UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

SPACE SCIENCES LABORATORY BERKELEY, CALIFORNIA 94720

TWX: UC SPACE BERK
(910) 366-7945
Remote Sensing Research Program
260 Space Sciences Laboratory
University of California
Berkeley, California 94720

September 17, 1975

Dr. A. E. Potter, Technical Monitor
NASA-Johnson Spacecraft Center
Mail Code TF 3
Houston, Texas 77058

SUBJECT: Quarterly Progress Report, Contract #NAS 9-14565
Contract for: Study-Development of Improved Photointerpretative techniques to Wheat Identification.

Dear Dr. Potter:

The objectives, approaches, supporting subtasks, and milestone schedules for the U.C. Berkeley LACIE tasks were given in the January 1975 proposal to JSC and refined at the beginning of the contract period (in conference with J. Deitrich and J. Garcia) as documented by the copy of the enclosed transmission (enclosure #1) to J. Garcia dated June 18, 1975.

The major activities engaged in during the first 90-day performance period (May 15, 1975 to August 15, 1975) for each of the three tasks within the contract were as follows:

TASK I (4.1.1.1a) Evaluation and Revision of Techniques for Manual Interpretation

A prime objective of this task is to quantitatively evaluate various LANDSAT image sets in order to determine the optimum sequential data set(s) needed for accurate identification of the wheat crop. In addition, this task will determine expected accuracies for interpreters using optimum and selected less than optimum (missing data effect or desire for early season estimation) sequential LANDSAT data sets. The activities of this first reporting period centered on data acquisition and preparation, and area familiarization. The specifics of these activities are described below.

Data files for LANDSAT imagery, ground data, and ancillary data were set up for the three agreed upon study regions - Kansas, North Dakota, and Montana, as well as for other possible regions of interest.

All data received from the J.S.C. technical monitor as well as that collected through our own efforts has been catalogued and filed for easy assess and to insure against loss.

A library search for literature pertaining to the geography, agricultural environment and agricultural management practices of the study regions was undertaken to provide ancillary data for study site familiarization of persons who would be involved within the overall project and specifically to provide adequate background material for training of interpreters involved in the testing phase of this task. The bulk of this activity has been completed for Kansas and will proceed for the other two study regions (North Dakota and Montana) in a phased manner relevant to the milestone plan for Task I.

Additional ancillary data in the form of weekly, monthly and annual reports from the respective states' Departments of Agriculture in cooperation with the U.S.D.A. Statistical Reporting Service have been subscribed to and are currently being received for all study regions.

Based on initial work on the five intensive test sites in Kansas, the data found to be most useful as background material and as area familiarization material have been:

1. The sequential full frame LANDSAT I color infrared color composites of the test site regions;

2. The U.S.D.A., Soil Conservation Service Soil Surveys for the counties within the two crop reporting districts considered (the southwest and the central crop reporting districts) -- particularly the sections of the reports which discuss the general agriculture of the county and the use and management of various land use classes such as dryland areas, irrigated areas and rangeland areas.

3. The weekly weather and crop reports for Kansas, the monthly crop production reports for Kansas, the annual crop acreage summaries for Kansas 1971-1974, and special reports on grain quality and grain varieties for Kansas published by the Department of Agriculture and the U.S.D.A. Statistical Reporting Service.

The five Kansas Intensive Test Site areas have been enlarged and copied onto color negative film for all dates of LANDSAT I color infrared composites received to date. For eight dates of LANDSAT I coverage of the Morton County, Kansas ITS for which we have not received color composites, but for which we have received the separate black and white positive transparencies, color infrared enhancements were made and enlarged to complete the sequential color infrared coverage for this site. The enlargements of the five Intensive Test Sites will be presented to the interpreters as part of the testing phase of Task I.

Analysis of the ground data, ancilliary data, and sequential imagery for the southwest and central crop reporting districts of Kansas is proceeding so that some preliminary screening can be done before the final design of the interpretation tests, and so that an efficient training package can be presented to the interpreters.

TASK II (4.1.1.2f(1) and 4.1.1.2f(3) Development of a Spectral Signature Extension Methodology

The primary objective of this task is to develop wheat analogue area stratification strategies and spectral signature prediction models to enable accurate, cost-effective wheat acreage determination in LACIE sample segments lacking ground data.

As specified in the contract, there are five major supporting subtasks to achieve this objective. Progress on each task for this progress report period is cited below. The overall U.C. Berkeley approach to the signature extension problem is summaried in enclosure no. 2. An updated supporting data request for this task, dated July 8, 1975 and sent to J. Garcia, is included here as enclosure no. 3.

SUBTASK I: Spectral Signature Extension Methodology Design

The first component of this subtask is the specification of a stratification strategy to delineate wheat analogue areas. A wide review of the literature (including the following important references: Pascale and Damario 1962, Peterson 1965, Klages 1942, and Nuttonson 1956) combined with U.C. Berkeley's previous work indicates that the following variables should be used for stratification:

- General soil type
- Land use/cropping practice
- Average growing season day-degree sums (April-June)
- Avertage growing season precipitation (April-June)
- Average last date of spring frost
- Average temperature and/or average minimum temperature for the coldest month of the year

These stratification variables are then used for delineation of analogue areas in one of four "first-case" ways that have been developed during the reporting period. Basically, these stratification techniques involve the following procedure. First, a single time of year is identified when soils/land use/cropping practice patterns of interest are most separable on LANDSAT false color 9 x 9 inch format transparencies. For the crop reporting district or state of interest, these LANDSAT transparencies from the specified period from a given year are then enlarged (approximately 1.5 times) to print form and mosaiced together.

Clear acetate is placed over the image mosaic. County and other crop reporting unit boundaries are then drawn on the acetate relative to the image according to available planimetric map references. With the aid of general soil association information (e.g. regional soil maps or county SCS soil reports) general land use/cropping practice descriptions and/or two dates of supplemental LANDSAT transparencies that bring out different aspects of soil/land use characteristics, a photo-interpreter delineates on the acetate overlay the areal extent of the soil/land use/cropping practice patterns as seen on the LANDSAT imagery.

Onto the soil/land use/cropping practice mosaic is overlain broad threshold lines of the temperature and precipitation stratification variables cited earlier. Generally, the isolines of the temperature variables are very positively correlated with one another, so that only one or two need be considered for stratification.

The resulting soil/land use/temperature/precipitation stratification defines unique wheat analogue area environments. Each of these environments should give rise to similar sequences of wheat spectral signatures during the life cycle of that crop, no matter where that analogue area type occurs in the world. Hence, variance among wheat signature and confusion crop/wheat signature mixes should be minimized, thus tending to maximize the success of signature extension models employed within each analogue area type.

The general signature extension (prediction) model selected for experimentation is a regression-like equation known as an analysis of covariance model employing both qualitative (e.g. crop type) and quantitative (e.g. a continuous variable such as temperature) variables. A regression equation will be employed as the "first case" expression of the more general model.

