-3.3%	10.6%
	CV	7%	5%	4%
YIELD				
QUINTAL/ HECTARE	FAS	10	10.9	11.3
	LACIE	10.7	10.6	10.5
	R/D	6.5%	-2.8%	-7.6%
	CV	9%	8%	8%
PRODUCTION				
MILLIONS OF METRIC TONS	FAS	17.1	18.6	19.3
	LACIE	14.3	17.5	20.1
	R/D	-19.6%	-6.3%	4%
	CV	11%	9%	9%

R/D = RELATIVE DIFFERENCE
CV = COEFFICIENT OF VARIATION

* = EFFECTIVE OPERATIONAL RELEASE DATE 14 DAYS FOLLOWING LATEST LANDSAT ACQUISITION DATE.

TABLE VI.— COMPARISON OF LACIE ESTIMATES TO GROUND-OBSERVED PROPORTIONS OVER WINTER WHEAT BLIND SITES IN THE U.S. GREAT PLAINS

Month	No. of segments	RMD, % (c)	Percent underestimated (d)
February	71	-30.6	83
March	95	-26.2	79
April	95	-26.2	79
May	95	-21.4	75
June	95	-15.7	72
July	95	-16.2	70
August	95	-15.2	71
September	95	-13.3	68
October	95	-13.7	68
Final	95	-13.2	68

TABLE VII.— COMPARISONS OF LACIE ESTIMATES TO TO GROUND-OBSERVED PROPORTIONS OVER ALL AVAILABLE SPRING WHEAT BLIND SITES IN THE U.S. GREAT PLAINS

Month	No. of segments	RMD, %	Percent estimated
August	33	-41.6	88
September	33	-25.6	82
October	33	-24.1	79
Final	33	-22.6	79

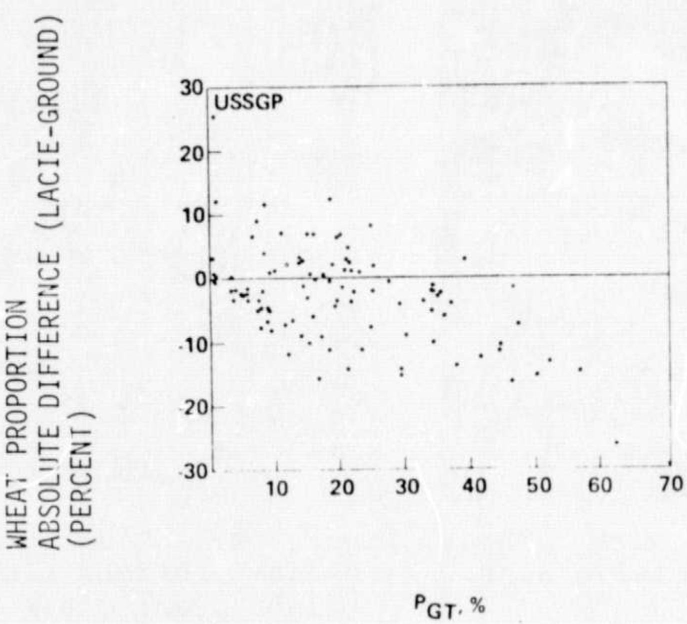


Figure 13a. Plot of winter wheat proportion estimation errors versus ground truth winter wheat proportions for blind sites in the U.S. Great Plains.

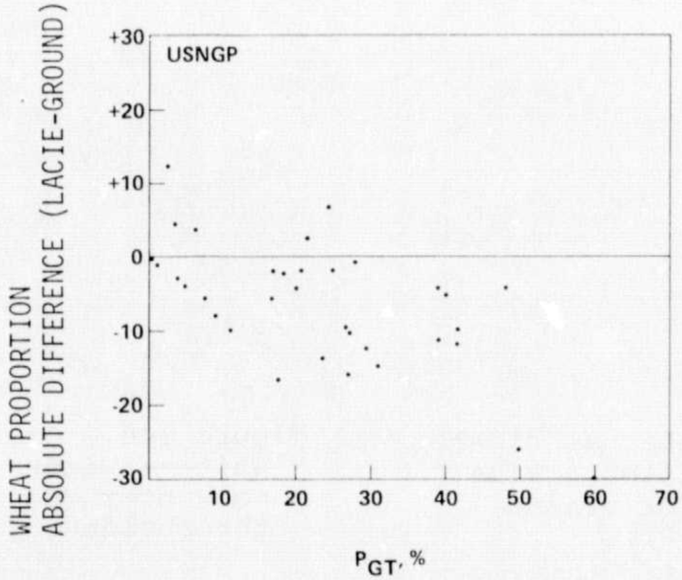


Figure 13b. Plots of spring wheat proportions estimation errors versus ground truth values for blind sites in the U.S. Great Plains.

