Reprinted from

# Eighth International Symposium

# Machine Processing of

# Remotely Sensed Data

with special emphasis on

# Crop Inventory and Monitoring

**July 7-9, 1982**

# Proceedings

# RESEARCH IN SATELLITE-AIDED CROP INVENTORY AND MONITORING

J.D. ERICKSON, J.L. DRAGG, R.M. BIZZELL, M.C. TRICHEL

National Aeronautics and Space
Administration/Johnson Space Center
Houston, Texas

## ABSTRACT

Developments of efficient and accurate automated information extraction procedures for analysis of multitemporal Landsat data in non-U.S. crop inventory and monitoring can provide a greatly improved capability for practical and affordable use on a global basis without requiring ground observations.

## I. INTRODUCTION

At the launch of Thematic Mapper, research in satellite-aided crop inventory and monitoring of global production continues to make advances toward practical, viable capabilities and improvement in current procedures. These capabilities, when added to current agricultural information systems, are expected to provide more timely and more accurate non-U.S. crop information than is now available.

While the overall objective of this research[1] is to develop technology for extracting agricultural information of various kinds, the focus is on improved non-U.S. production forecast technology which will be evaluated by the U.S. Department of Agriculture (USDA) for possible integration into its information systems. More specifically, the objective is to develop procedures for using aerospace remote sensing and related technology to provide more objective and reliable crop area forecasts at several times during the growing season and with improved preharvest estimates for selected countries and crops.

We have adopted four categories of factors that reflect key characteristics of technology and which guide us in the research. These are timeliness, affordability, general applicability, and accuracy. Timeliness emphasizes the quick extraction of information. Timeliness is associated with early-season estimates as well as estimates made thoroughout the season. Affordability reflects efficiency and inexpensiveness. General applicability refers to techniques that are applicable for foreign crop regions, while remaining objective and improvable procedures. Accuracy reflects the degree of bias and variance over time and the responsiveness to factors affecting departures from average.

## II. TECHNICAL APPROACH

The technical approach is to provide satellite-based, objective estimates of area, yield, and production as one input into a comprehensive, multidata-source information system. The conceptual framework involves estimating crop area and yield for specified regions and multiplying them together to obtain production at the regional level. Automatic data processing approaches are considered necessary to provide objective, timely, and reliable estimates. Our research has been focused on the area estimation component; hence, further discussion will be primarily oriented to that component. Area estimates for regions are derived by processing statistical samples (called segments) of satellite digitial image data (Landsat MSS). The desire is to estimate crop area periodically throughout the season, from the time of planting through harvest.

Current approaches in crop area estimation extensively utilize statistical sample survey methodology. Sampling methodology allows reliable estimates to be made by processing only a very small portion of the data. Expansion of the sample estimates to a regional estimate is referred to as aggregation and is a statistical process that is made more complex by non-response (generally because of intervening cloud cover) and partial response.

Since non-U.S. ground data may be available in limited quantities for research studies but would not be available to support an operational system, and because of efficiency and timeliness considerations, our approach to non-U.S. crop forecasting does not require ground observations.

The variability in the crop scene environment among crop regions and countries is quite large. It is also strongly desirable to minimize human interaction. Hence, recent approaches have concentrated on the development of robust procedures that are largely self-adaptive, in terms of recognizing crops based on their spectral signatures over time and distance. Some adaptation to specific crop regions is still necessary. The foremost challenge in satellite-oriented global crop forecasting is crop estimation without the ground observations that serve to train (adapt) remote sensing procedures which use them.

Technology development has been based on use of extensive ground observations obtained in the United States where the reliability of the observations is understood. Regions in the United States that are similar to foreign crop regions of interest have been selected as study areas. Since true analogue regions do not exist, these foreign similarity regions in the United States are not completely suitable. Incremental testing over domains of greater variability of independent data sets is required. Such testing serves to establish levels of performance, isolate the technology deficiencies, and suggest required avenues of resolution. Even so, testing on such data sets is limited to the variability of available characteristics. However, by using the available data to establish the sensitivity of the methodology to these parameters, simulation techniques can be employed to establish the performance over a wider variety of conditions.