The regression equation will have as its objective the prediction of the mean spectral response for a given LANDSAT band for a given crop residing in a recognition segment. The "prediction" or "independent" variables in the regression equation used to predict the dependent variable, (mean spectral response) have been selected to be:

- Crop maturity codes from A.I.s
- Soil color codes from A.I.s
- Growing season day degree sums
- Planting season day degree sums
- Growing season precipitation
- Planting season precipitation
- Sun angle
- Scan angle
- In-segment LANDSAT spectral data such as the BAND 7/BAND 5 ratio (crop calendar correlated)
- Atmospheric transmittance (generated from ground horizontal visibility data obtained at airports)

The regression coefficients are estimated for these predictor variables from data in training segments for wheat and major confusion crops for each phenological phase.

Signature predictions can then be made for LANDSAT bands 4, 5, 6, 7 for a given phenological phase on a field basis or for clusters identified within the recognition segment through non-supervised classifiers. Crop identification is then attached to the fields and/or clusters by statistically testing for the similarity between the predicted signatures (assuming a given crop - one regression equation set per crop per phenological phase) and the actual signatures in each LANDSAT band for the given field or cluster in the recognition segment. Statistical tests of similarity include single band students-t tests and multiple band multiple comparison tests such as those given by Scheffe (1959).

The crop type having the least significant statistical difference between predicted versus actual spectral signatures is assigned to the field or cluster in question in the recognition segment according to mean and covariance statistics for crop types obtained from the individual fields or in the statistical clusters in the recognition segment identified by crop type in the previous step.

SUBTASK 2: Acquisition and Preprocessing of LANDSAT and Supporting Data

Work to date has focused on the southwest and central crop reporting districts in Kansas. All LANDSAT 9 x 9 inch transparency data have been obtained for these areas for the 1973-74 crop year except for the color infrared composites of Morton Co., Kansas as noted in the Section on Task I. 1974-75 crop year LANDSAT transparency data acquisition is as yet incomplete.

JSC digital data for the LACIE 1973-1974 Kansas R&D segment set is on hand with the associated crop I.D. and ancilliary information. Similar digital SRS segment data and associated descriptive information for the Kansas 1973-74 crop year is now arriving from JSC. U.C. Berkeley is presently reserving a request for full or partial scene LANDSAT digital data for a portion of Kansas to allow a broader segment population on which to test its technique.

Regional soil map data for Kansas has been obtained and the published county-wide soil survey data (USDA S.C.S.) for the two crop reporting districts is on file. Climatic ground station data for the 1972-1975 time period for Kansas has been obtained from the U.S. Department of Commerce, Environmental Data Service. Airport ground horizontal visibility data for selected dates 1973-1974 in Kansas is now available. Historical data, crop calendar, and cultural practice information to be used in TASK II parallels that obtained and described for TASK I.

Data for the complete western two-thirds of Kansas desired for stratification by our LACIE technical monitors is incomplete with respect to LANDSAT '73-'74, '74-'75 transparency coverage. Most of this data should be acquired for stratification purposes in the next two months.

Most data for the North Dakota Study Area as well as for the other to be defined study area (another state or portion of a foreign country) is not yet available. The exceptions to this statement include complete ground station climatological data for the period 1972-1975 for North Dakota, Texas, Illinois, Indiana, Montana, Washington, and the world as published by the Department of Commerce. Also significant historical, land use, cropping practice, and general soil/climate data is presently being obtained for the Soviet Union. The Russian data is being obtained for a tentative U.C. Berkeley LACIE TASK III effort, but could be used for TASK II signature extension research as well.

SUBTASK 3: Stratification for Analogue Areas

Stratification using one form of "first case" stratification strategy discussed previously will soon (2 weeks) be completed for the southwest crop reporting district in Kansas. Completion of first-case stratification of wheat analogue areas should be completed in the central Kansas crop reporting district soon thereafter. Stratification variables receiving the most significant attention at this iteration are general soil/land use/crop practice strata, growing season (first-case definition -- the three month period following the month of the vernal equinox, i.e. for Kansas, the April to June period), day-degree sum, and the growing season average precipitation. The average last-date-of-freeze and the vernalization factor of average temperature or average minimum temperature in the coldest month of the year (correlated with indices of number of days below 5^{o} C., etc.) are being considered for inclusion in the stratification scheme.

Stratification for the remainder of the western two-thirds of Kansas has been limited primarily due to the lack of several LANDSAT I transparencies which have been requested through our TASK II monitor and are, as we understand, forth-coming, and to manpower limitations.

Stratification for wheat analogue area definition in other states (countries) has not yet begun as per previous milestone schedule guidelines.

SUBTASK 4: Determination of Numerical Constants in the Spectral Signature Prediction Equations

Predictor variable values are being calculated for 1973-'74 LACIE sample segments in Kansas. For a simulated set of training segments, regression coefficients for each prediction variable will be estimated according to least squares analysis of paired LANDSAT band, mean spectral data and corresponding values of signature prediction variables.

Regression constants will be estimated for each regression equation needed to represent each important combination of crop type and phenological phase. These regression equations will then be applied to fields and clusters not used to develop the regression coefficients, but for which the true crop identity is known, to confirm that each equation is accurately predicting the expected signature for the crop in question. When predicted and actual signatures for the same crop do not match according to the statistical similarity tests, then the training segment and training set selection procedures will be modified.

Regression coefficients are being determined for the southwest and central Kansas crop reporting districts as first priority. These two crop reporting districts contain all five intensive test sites in Kansas. Temperature, precipitation, and transmittance prediction variables are being calculated for a sample segment as a whole according to a weighted average of values from the nearest ground meteorological stations. The weight for a given station value is inversely proportional to its distance from the segment center. Other data such as in-segment LANDSAT band ratios and crop maturity is presently being considered on a field or cluster specific basis.

SUBTASK 5: Application of the Spectral Signature Prediction Equations to Corresponding Simulated or Actual Remote Analogue Areas.

This last step in approach to signature extension awaits further development in the foregoing steps. As a first priority, signature extension by use of regression prediction equations within the same wheat analogue area type will be performed in and between the southwest and central Kansas crop reporting districts. Signature extension over larger areas of Kansas and other states will follow according to the milestone schedule.

TASK III: (4.1.3.1b) Variable Probability Sampling for Acreage Estimation

The objective of this Task is to develop and test photo interpretation techniques for definition of wheat sampling strategy. Essentially, this task involves the determination of the photo interpreter's ability to provide a quick estimate of wheat acreage in the population of sample segments in order to provide better wheat density information on which to allocate sample segments for computer information. It is important to note that in this first sample stage photo-interpreter information can also provide a much more precise estimate of wheat acreage by reporting unit and potentially, significantly reduce the number of the much more expensive computer analyzed segments needed at the second stage (phase) of the sample. In other words, if quick-look photo interpreter estimates of wheat acreage by sample segment are significantly correlated with computer results and/or ground truth (R&D), then significant improvements in wheat acreage estimation accuracy, precision, and cost may be obtainable.

Work to date on this task has concentrated on selecting areas for testing and demonstration. Two crop reporting districts in Kansas, the southwest and central, have been sleected for the winter wheat type. Sample grids (a matrix of 5 x 6 mi sample segments) are being constructed for overlay onto LANDSAT transparencies and signature extension LANDSAT print mosaics. The sample grids themselves are being constructed for the state of Kansas. The portions overlaying the two crop reporting districts mentioned will then be used for photo-interpretation.