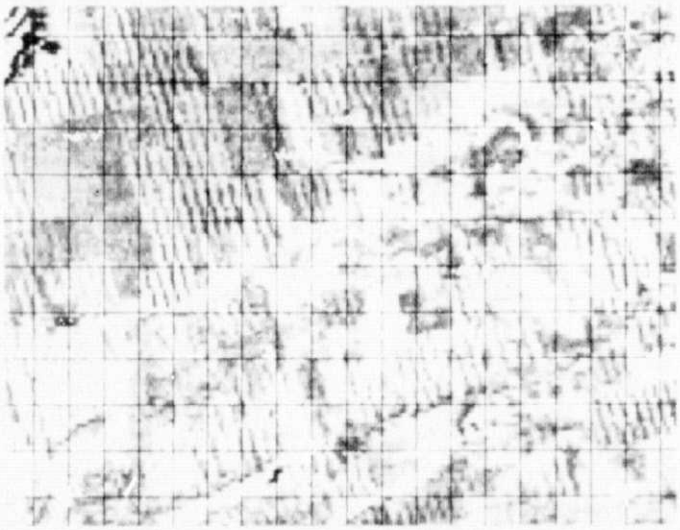


Figure 14a. Color IR image. Wheat emergent stage; W-Winter grains, N-non-winter grains.

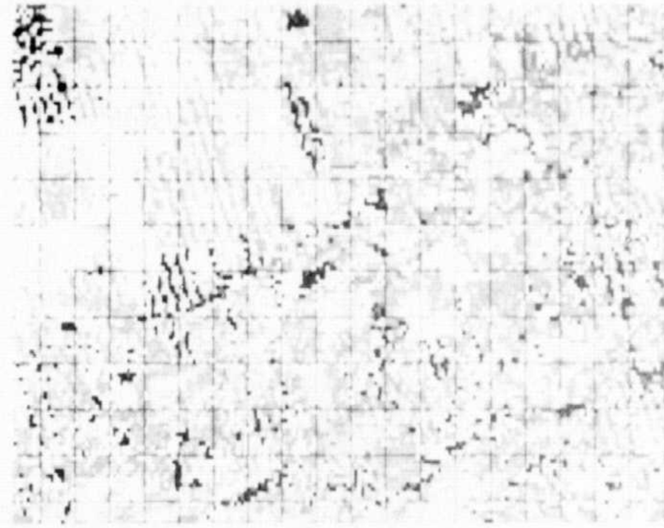


Figure 14b. Cluster map. Winter grains – bright blue, bright green, bright cyan; other colors - non-winter grains.

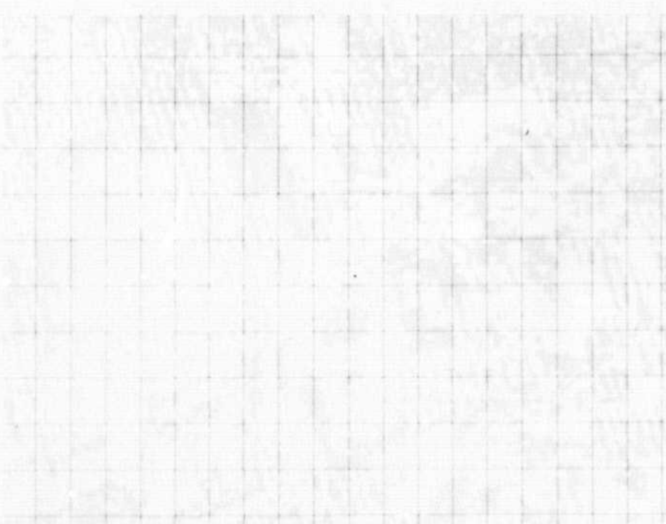


Figure 14c. Conditional cluster map. Green - Winter grains; Yellow - nonwinter grains; Blue - conditional.

Figure 14d. Classification map. White - Winter grains; Gray - nonwinter grains; Black - thresholded.

Figure 14. SMALL FIELDS COLOR INFRARED IMAGE. Cluster map, conditional cluster map, and classification map for Fergus County, Montana Segment, 11 Nov. 1976.