## III. DESCRIPTION OF EXPERIMENTS

The objective in our research experiments has been to develop and evaluate state-of-the-art technologies for estimating the crop area of spring small grains, summer crops, corn, and soybeans, in order to obtain an understanding of where limitations in performance exist and so that R&D efforts can be focused there. These experiments represent the first "independent" sub-system level sensitivity tests against the adopted performance criteria of timeliness, affordability, applicability, and accuracy.

The technology developed for crop are estimation of spring small grains and that for summer crops, corn, and soybeans reflected two different stages of technical maturity. For spring small grains, a technical breakthrough has been achieved (figure 1). A highly efficient state, from a man/machine systems point of view, has been attained by modeling the interpretive intelligence of skilled and expert image analysts in an automated and objective form - simple, artifical intelligence. Area estimation technology for the summer crops, corn, and soybeans, being 4 years later in initiation, was in a less mature stage of development. Thus, two separate experiments (one for spring small grains and one for summer crops, corn, and soybeans) were designed to optimize the evaluations of these technologies.

## A. SPRING SMALL GRAINS

Area estimation technology available for evaluation produced regional aggregations as well as component crop area proportion estimates within sample segments.

The segment proportion estimation component was represented by two different technical approaches to the same basic identification scheme of determining different spectral crop appearance development over time. Both approaches exhibit an efficiency improvement through automation of an order of magnitude, when compared to previous analyst-intensive procedures (figure 2). The first procedure (SSG3C) is a completely automated modeling[2] of the previous manually intensive, skilled and expert analyst interpretative procedure. The basic functions consist of: acquisition selection based on meteorological variables and anticipated Landsat spectral responses for spring small grains; a transform of Landsat data to a level of greenness observed; an automatic, multitemporal, pixel-labeling logic based on a hierarchical process; and a proportion estimator based on a systematic sample of the labeled pixels. SSG3B (same technical approach as SSG3C) which allows an analyst override of the automatic, acquisition-selection module was also developed[2] and evaluated in this experiment.

The second proportion estimation approach (SSG4) was also completely automatic[2]. The basic functional differences between this approach and that of the SSG3 were: the Landsat data are transformed by a model which "normalizes" color; a "field" finding algorithm groups the individual pixels into quasi-fields; the quasi-fields are labeled by a multitemporal logic of "green/not-green" sequence; and the proportion estimate is derived from an enumeration of the labeled fields plus an adjustment for estimated omission rates.

A major objective of the experiment was to ascertain if the cost benefits of procedure automation could be achieved with comparable accuracy to the analyst-intensive procedures.

The Landsat data used for the small grains experiment consisted of sample segments collected over a 4-year period (1976-1979) covering the U.S. Northern Great Plains (North Dakota, South Dakota, Montana, and Minnesota) and Saskatchewan, Canada (figure 3). Only segments that had ground observations for evaluation were analyzed, with the exception of 1978 samples of North Dakota in which all allocated Landsat segments were included to support aggregation studies.

The 1976 crop year in this region was warmer than average. The crop season was earlier than average, as was the case in 1977, even though the temperature was near average. The 1978 crop year was cooler than average, which gave normal to late planting. Late planting was experienced in 1979 as well.