Photo interpretation will proceed in the first-case demonstration on early July 1973 LANDSAT imagery. This single growing season date for western and central Kansas appears to contain a maximum of wheat acreage information with a minimum of crop confusion information for the quick look approach. At this time of year in this winter wheat area most dryland and irrigated wheat fields have a distinctive yellow to beige tone on the false color LANDSAT data. Other regions may require more than one date of LANDSAT imagery for optimal results with the "quick-look" aporoach. Further research on the Kansas areas could also indicate a multidate requirement.

Photo interpreter training keys/aids are presently being developed for the two Kansas crop reporting districts cited earlier. These keys are based on cropping practice/soil relationships and crop calendar information as they relate to the image features of wheat on the LANDSAT data selected for interpretation. The training aid information is being derived from data sources cited in U.C. Berkeley Tasks I & II.

Demonstration sites for areas other than Kansas are yet to be finally determined. Recommendations from our technical monitors indicate that Montana or Idaho would be good regions in which to demonstrate the "quick-look" wheat acreage estimation technique for a mountainous area where small agricultural areas are found within interfingering valleys. But the decision as to the final area of study within one of these states awaits an evaluation of the availability of cloud-free properly timed with crop phase imagery. Evaluation of available imagery is now in progress.

Foreign country discussions have centered on the Soviet Union, particularly one oblast or a portion of an oblast. U.C. Berkeley has been studying Soviet Union historical wheat data, climate, soils, and existing LANDSAT data to determine the location and relative concentrations of winter, spring and mixed wheat in that country. The purpose of this Soviet analysis, now complete, was to enable U.C. Berkeley to make a recommendation to our LACIE Task III monitor concerning the location of "quick-look" wheat-acreage demonstration sites.

The recommended Russian areas are either a mixed wheat area in the region defined by the cities of Penza and Kuyushev (north and NNW of the Caspian Sea) or a spring wheat area in northern Kazakh S.S.R. NE of the Caspain Sea). The former area is an established small grains producing region while the latter is located in a "new" agricultural lands area brought into production by irrigation. The latter region has a climatic regime apparently similar in many respects to the Southwest U.S. Variability in wheat acreage from year to year in the new lands area may be quite high. The decision as to which region should be used for demonstration of the Task III "quick-look" acreage estimation technique depends on the sensitivity of the LACIE acreage estimation formulas to the two types of Soviet wheat producing situation cited.

These foreign country recommendation study results, just obtained, have yet to be discussed with our Task III technical monitor.

Literature Cited

Klages, K.H.W. 1942. Ecological crop geography. The Macmillan Company, New York. 615 pp.

Nuttonson, M.Y. 1956. A comparative study of lower and upper limits of temperature in measuring the variability of day-degree summations of wheat, barley, and rye. American Institute of Crop Ecology, Washington, D.C. 42 pp.

Pascale, A.J. and E.A. Damario. 1962. Agroclimatic wheat crop types in the world. In: Biometeorology. Proceedings of The Second International Bioclimatological Congress held at The Royal Society of Medicine, London K-10 Sept. 1960. Pergamon Press, New York.

Peterson, R.F. 1965. Wheat; Botany, Cultivation, and Utilization. Interscience Publishers Inc., New York. 422 pp.

Scheffe, H. 1959. The Analysis of Variance. John Wiley & Sons, Inc. New York. 477 pp.

If there are any questions concerning this report, please feel free to give us a call.

Sincerely yours,

Claire M. Hay
RSRP-LACIE Project Leader

Randall W. Thomas
Technical Services
Branch Manager

cc: Messrs. John Deitrich
Joe Garcia
Al Feiveso
Joe Downs

Enclosure #1

JUN 18 1975

1.1 Task I: Evaluation and Revision of Techniques for Manual Interpretation

1.1.1 Objectives

The primary objective of this proposed task is to evaluate various manual interpretation techniques for the identification of wheat on a world wide basis and to revise them as necessary. Specifically, the techniques will entail the human analysis of LANDSAT imagery and ancillary data for the purpose of providing reliable training data for automatic classification systems in remote areas where ground data are not available.

The emphasis of the proposed manual interpretation work will include the following subtasks: (a) the quantitative evaluation of various LANDSAT image sets, and (b) the definition of the best procedure for the image analyst to use to identify wheat. Note that these two subtasks are closely related. For example, an optimum set of imagery to be tested in Subtask A may have been defined on the basis of temporal information determined in Subtask B. Initially, however, these two subtasks, will be evaluated as independently as possible of each other.

A. Quantitative Evaluation of LANDSAT Imagery

Because of the numerous multitemporal image sets and multiband/multidate image combinations that can be produced from LANDSAT data, the determination of the best set of temporal images needed for accurate crop identification will be a major objective of the proposed study. Specifically, the interpretation accuracy that can be expected from interpreters using 1) optimum photogenic date or combination of dates (sequential technique) during the growing season; and 2) less than optimum photogenic dates or combination of dates during the growing season (i.e. other than an optimum date due to missing images as a result of cloud cover, etc. or to incomplete optimum image set in an attempt to make estimates of wheat acreage early in the growing season) will be determined from interpretation tests.

B. Image Analysis Techniques

The objective of this subtask is to define and evaluate those materials, equipment, and techniques which are required by the image analyst to complete the task of identifying wheat. This will include, in part, determining the utility and optimum mix of the following materials and techniques:

1. reference materials -- crop keys, crop calendars, data from analogous areas, etc.
2. auxilliary data -- high altitude photos, low altitude photos, historical agricultural data, spectral data, meteorological data, etc.
3. convergence of evidence -- field size, cropping patterns, cultural features, etc.
4. conference system -- joint interpretation with two or more analysts.

1.1.2 Approach

Wherever possible, a quantitative evaluation of all the subtasks will be made, in order to determine how the accuracy levels and associated errors vary with the different data sets that are examined. It must be emphasized that in order for any of the results to have meaning with the LACIE, all existing ground data, image sets, crop calenders, cropping practice information, and LANDSAT computer compatible tapes, must be made available to the Remote Sensing Research Program. Not only will this make the results more meaningful with regard to LACIE, but it will minimize duplication of effort between participating organizations.

A. Quantitative Evaluation of LANDSAT Imagery

Image interpretation tests will be given to determine the expected interpretation accuracy of optimum image sets for the identification of wheat. These tests will be given to skilled image analysts from both the RSRP and the LACIE Lockheed support group. Test results will be analyzed using an analysis of variance design so that the levels of accuracy attainable from the different sets and their confidence limits can be determined.

Most of the image sets that will be tested will be produced optically, photographically, and/or electronically by personnel from the Johnson Spacecraft Center and the RSRP. In order to completely evaluate the state-of-the-art in image production, however, a limited number of image sets will be purchased from educational facilities and commercial organizations for testing purposes.

B. Image Analysis Techniques

The results of the image interpretation tests given to evaluate the image sets will be used also to evaluate the image analysis process, in order to determine (1) the contributions to identification accuracy made by various crop keys and ancillary data, and (2) the effect that missing data have upon usefulness of these data sets.

1.1.3 Expected Results

The expected results include the documentation, evaluation, and revision of techniques for the manual interpretation of LANDSAT scanner data for the purposes of training a computer algorithm to identify wheat.

These results will include: (1) a description of the optimum LANDSAT image sets for identification of wheat along with the accuracies of interpretation expected using the optimum sets as well as less than optimum sets, and (2) a description of the necessary and sufficient ancillary data needed by the interpreters.

1.1.4 Milestone Plan

The milestone plan for Task I is given in Table 1. It should be noted that all time lines will shift to the right proportional to the delay in receiving LANDSAT and auxilliary data.

TABLE 1. MILESTONE PLAN FOR LACIE TASK I.

Subtask A. Quantitative Evaluation of Landsat Imagery Set(s).

1. Determination of optimum sequential data set(s) for Kansas test sets (winter wheat areas). (Data expected from JSC by 7/15/75.)

   a. Rank all dates according to their effectiveness in identifying wheat. (10/15/75)

   b. Description of optimum sequential data set(s) for Kansas test sites - best 4 or 5. (10/15/75)

   c. Interpretation accuracy expected from optimum data set for Kansas test site as determined from the interpretation tests. (10/15/75)

   d. Interpretation accuracy expected from less than optimum data sets for Kansas test sites. (10/15/75)

2. Determination of optimum sequential data set(s) for North Dakota test sites (spring wheat areas). (Data expected from JSC by 8/1/75.)

   a. Rank all dates according to their effectiveness in identifying wheat. (12/31/75)

   b. Description of the optimum sequential data set(s) for North Dakota spring wheat sites. (12/31/75)

   c. Interpretation accuracy expected from optimum data set for North Dakota spring wheat test sites. (12/31/75)

   d. Interpretation accuracy expected from less than optimum sequential data sets for North Dakota spring wheat test sites. (12/31/75)

3. Determination of optimum sequential data set(s) for an area of mixed winter and spring wheat production (possibly Montana or South Dakota). (Data expected from JSC by 8/1/75.)

   a. Rank all dates according to their effectiveness in identifying winter and spring wheat. (3/1/76)

   b. Description of optimum sequential data set(s) for an area of mixed spring and winter wheat production. (3/1/76)

c. Interpretation accuracy expected from optimum sequential data set for an area of mixed spring and winter wheat production (3/1/76)

d. Interpretation accuracies expected from less than optimum data sets for an area of mixed spring and winter wheat production. (3/1/76)

Subtask B. Evaluation of Image Analysis Techniques

Description of necessary and sufficient data inputs (imagery and ancillary data) for the interpretation of winter wheat based on the Kansas data, spring wheat based on the North Dakota data, and mixed spring and winter wheat based on the data of an area of mixed production-possibly Montana or South Dakota. (5/14/76)

a. Large area full frame input to the interpreter vs. small area test site only (5 X 6 mile segment) input to the interpreter. (5/14/76)

1.2 Task II: Development of a Spectral Signature Extension Methodology

1.2.1 Objectives

The primary objective of this task is to develop analogue stratification strategies and spectral signature prediction models that will allow accurate, cost-effective classification of wheat areas in agricultural regions of the world lacking current ground data.

The development of this spectral signature extension methodology will involve five supporting subtasks:

A. Spectral Signature Extension Methodology Design

This subtask will involve (1) development of stratification strategy for analogue wheat areas, (2) selection of the spectral signature prediction model to be employed and a definition of variables utilized in the model, and (3) design of supporting sampling strategies. These design specifications will guide and coordinate progress on the remaining subtasks.

B. Acquisition and preprocessing of LANDSAT and Supporting Data

This subtask will include acquisition and indexing of LANDSAT data and supporting ground and aerial data, as well as meteorological, soil, and plant culture information as required by design specifications for stratification and model variables. Transformation of data planes to a common LANDSAT coordinate system and subsequent information merging will also be performed where necessary. Integrated information management as this stage will permit most efficient processing in the remaining three subtasks.

C. Stratification for Analogue Areas

Application of the criteria for delineating analogue areas over specified regions will be performed with subsequent refinement of stratification methodology. Analogue areas are defined as agricultural lands potentially giving rise to similar wheat signatures (spectral and spatial) over the life cycle of wheat due to similar climate and soils conditions as well as similar cultural practices. After strafification, each agricultural stratum will define a particular wheat analogue area type. Within a given wheat analogue area type there will exist non-remote analogue areas (current ground data available) and remote analogue areas (no significant current ground data available). Stratification of regions for analogue areas is designed to significantly reduce variation in predicted spectral signatures and also the cost of signature extention and stratification.

D. Determination of Numerical Constants in the Spectral Signature Prediction Equations

For wheat and non-wheat training fields selected by image analysts in non-remote analogue areas parameters will be estimated for the spectral signature preduction equation for each analogue area type. Tests for statistical significance of variables included in each equation will then be conducted. A determination of the accuracy of the signature prediction models within non-remote analogue areas will allow further refinement of model parameters. The objective of this subtask is to obtain the ability to generate accurate channel means and variances for each wheat analogue area type under varying conditions.

E. Application of the Spectral Signature Prediction Equations to Corresponding Simulated or Actual Remote Analogue Areas

Within remote analogue areas, data from wheat and non-wheat fields located by image analysts will be entered into the spectral signature prediction equations particular to those strata. Predicted spectral estimates will then be statistically compared by channel to the corresponding actual remote area spectral signature means. If statistically significant differences occur in critical channels, then retraining for wheat and non-wheat locations in remote areas will be initiated. Otherwise, wheat areas will be classified in the given remote analogue areas according to spectral statistics particular to each remote area's training fields. Finally, confidence intervals around wheat acreage estimates in analogue areas will be constructed and checks on classification accuracies will be performed where possible.

1.2.2 Approach

The five subtasks outlined in Section 1.2.1 for the development of an accurate, cost-effective spectral signature extension methodology will be conducted in a coordinated manner both stepwise and in some instances simultaneously. The overall relationship of the five subtasks is shown in Table 2. Methodology design and data acquisition must necessarily be the overlapping first steps followed by analogue area stratification, model parameter determination, and finally model application to simulated or actual remote analogue areas for wheat classification. On time completion of all tasks will be dependent in a sliding scale manner on the on-time RSRP receipt of funds and data from LACIE management. Specific approaches for each of the five subtasks will be as follows:

A. Spectral Signature Extension Methodology Design

Stratification strategy for analogue areas will involve 1) a specification of important stratification variables including soil and meteorological conditions, cultural practices, wheat types and confusion crops grown, and 2) selection of human and human-automatic techniques to allow rapid delineation of strata according to stratification variables of interest. An important point to note is that a given wheat analogue area type must occur within the same agricultural stratum. This means that a given stratum will be disjoint, that is, it will occur in several separate places over the globe.

The spectral signature prediction model has been tentatively selected as the analysis of covariance model. This is a linear probabilistic relationship that includes qualitative factors and their interactions as in analysis of variance models while at the same time providing for quantitative variables. Hence the statistical significance of several spectral signature modifying factors in a p-way layout can be cost-effectively and realistically tested after correction for regression on a quantitative string of other variables also impacting on spectral signature. The dependent variable in the analysis of covariance model will be taken to represent the predicted mean value for a given spectral and/or spatial information channel for a given analogue area according to the values of the signature affecting variables existing in the given analogue area at a given time.