# TECHNICAL BREAKTHROUGH
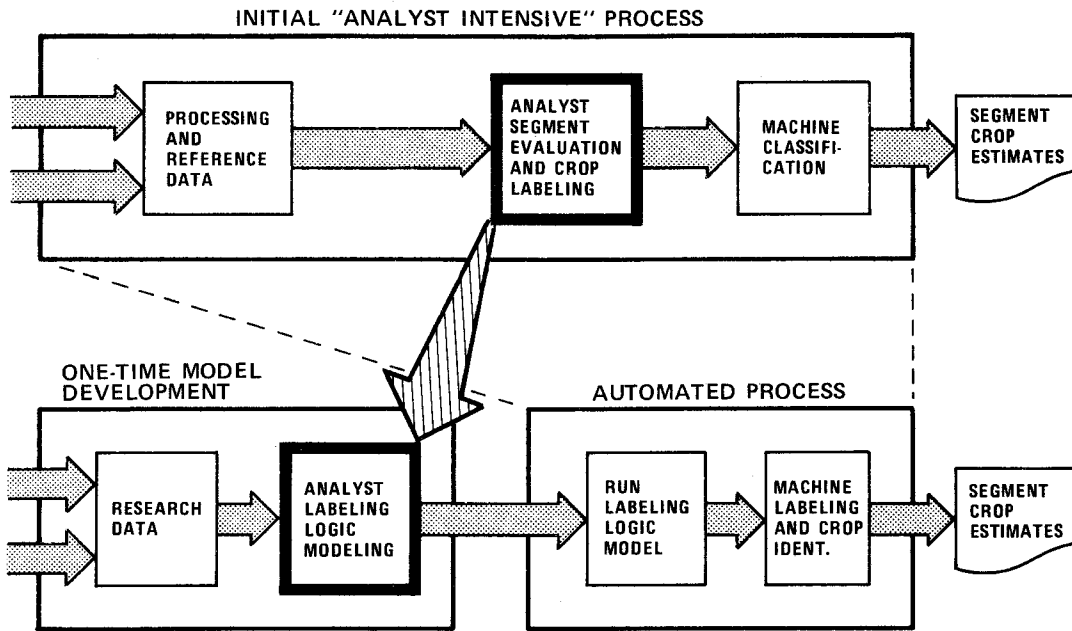
INITIAL "ANALYST INTENSIVE" PROCESS



Figure 1.- Technical breakthrough (modeling of analyst function enabled automatic processing).
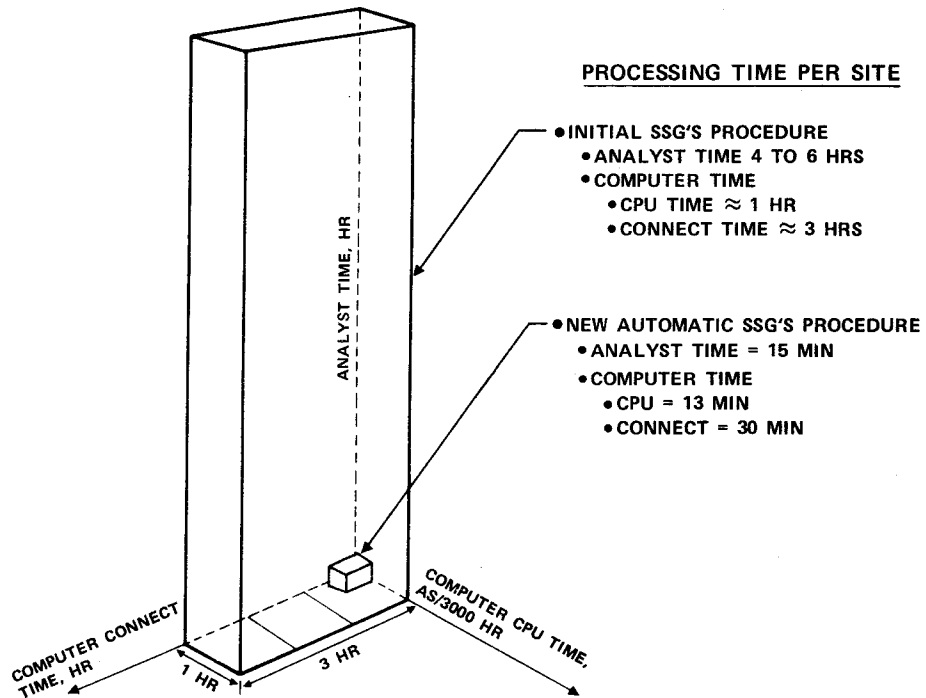


Figure 2.- Improved efficiencies of automated information extraction process.