The selection of which qualitative and quantitative factors and their interactions to include in the analysis of covariance model will depend on factors identified in the literature, in RSRP research, and in forums with other LACIE signature extension working group members as having significant effects on spectral signature. Moreover, any factor or variable included in the model may be complex. That is, raw temperature or humidity values may not be used directly, but some physiologically related quantities generated by deterministic relationships could be employed in the model instead.

Supporting sampling strategies utilizing LANDSAT, aerial, and ground data will be used to calibrate signatures where ground data are obtainable to improve the precision and accuracies of the spectral signature extension equations. These sampling strategies will be designed to be cost-efficient and workable in real-time wheat acreage estimation situations. Moreover, estimation of variables employed in the signature prediction models will be based on sampling plans involving cost-efficient combinations of data.

B. Acquisition (and Preprocessing) of LANDSAT and Supporting Data

The approach to an organized and readily manipulatable data system will involve firstly an acquisition and indexing of LANDSAT digital tapes and transparencies for areas specified in the study. Digital tape reformating, calibration, study area extraction, and if necessary, generation of a common LANDSAT multidate coordinate system will follow.

Acquisition and indexing of supporting aerial and ground data will be performed in non-remote analogue study areas and auxillary information planes, including topographic, meteorological and cultural, will be obtained to the extent possible for all study areas.

A generation of auxillary data planes providing information important in the stratification process and the analysis of covariance spectral signature prediction model will then be performed and, when necessary, transformed to the common LANDSAT multidate coordinate system.

Finally the LANDSAT and auxillary data planes will be merged where appropriate for stratification, model development, and wheat classification.

C. Stratification of Analogue Areas

Stratification for wheat analogue areas in specified study regions according to variables determined important in defining homogeneous wheat producing locations will be conducted according to manual or optionally manual-automatic technologies. Manual approaches will involve delineation of analogue areas on hardcopy false-color enhancements of LANDSAT and superimposed auxillary data combinations. Auxillary data may include general climatic, soil fertility, drainage, and cultural practice indices. Human-automatic stratification techniques might involve human location of analogue area boundary coordinates on manipulatible TV monitor displays of merged LANDSAT climatic, soil, and cultural practice information.

Analogue area stratification technology development for the period to December 15, 1975 will focus on the winter wheat areas of the State of Kansas, particularly the designated LACIE intensive test site(s) where special data for signature extension research will be collected by LACIE cooperators. Delay in RSRP receipt of data will cause delays in task completion.

For the period December 16, 1975 to May 14, 1976 analogue area stratification technology development will focus on spring wheat production areas of North Dakota, particularly the designated LACIE test site(s) where special data for signature extension research will be collected for LACIE cooperators. Again, delay in RSRP receipt of data will potentially cause delays in task completion.

D. Determination of Numerical Constants in the Spectral Signature Prediction Equations

Image analyst selection of training fields for estimation of spectral signature prediction equation parameters (i.e. numerical constants) will be based on crop identification information gained in non-remote study areas. Where possible, crop identification information will be determined from analysis of LANDSAT, aerial, and ground data collected in cost-effective sampling schemes. These sampling plans may be the same as those designed for LACIE wheat acreage over non-remote analogue areas.

Estimation of model parameters will proceed according to least-squares criteria for data derived from the geometrically coincident LANDSAT and auxillary information planes occupying selected training field locations. Training fields will be located so as to provide information for a complete p-way layout with at least two samples per cell. Thus all parameters for interactions of interest in the specttal signature prediction model should be estimable.

All parameters for the p qualitative (presence-absence) variables or their interactions in the analysis of covariance model will be determined by calculating their values according to usual analysis of variance procedures and then subtracting from each such numerical constant a correction for regression on the quantitative (continuous) variables not controlled in the p-way layout. For instance, assume that the number of qualitative variables, p, is equal to three. Further assume that these three qualitative variables consist of wheat analogue area type, life cycle stage, and class of irrigation frequency. Then the effects (expressed as numerical constants) of those variables on wheat signature would be calculated according to usual analysis of variance rules and then corrected by regression of wheat signature on such quantitative variables as sun angle, inches of rain for a given time period, etc. In all cases the values of the dependent variable will be generated from the mean numerical values for given information channels on given training fields. In this way, the true importance or magnitude of contribution to a predicted spectral signature can be assessed for each combination of qualitative factors after adjustment for the effect of many freely varying environmental variables.

F-tests and multiple comparison tests including the Scheffe method will be made to determine the statistical significance of qualitative and quantitative variables and associated interactions included in the spectral signature prediction equation for each wheat analogue area type. Those terms not found to be significant will be dropped in order to streamline the prediction model. In this way a cost-effective model will be generated utilizing only data found to be significantly correlated with wheat signatures. Furthermore, estimates of variables important to signature prediction will be based on a cost-efficient sampling process utilized in the preprocessing subtask.

Development of software for the analysis of covariance prediction model and associated tests of statistical significance for model parameters will be designed for compatibility with current data processing systems.

A test will then be conducted in non-remote areas for the accuracy of wheat field signature identification by the particular spectral signature prediction model associated with each wheat analogue area type. Where signatures for known wheat fields are not accurately estimated, then further training field location and model parameter estimation will be conducted until statistically adequate signature predictions are made.

E. Application of the Spectral Signature Prediction Equations to Corresponding Remote Analogue Areas

Image analyst selection of training fields in remote or simulated remote analogue areas will be based on inferences of wheat presence from various LANDSAT spectral and spatial enhancements and on auxiliary data such as crop calendars. Other LACIE work on remote identification of wheat will have important application here. Simulated remote areas will consist of non-remote locations treated as if they were remote. Aerial and ground data for these simulated remote locations will allow subsequent determinations of the accuracy of wheat classification after spectral signature extension.

For wheat training fields selected in remote or simulated remote analogue areas, merged LANDSAT and auxiliary data will be directly substituted into the spectral signature prediction equations particular to the given wheat analogue area types. The resulting signature estimates by channel will then be statistically compared through system compatible software with known corresponding values from remote areas. These determinations of statistical significance will be tests of the null hypothesis that remote training fields are in fact wheat. Candidate tests include simple pairwise t-tests between predicted versus known channel means and multiple comparison tests between several predicted versus known channel means.

If statistically significant differences in key channels occur, then retraining for wheat fields in remote areas must occur. This feedback to the remote wheat identification task will indicate whether further refinement of associated remote identification techniques is necessary for specific wheat analogue area types. Retraining will enable generation of a new set of wheat spectral signature predictions.

If no critical statistically significant differences in channel means occur between known and predicted wheat signatures then wheat will be classified in the remote areas according to usual methods. Specifically, training statistics from fields identified as wheat by the spectral signature prediction equation will be used to classify wheat acreage in sample units for the particular remote analogue area. The size, shape, and number of sample units selected per analogue area will of course depend on the wheat acreage estimation procedures used for LACIE. For most accurate wheat classification, it may be necessary to identify non-wheat confusion crop training locations also by spectral signature prediction equations and then use the statistics for those non-wheat types in the classifier as other classes.

Confidence intervals around wheat acreage estimates will be formulated according to usual methods. It is important to recognize the overall significance of the confidence interval argument in the proposed spectral signature extension approach outlined here. Among other applications, confidence intervals will be constructed around 1) estimated raw data variables, 2) estimated variables used in the spectral signature equations, and 3) average channel values estimated by the prediction equations for remote areas, in order to determine the reliability of the ultimate acreage estimate. These confidence intervals will also be used as criteria to adjust spectral signature prediction results when input data are lacking or incomplete in remote areas. As such the intervals will constitute a portion of the sensitivity analysis for missing data.

1.2.3 Expected Results

The primary results from the proposed investigation will include a documentation of techniques employed in the five subtasks and an evaluation of the timeliness, cost-effectiveness, and accuracy with which wheat can be classified in simulated remote analogue areas. Any workable technology that is developed in these tests to employ spectral signature extension methodology for large, real-time global inventories should be transferable to user agencies. This technology transfer will involve specifications for hardware modification in initial and more advanced spectral signature extension methodologies and will also include all software particular to the signature extension process.

1.2.4 Milestone Plan

The milestone plan for Task II is given in Table 3. It should be noted that all time lines will shift proportional to the delay in receiving LANDSAT and auxiliary data.

TABLE 3 MILESTONE PLAN FOR LACIE TASK II.

1. Initial Spectral Signature Extension (SSE) Methodology Design. (Most of Methodology Design will be developed using Kansas data.)

   a. Development of a stratification strategy: variables, techniques, wheat life cycle relations between analogue areas, and criteria for analogue areas. (Brief report or oral presentation for Kansas (9/1/75)
   final report due (12/15/75)

   b. Selection of SSE prediction model type(s) and variables
   the model - 7/1/75
   the variables - 10/1/75)

   c. Design of statistical tests for variable inclusion within the SSE model(s). (9/15/75)

   d. Specification of statistical decision algorithm to determine significant differences between actual and predicted signatures. (9/15/75)

   e. Design of SSE sensitivity analysis tests to quantify performance and isolate error sources. (Preliminary report - 10/15/75)
   Final report - 12/15/75)

   f. Design of SSE supporting sampling strategies. (10/1/75)

2. Acquisition of Landsat and supporting data for selected intensive test sites.

   a. Identify data requirements for Kansas (6/20/75)
   Identify data requirements for North Dakota. (9/1/75)
   Identify data requirements for Other. (9/15/75)

   b. Acquisition of needed data for Kansas (7/15/75)
   Acquisition of needed data for North Dakota (10/1/75)
   Acquisition of needed data for Other. (10/15/75)

3. Stratification of Analogue Areas.

   Application of stratification criteria to selected [illegible]

   a. Kansas (Final report - 12/15/75)

   b. North Dakota (Final report - 5/14/76)

c. Other (Final report - 5/14/76)

4. Determination of Numerical Constants in SSE Prediction Equations including Variable Significance Testing and SSE Prediction Accuracy.

Documentation for selected ITS for

a. Kansas (Final report - 12/15/75)

b. North Dakota (Final report - 5/14/76)

c. Other (Final report - 5/14/76)

5. Application of SSE Prediction Equations to Corresponding Simulated Remote Analogue Areas for selected ITS

a. Applications of tests for statistical differences between predicted and actual signatures in fields located in current ground dataless or simulated ground dataless areas which are identified by photo interpretation as being wheat. Checks for classification accuracy where ground data is actually available.

1. Kansas (12/15/75)
2. North Dakota (5/14/76)
3. Other (5/14/76)

b. Identification of optimal SSE prediction information/photo interpreter remote wheat field location feedback procedures.
(Preliminary assessment - 12/15/75)
final report - 5/14/76)

6. Sensitivity Analysis in Selected Intensive Test Sites to Quantify SSE Prediction Performance Relative to Input Types, Quality and Quantity. (Preliminary assessment - 12/15/75
final report - 5/14/76)

1.3 Task III: Variable Probability Sampling for Acreage Estimation (Task 4.1.3.1b, (RSRP, U.C. Berkeley)

1.3.1 Introduction

The human's ability to recognize wheat growing areas and make quick estimates of wheat acreage using inexpensive photographic products can be incorporated into a variable probability sampling scheme. Under this procedure, sample units (segments) are selected with probability proportional to estimated size (PPES) of the acreage of wheat determined by ocular techniques.

1.3.2 Objective

The objective of this task is to develop and test photo interpretation techniques for definition of wheat sampling strategy and to define the PPES estimation formula associated with the sampling strategy. The desired product is a variable probability sampling scheme utilizing image interpretation that may be used, within specified application criteria, to allocate samples for unbiased wheat acreage estimation.

1.3.3 Overall Approach

The extent to which photo interpreters can be used to allocate samples in an unbiased and cost effective manner for wheat acreage estimation will be determined. To do this, seven subtasks will be accomplished.

A. Photo Interpreter Wheat Estimates

The ability of the photo interpreter to estimate wheat acreage of other variables that can be correlated with wheat acreage and the extent of the correlation between the photo interpreter's estimate and computer estimate will be determined. This will be done in areas defined for RSRP LACIE Study within the U.S. that represent, as well as possible, the conditions encountered in the non-U.S. areas. The ability of the photo interpretation scheme to handle the extremes in field size found between specified areas such as found in the U.S.S.R., China and India, will be specifically considered through the use of U.S. analogues and the actual imagery from the countries.

B. Definition of PPES Estimation Formulae

As a result of the investigation outlined in (A), detailed definition of the PPES estimators, the derivation of the estimators, the derivation of the estimators of variance, presentation of sample size and sampling rate formulas, and the sample allocation method to meet the system's constraints will be provided.

C. Estimation of Sample Size

After determining the correlation between the computer predictions and human predictions of wheat acreage, the sample size requirements specific to the wheat growing regions under study by RSRP in the LACIE will be estimated.

D. Human-Computer Analysis of Segment Data

If the human input proves cost-effective, a general procedure for human and possible computer analysis of segments will be developed for the various wheat producing areas of the world. This will include the documentation of the advantages, as well as the limitations, unusual situations that may be encountered in the application of these techniques and the image types, and procedures required to estimate the selected wheat correlated parameters for spring, winter, and mixed wheat areas.

E. Effect of Missing Data (minimum effort)

The best possible way to handle the problem of imagery which is missing and which is also important to the PPES-photo interpretation scheme will be determined. This procedure may entail the use of existing maps, meteorological data, and historical acreage data to generate surrogate images to use in the photo interpretation estimation procedure.

F. Change Detection (low priority)

To the extent that any significant changes can be detected within the constraints of the existing sequential data base for the intensive sites being studied, techniques will be evaluated to assess the impact of land use shifts on the sample allocation procedure. The PPES allocation using previous years of data will undoubtedly reduce the ability of the system to detect this type of change. However, this problem can be overcome through a partial replacement sampling scheme that will allow the inclusion of areas where changes are likely to occur.