1982 *Machine* Processing of Remotely Sensed Data Symposium

**(A) 1976 SITES (WEATHER: WARM; SEASON: EARLY)**

**(B) 1978 SITES (SEASON: EARLY)**

**(C) 1977 SITES (WEATHER: COOL; SEASON: NORMAL TO LATE)**
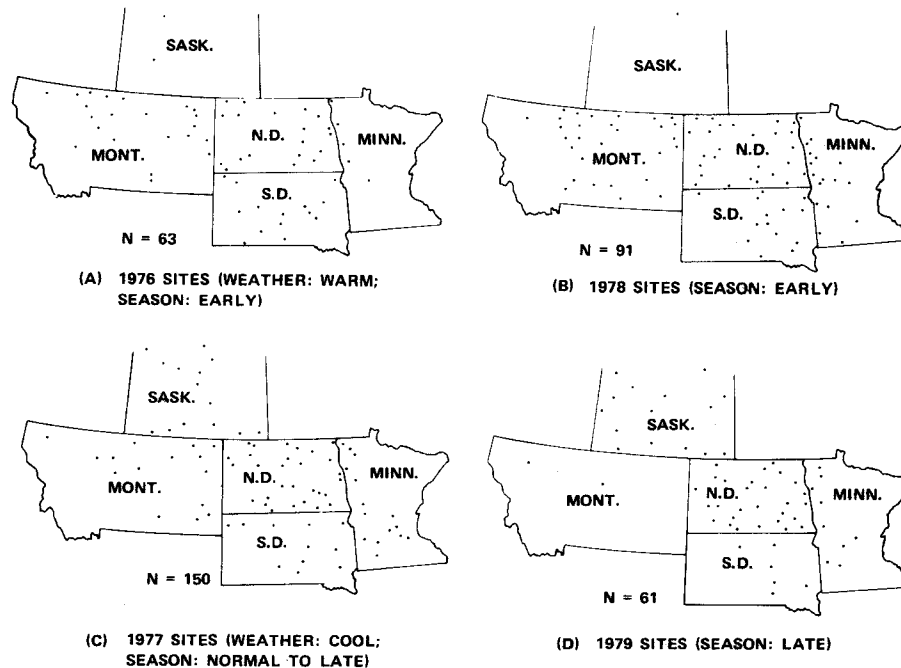
**(D) 1979 SITES (SEASON: LATE)**

Figure 3.- Site locations and meteorological summary for U.S. and Canada spring small grains.

Ideally, the data set for this experiment would be completely representative of non-U.S. crop regions. While it is not, this is the most extensive data set ever used in testing, including about 9,000 square miles of ground observations spread through four states and one province over 4 crop-years.

B. SUMMER CROPS

The level of evaluation of summer crops, corn, and soybeans technology was more limited in both scope and purpose, due to its less mature state. The results of this experiment would provide identification and quantification of the major subcomponent contributors to the proportion estimation error for future developmental modification prior to the development of an automatic processing approach.

The summer crops, corn, and soybeans proportion estimation technology (called CS-1) performed functions similar to the small grains procedures, such as: Landsat data transformation and feature extraction; target definition and stratification; labeling; sampling; and proportion estimation. However, the design of the experimental procedure was structured and analyst-intensive, to allow the tabulation of results at each detailed procedural step for subsequent evaluation and performance analyses.

The Landsat data for the summer crops, corn, and soybeans experiment consisted of 18 sample segments from the 1978 crop year and 10 sample segments from the 1979 crop year. A large data base for evaluation, such as was available for small grains, does not yet exist for corn and soybeans. These segments were drawn from the U.S. Central Corn Belt states of Iowa, Illinois, Indiana, and Missouri.