1.3.4 Detailed Task Descriptions

The seven sub-tasks outlined in Section 1.3.3 will be conducted in a stepwise, coordinated manner. On-time completion of all tasks will be dependent upon the timely receipt by the RSRP of data from LACIE management. The specific approaches for each of the seven sub-tasks will be as follows:

A. Photo Interpreter Wheat Estimates

Optimal human image-interpretation methods for wheat acreage estimation will be defined, described, and quantified. "Optimal" will be defined as that combination of photo interpretation techniques and image sets allowing the highest correlations between photo interpreter wheat acreage estimates and "known" (e.g. computer or ground truth) acreages within given cost and time constraints. This research will be based on (1) the optimal image interpretation technique combinations defined in the RSRP's LACIE Tasks I (4.1.1.1a) and II (4.1.1.2f(1) and 4.1.1.2f(3), and (2) investigations to determine the optimal image sets along with photo interpreter techniques that will be conducted specifically for Task III (4.1.31b).

Optimal image-interpretation technique combinations will be defined for major wheat producing types and spatial arrangement patterns within LACIE RSRP Study areas that represent, as well as possible, conditions encountered in U.S. and non-U.S. areas. As a minimum, the ability of the PPES photo interpretation scheme to handle extremes in field size and spatial arrangement between the USSR, China, and India will be specifically considered through the use of U.S. analogues and the actual imagery and supporting auxillary data from the countries. However, in order to insure adequate evaluation of the photo interpretation techniques, the PPES scheme, and RSRP's Task I and Task II techniques, data from all countries involved in the LACIE experiment should be available for research.

B. Definition of PPES Estimation Formulae

Detailed definition of the PPES estimators, the derivation of the estimators, and the estimators of variance will be conducted with a view to current and eventual LACIE system constraints. An example of a current constraint that would affect the mathematical representation of the estimators is the systematic nature of eligible sample segment location imposed by the GSFC system. Estimator bias and precision problems associated with such current constraints will be identified and contrasted with situations associated with alternative PPES derivations under less restrictive LACIE sampling segment locations guidelines.

Derivation of sample size formulas will be presented along with the overall sample design under which sample segments may be allocated to specified LACIE countries, regions, zones, and acreage strata. Sample size formulas and sampling designs will be specific to given LACIE system constraint sets. Countries, regions, and zones will be as defined by LACIE cooperators and management. Acreage strata will be defined on the basis of approximately constant levels of wheat density as photo interpreted and checked with historical data so as to minimize the variation in wheat acreage estimates.

C. Estimation of Sample Size

For a given LACIE constraint set and associated PPES estimators and sample size formulas, sample segment sizes will be calculated for specific country, region, zone, acreage strata within LACIE RSRP Study areas. This task involves acquisition (from JCS), indexing, preprocessing and data merging not already performed under Tasks I and II for the areas of interest. In order to calculate sample sizes, acreage strata must be defined within LACIE reporting units, photo-interpreter estimates of wheat acreages must be obtained, and in some cases ground wheat acreage data variance calculated where such information is available. In addition, the cost structure of the PPES acreage estimation technique will be determined including sample unit costs at various stages or phases in the sampling acreage estimates will determine, in part, optimal sample rates for the respective estimation techniques. These calculations will be reported for various regions at times specified in RSRP's LACIE Task 4.1.3.1b milestone plan, Table 4. Delay in RSRP receipt of data will potentially delay sample size reporting.

D. Human-Computer Analysis of Segment Data

Based on the work in the foregoing four sub-tasks, a general procedure will be outlined for combining photo-interpretation results and computer estimates in wheat acreage estimators. The PPES sample design will be specifically considered, though the application of the techniques developed to date in other sample designs employed in LACIE will be considered as well. Included will be a documentation of advantages, limitations, and unusual situations to be encountered in the use of the human-computer segment analysis for wheat acreage estimation. Image types and interpretation procedures required to estimate selected wheat correlated parameters for major wheat producing area types considered in LACIE RSRP Study areas will also be given.

E. Effect of Missing Data (minimum effort)

When missing ERTS imagery for given wheat production areas occurs, either due to cloud, hardware, or processing problems, techniques will be developed to generate estimated photo-interpreter wheat acreage estimates. The approach will be (a) to generate and apply regression relationships between previous photo-interpreter estimates for the missing data areas and current photo-interpreter and/or computer estimates in similar areas where data is not currently missing; or (b) generate and apply regression relationships between most recent photo-interpreter estimates and the following: existing map data, historical acreage data including past photo-interpreter and computer estimates, current computer estimates, and/or indices of meteorological conditions. Application of the regression relationships will allow least-square estimates of current photo-interpreter wheat acreage estimates in areas currently missing data. Combined with this missing data approach will be a preliminary analysis of its effect on PPES estimator cost, accuracy, and precision performance.

F. Change Detection (low priority)

It is presently conceived that complete photo-interpreter estimates of wheat acreages will be formulated only at the initial inventory period for those areas selected by LACIE management for subsequent photo-interpreter PPES application. In order to inexpensively detect land use conversion to or from wheat production, it is proposed that a photo-interpreter sample check of wheat acreage estimates be made at subyearly, yearly, or multiyearly intervals. Based on photo-interpreter results, sample segments for computer analysis can be reallocated by a partial replacement strategy at similar intervals in order to minimize bias in the wheat acreage estimates.

In addition, photo-interpreter line stratification of land use change as it relates to efficient sample reallocation will be examined.

1.3.5 Resource Data Requirements	Date Required by RSRP
I. Existing mosaics for State of Kansas	7/15/75
II. Existing mosaics for other states (Idaho or Montana?)	10/1/75
III. Existing mosaics for selected Foreign Area (Oblast?)	10/15/75

		Date Required by RSRP
IV.	Updated state and country mosaics as they become available	as available
V.	Selected LANDSAT color infrared composite transparencies	
	a) Kansas	7/15/75
	b) Other state (Idaho or Montana?)	10/1/75
	c) Selected Foreign Area (Oblast?)	10/15/75
VI.	Supporting cultural practice, climatic, wheat acreage, wheat variety, yield, soils, and other support data (TBD) for the above-named states and countries when not available through other LACIE U.C. Berkeley Tasks.	
	a) Kansas	7/15/75
	b) Other state	10/1/75
	c) Selected Foreign Area	10/15/75
VII.	Field identification data, including SRS and other available information for above-named states and countries when not available through other LACIE U.C. Berkeley Tasks.	
	a) Kansas	7/15/75
	b) Other state	10/1/75
	c) Selected Foreign Area	10/15/75
VIII.	Detailed description of capabilities and restrictions of hardware and software components (existing and developing) available to LACIE and JSC and Goddard	7/15/75
IX.	Complete description of the current LACIE sampling plans for the USA and other LACIE countries	7/15/75
X.	Description of other sampling R & D approaches, locations of application (e.g. particular areas and reporting district types of Kansas), and estimator results and variances.	7/15/75
XI.	Crop/land use (wheat and other crops/land use types where available) proportions from computer classification for segments in areas defined for study. Also, computer classification accuracies by segment by crop/land use type where available.	
	a) Kansas	7/15/75
	b) Other state	10/1/75
	c) Selected Foreign Area	10/15/75

TABLE 4 MILESTONE PLAN FOR LACIE TASK III.

1. Determination of Photo Interpreter Wheat Estimation Abilities for selected North American (NA) Wheat Producing Areas.

   a. Kansas (counties) (Results 12/31/75)
   b. Idaho (North Dakota?), Other state (5/14/76)
   c. Selected Foreign Area (5/14/76)

2. Definition of PPS Estimation Formulae

   Includes the sample design for the wheat acreage estimation problem along with the associated estimators, estimators of variance, sample size and rate formulas, and the specification of wheat acreage stratification strategy.

   (Preliminary documentation - enough to implement plan - 12/31/75
   Final report - 5/14/76)

3. Estimation of Sample Size

   Includes any additional data acquisition from JSC and data processing not already specified by Tasks I and II; also includes acreage stratification, photo interpreter acreage estimation, selected ground data variance analysis, sampling cost analysis.

   Analysis for selected North American areas:

   a. Kansas (12/31/75)
   b. Other states (North Dakota, Idaho??) (5/14/76)
   c. Oblast? (5/14/76)

4. Results of Wheat Acreage Estimation including proportions and variance of proportions and related statistics summarized by counties and/or other reporting districts thought to be desirable. Included also, if necessary, are maps of counties and/or other reporting districts showing acreage strata and associated proportion estimates.

   a. Kansas (12/31/75)
   b. Other state (5/14/76)
   c. Oblast (5/14/76)

5. Develop a General Procedure and Criteria for Human and Computer Analysis of Sample Segments (5/14/76)

6. Determine Methods to Develop Photo Interpreter Estimates When Missing Data Occurs (5/14/76)

7. Develop techniques to Detect Land Use Change Affecting Wheat Acreage Estimates (5/14/76)

Enclosure #2

U.C. BERKELEY SIGNATURE EXTENSION APPROACH

(1) Stratify regions into analogue areas tending to give rise to a similar sequence of wheat signatures.

(2) Identify land cover signatures in recognition segments by utilizing land cover, phenological phase specific signature prediction regression equations developed in training segments.

(3) Classify recognition segment based on training signatures from the recognition segment identified by the process in (2).

STRATIFICATION

Analogue areas are strata of specified extent defined by relatively homogeneous combinations of values for environment variables such as

- General Soil Type
- Land Use/Cropping Practice
- Growing Season Day-Degree Sum
- Growing Season Precipitation Sum
- Average Last Date of Spring Frost
- Average Temperature for the Coldest Month

Analogue area boundaries may be updated when desired, generally every several years.

Spectral Signature Statistics for General Cover Types Within Recognition Segments

(1) Within a given analogue area type, signatures are extended between the training segment(s) and the recognition segment(s) by use of regression equations.

(2) There is one regression equation developed from training segments for each combination of major cover type, major phenological phase, and LANDSAT channel.

(3) Each regression equation is designed to predict the spectral signature in a given LANDSAT band for a given cover type according to predictor variables derived from the recognition segment such as

- Crop Maturity Codes from A.I.s
- Soil Color Code from A.I.s
- Growing Season Day-Degree Sum
- Planting Season Day-Degree Sum
- Growing Season Precipitation Sum
- Planting Season Precipitation Sum
- Sun Angle
- Scan Angle
- In-Segment LANDSAT Spectral Data Such as the 7/5 Ratio (Crop Calendar Related)
- Atmospheric Transmittance (Generated from Ground Horizontal Visibility Data)

(4) Predicted signatures are statistically matched to fields or clusters in the recognition segment

(5) The spectral statistics from the matched fields or clusters in the recognition segment are then used for recognition segment discriminant analysis.

Enclosure #3

UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO | SANTA BARBARA • SANTA CRUZ

SPACE SCIENCES LABORATORY | BERKELEY, CALIFORNIA 94720

TWX: UC SPACE BERK
(910) 366-7945
Remote Sensing Research Program
260 Space Sciences Laboratory
University of California
Berkeley, California 94720

Mr. Joe Garcia
Code TF3
NASA-Johnson Spacecraft Center
Houston, Texas 77058

July 8, 1975

Dear Joe:

Enclosed please find a list of the U.C. Berkeley Remote Sensing Research Program's (RSRP) research and development data needs for the LACIE spectral signature extension task. It would be desirable to have as much of this data as possible for the LACIE segment sites as well as for the intensive test sites. It is expected that this data will be provided through task II when not available or duplicated by the other RSRP LACIE task's data requests or when not readily available through public sources.

Some of the items on the list would theoretically not be used in an operational system, but are necessary for the research and development of the system.

Some of the data items are already in hand as you know, but are included for a full documentation of the data requirements for this task. Also I have made notes on the availability from public sources of certain items of data so that if other researchers have similar data needs you can advise them of their ready availability through public sources.

Sincerely,

Claire M. Hay

Claire M. Hay
RSRP LACIE Project Leader

Randall Thomas

Randall W. Thomas
RSRP

RSRP LACIE Signature Extension R & D Data Needs*

1. County historical crop data
   Partly available in U.S.D.A. Crop and Livestock Reporting Service's annual crop acreage summaries by counties. Need years '71 through '75 for all test sites.
2. Crop calendar information
   Normal crop calendars for the test site areas with adjustments for the '73-'74 & '74-'75 seasons.
3. Cultural practice information for specific regions within the study:
   a) amount of irrigation practiced, what system used, location map of irrigation projects within the regions of the test sites.
   b) crop rotation practices, fallow periods, double cropping or overcropping practices specific to given crop types and as specific to geographic regions such as crop reporting districts as possible.
   c) type and prevalence of fertilizer applications by specific crop or soil conditions. Similarly for insecticide and herbiside treatments.
   d) names and characteristics of major varieties of specific crops as specific to geographic regions (crop reporting districts) as possible.
   e) significant variation in seeding rates and plant densities due to different planting practices by crop as specific to region as possible.
4. Base map photography of test sites
5. Any existing LANDSAT coincident aircraft photography with coincident or nearly coincident ground data for test sites.
6. LANDSAT color infrared, 9" X 9" transparencies to cover the entire state(s) of study. For Kansas one each within the follwing time periods:
   a) 1st week of July
   b) 1st week of August

c) mid-October (and/or mid-March)

7. Selected CCTs to cover test site locations

8. Soils data consisting of:

   a) the 7th approximation classification to the subgroup level by state.

   b) Local soils maps to the series level.(these are partly available in the U.S.D.A. S.C.S. soil survey reports for specific counties, however, not all counties have been mapped as yet or the reports are at this time not easily available to the public). These reports should contain information on the profile characteristics of the soils, the available water holding capacity of specific soils, etc.

9. Climatic data as specific to test site locations as possible. The data should include daily temperatures, daily maximum temperature, daily minimum temperatures, daily precipitation, and daily Rn (all wave net radiation input). This data is partly available through the Department of Commerce, Environmental Data Service's (Asheville, N.C.) monthly state climatalogical data reports.

10. Atmospheric data as specific to test site locations as possible. The data should include:

    a) horizontal visibility data such as reported by airport weather observation stations.

    b) skycover

    c) % possible sunshine

    d) wind vector

11. Ground data for test sites specific to individual fields consisting of:

    a) crop identification

    b) variety

    c) stand quality with specifics of any stress conditions (causes of stress)

12. Agronomic information--any publications dealing with crop behavior and management requirements by variety.

* Data requested should be provided when not available through other RSRP LACIE tasks or individual RSRP effort.