IV. RESULTS

A. SPRING SMALL GRAINS EXPERIMENT[3]

North Dakota Aggregations. Evaluation of these procedures, developed specifically to estimate non-U.S. spring small grains area, were carried out over areas in the United States (and Canada) where reliable reference data were available before attempting adaptation. The results of North Dakota aggregations for 1978 data are (percent relative error and coefficient of variation): +3.5 (C.V. = 4.4), +6.9 (C.V. = 4.3), and -9.2 (C.V. = 4.6) for SSG3B, SSG3C and SSG4 respectively.

The performance is expressed relative to the published USDA estimate of the spring grains acres harvested in North Dakota during 1978 (13.12 million acres).

Subsystem Level Results. The sample segment proportion estimation accuracy performances are shown in figure 4. The results shown include all segment analyses including machine errors, clerical errors, and otherwise outliers. (These unedited results are also included in the aggregations.) The intent was to identify and quantify all potential contributions to error, without regard to producing the best estimate. By not thresholding seemingly obvious outliers, we optimize the probability of isolating the major subcomponent error contributors. Even with this, it was somewhat encouraging that the mean absolute error in most cases was less than 10 percent. Also, the results when compared to the more labor intensive historical procedures verify what was expected, i.e., generally comparable with lower bias and somewhat larger variance.

It is of interest to get an indication of the possible performance of the technology for a key foreign region. Similar performance was obtained for segments within a previously determined USSR Foreign Similarity Region (FSR).

In general, the automatic procedures currently provide estimates for fewer segments (50 percent to 65 percent of total allocated)

compared to the historical manual procedure (75 percent).

In terms of the other performance criteria, figure 2 shows a summary of key efficiency parameters related to affordability. Although neither the historical nor the current automated technologies were engineered nor implemented for an operational environment (substantial overhead is necessary in an experimental mode for recording of intermediate output), the relative improvements exhibited by the automated technologies is obvious. The timeliness in the growing season is some 30 days prior to harvest.

In summary, this first-time evaluation of highly automated spring small grains area estimation technologies was very encouraging. There was definite improvement in near-harvest estimation efficiency with modest losses in accuracy as compared to the best previous analyst-intensive approaches. For the 1-year, one-state (1978 N.D.) aggregation, no glaring deficiencies were noted. The attractive characteristics of low processing cost, objectivity, repeatability, modularity, and adaptability packaged within an automated framework make the outlook for meaningful advancements very optimistic.
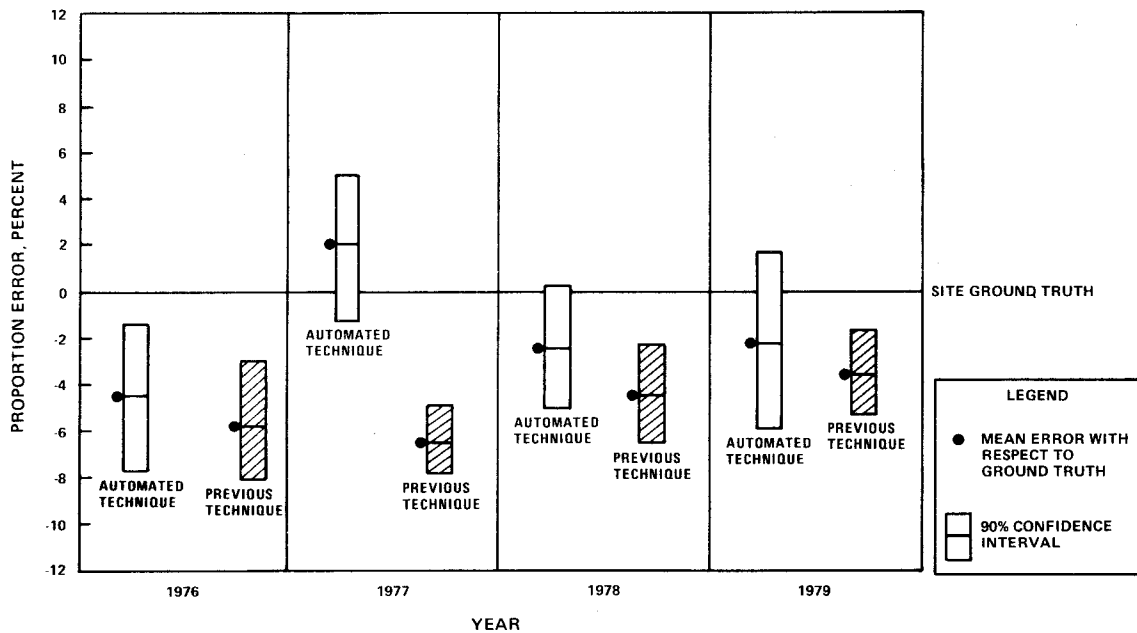


Figure 4.- Accuracy comparisons of spring small grains techniques.

## B. SUMMER CROPS, CORN, AND SOYBEANS EXPERIMENT[4]

This experiment in summer crops, corn and soybeans proportion estimation technology represents our first attempt to take advantage of an improved systems approach to research and allowed the effective utilization of this environment for more iterations of development, testing, and feedback to research than heretofore achievable in a given time (such as 6 months).

Figure 5 shows accuracy results in terms of mean error and 90 percent confidence limits for the previously described crop year (1978-79) data.
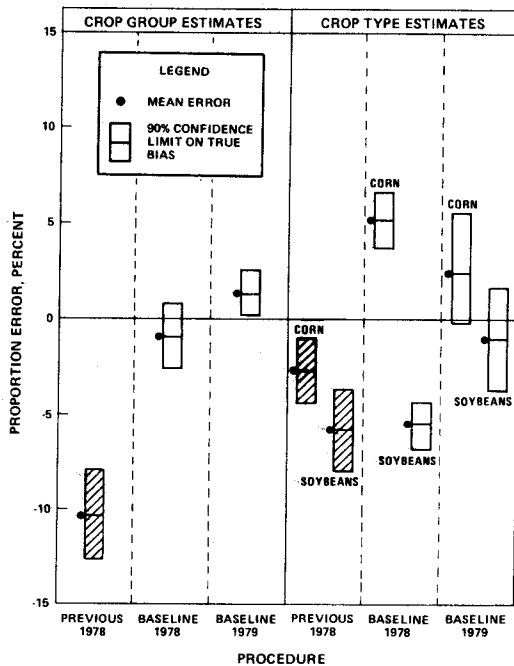


Figure 5.- Comparison of corn and soybean baseline subsystem with previous results.

The following are some additonal observations:

1. There is a very accurate estimate of crop group (summer crops) with RME less than 2 percent in both years.

2. For crop type (corn and soybeans) there was a significant bias (corn RME=15 percent, soybeans=19 percent) in 1 year (1978).

3. The crop type bias for the other year (1979) was not significant (corn RME=6 percent, soybeans=3 percent).

4. The direction of the bias for crop type appears to be consistent between years (corn tends to be overestimated, soybeans tends to be underestimated).

5. The standard deviations for all estimates are relatively consistent and comparable to those achieved in previous "best" spring small grains technologies.

6. Compared to a previous procedure for the 1978 crop years, CS-1 exhibited a significantly lower bias in estimating crop group (summer crop RME -1.5 vs. -16.4 percent) with lower standard deviations for both crop groups and crop types.

In terms of the other performance criteria the following was observed:

1. Timeliness - processing to crop type is achievable after corn tasseling which is 30-45 days prior to harvest. This is about the middle of August in the Corn Belt.

2. The processability rates with the CS-1 technology were quite high, typically 50 percent to 75 percent.

Quick identification of subcomponent contributors was achieved. These specific results have led to the earlier-than-planned development of a more-automated summer crop, corn, and soybeans proportion estimation procedure. The process of making the developmental modifications for a new version (CS-1A), conducting a verification test, then designing and implementing a semiautomatic version (CS-1B) has already been accomplished. Results of early verification testing over a sample (10 segments) of the 1978-79 data set show excellent summer crop accuracy[5] and corn and soybeans accuracy of about 10 percent relative mean error and standard deviations of 4 to 6 percent.

The results of this experiment are very encouraging because the develomental time frame including procedure development, testing, identification of limitations, and procedure improvement was accomplished in about one-fifth the time of previous research in this area. Additional improvements are expected in this technology.

## V. SIGNIFICANCE OF RESULTS

The significance of the present results is threefold:

1. We were able to adequately model the subjective human analyst with an objective process and achieve reasonable accuracies.

2. We were thereby able to develop an information extraction technology which was not prohibitively costly either in terms of manual effort or computational resources. It is affordable within a reasonable standard.

3. If the results reported here could be achieved for foreign regions, substantial improvements in global crop information would result.

This should not suggest that the non-U.S. inventory and monitoring problem is solved as several key research problems remain. Briefly, the most critical of these are:

1. Improving the automated selection of acquisitions.

2. Finding methods to estimate crop areas much earlier in the season.

3. Finding information extraction methodologies which reduce the quantity and quality of the data required, thus reducing data costs.

4. Finding methods for adapting these objective analysis methods to other crops and regions without requiring information which is not available in non-U.S. situations.

5. Early quantification of the benefits of improved performance of this technology due to the features offered by the Thematic Mapper.

6. Doing adequate testing and evaluation to understand the technology performance.

We feel that the present results represent a real breakthrough both in approach and practicality.

## ACKNOWLEDGEMENTS

We wish to thank the many professionals of the joint agency and contractor team which have contributed to the research reported here.

## REFERENCES

1. AgRISTARS Annual Report for FY81, AP-J2-04225, NASA Johnson Space Center, Houston, January 1982.

2. J.T. Waggoner and D.E. Phinney, "Project Procedures Designation and Description Document," Vol. I., FC-L1-00715/JSC-17154, NASA Johnson Space Center, Houston, June 1981.

3. FCPF Preliminary Technical Results Review /Spring Small Grains, Vol. I., FC-J1-04175/JSC-17433, NASA Johnson Space Center, Houston, September 28, 1981.

4. FCPF Preliminary Technical Results Review /Corn and Soybeans, Vol. II., FC-J1-04175 /JSC-17433, NASA Johnson Space Center, Houston, September 29, 1981.

5. ITD Semi-Annual Review to Level I, IT-J2-04267/JSC-17830, NASA Johnson Space Center, Houston, April 15, 1982.

## BIOGRAPHICAL DATA

Jon D. Erickson. Now managing the Inventory Technology Development (ITD) project of AgRISTARS, he has extensive experience in the administration of advanced multispectral analysis research and development programs. He received a B.S.E. in Science Engineering in 1959, an M.S.E. (1962) and a PhD: (1966) in Nuclear Engineering, all from the University of Michigan.

James L. Dragg. Chief, Agriculture Applications Branch, he also managed the LACIE Application Evaluation System, the LACIE Transition Project, Apollo photogrammetric analysis. He completed a B.A. in Geography and Mathematics at the University of Oklahoma in 1958, an M.S. in Geodetic Science at Ohio State University in 1964, and a year of graduate-level mathematics at the University of Oklahoma in 1965.

Robert M. Bizzell. Now managing AgRISTARS ITD Experiments, he also managed the LACIE CAMS operations and has extensive experience in managing projects for development of Landsat data analysis techniques. He completed a B.S. in Mathematics and Physics at S.W. Oklahoma State University in 1964, an M.S. in Geodetic Science at Purdue University in 1966, and post graduate course work at St. Louis University during 1967-69.

Milton C. Trichel. Now managing AgRISTARS ITD Technology Integration operations, his experience includes Agristars Supporting Research, LACIE and LACIE Tansition Research, and management of the LACIE Classification and Mensuration Subsystem. He received a B.A. in Engineering in 1962, a B.S. in Electrical Engineering in 1963, and an M.S. in Space Physics in 1967, all from Rice University.