Energy, Mines and Resources Canada
Énergie, Mines et Ressources Canada
Canada Centre for Remote Sensing
Centre Canadien de Télédétection

THE USE OF LANDSAT, GROUND DATA AND A REGRESSION ESTIMATOR FOR POTATO AREA ESTIMATION

R. Ryerson
J-L Tambay
R. Plourde
J. Harvie

Users' Manual 83-2
January 1983

THE USE OF LANDSAT, GROUND DATA AND A REGRESSION ESTIMATOR FOR POTATO AREA ESTIMATION

R. Ryerson, J. Harvie
Canada Centre for Remote Sensing

J-L Tambay, R. Plourde
Statistics Canada, Ottawa

Users' Manual 83-2
January 1983

Canada Centre for Remote Sensing
Department of Energy, Mines and Resources, Ottawa

également disponible en français

Minister of Supply and Services Canada 1983
Cat. No. M77-12/83-2E
0-662-13164-9

TABLE OF CONTENTS

Abstract

This manual outlines a cost effective procedure for potato area estimation using either geometrically corrected or standard Landsat data, ground data as collected by Statistics Canada, and a regression estimator. The regression estimator relates the area of potatoes known (from ground work) to be in a segment to what the satellite estimates to be in the same segment. The estimator is then used to correct the initial satellite estimate for the whole region covered by Landsat MSS data.

Acknowledgements

A number of individuals with Statistics Canada and CCRS not named as authors have contributed their expertise to this work. These include Drs. R. J. Brown and F. Ahern of CCRS for work on method development, Ms. B. McLaughlin of Statistics Canada for designing field work procedures, and Mr. R. Dobbins of Statistics Canada and Ms. D. Robert, formerly of Intera Environmental Consultants for technical assistance. The methods are based to a certain extent on those developed in the Statistical Reporting Service, United States Department of Agriculture. Ground data used in some of the Appendices was provided by Dr. P. Mosher, formerly of the New Brunswick Department of Agriculture and Rural Development. The manuscript was critically read by W. Wigton and R. Sigman, SRS, USDA, Washington, Drs. R.J. Brown and J. Cihlar, CCRS and Mrs. N. Chinnappa, of Statistics Canada.

A Note on Authorship

This manual incorporates parts of an earlier manual which described methods which have now been improved substantially. The present authorship includes the individuals primarily responsible for method development (Ryerson and Tambay), training/technology transfer, and application and evaluation (Ryerson, Tambay, and Plourde). Much of the material provided by Harvie for the earlier publication also appears here. Apendix B on field procedures was prepared by C. Bradshaw. The senior author is now seconded to Statistics Canada and the junior author is at Simon Fraser University.

1. INTRODUCTION AND PURPOSE

The potato crop is of considerable importance in Canada[1], and in New Brunswick and Prince Edward Island it is the single most important crop under cultivation[2]. Given wide fluctuations in potato production over the years, it has been deemed necessary to develop a better procedure to estimate production for planning the handling and marketing of the crop. The purpose of this manual is to describe the procedure developed by the Canada Centre for Remote Sensing (CCRS) and Statistics Canada which, through the use of remote sensing, permits the timely estimation of potato acreage for any given region[3].

2. BACKGROUND

2.1 THE PROBLEM

The large fluctuations in potato production are among the major factors contributing to the difficulties in long-term economic and agricultural production forecasting and in developing a comprehensive marketing policy. By providing a more reliable estimate of production for a given year or given potato producing region, industry or government agricultural departments would be able to plan a more orderly handling and marketing scheme than currently exists. This will, in turn, tend to reduce fluctuations in production. Since production controls land value and equipment purchasing decisions at the farm level, stable productivity will relieve some of the economic planning problems at the farm and industry levels. This will contribute to increased consistency of employment opportunities in the secondary processing industry.

- -

[1]In 1981, total farm cash receipts for potatoes equalled $308,761,000. This made it the sixth most important crop in Canada behind wheat, barley and rapeseed, corn and tobacco. (Statistics Canada, Farm Cash Receipts, 1981, Catalogure 21-001 Monthly)

[2]In New Brunswick and Prince Edward Island, potatoes represented 72 and 75 percent of the total crops farm cash receipts for 1981, a value of $57.3 and $79.2 million respectively. (Statistics Canada, Farm Cash Receipts, 1981, Catalogue 201-001).

[3]Much of the information included in this manual has been reported in the scientific literature (Mosher, Ryerson and Strome, 1978; Ryerson, 1977; Ryerson and Dahms, 1975; Ryerson et al, 1978; Ryerson et al 1980; and Ryerson et al 1981). A previous manual did not include the statistical procedures incorporated here, although it did include a back-up method based on aerial photography.

In New Brunswick, the essential conditions which must be met to obtain a useful area estimation are that they be available in late August and that they be very accurate, with a standard deviation of the estimate under 5 to 6%. Using such data, international and domestic sales and pricing can be estimated and planned before the potatoes are harvested.

2.2 HOW IS IT SOLVED WITHOUT REMOTE SENSING?

At present, Statistics Canada uses indicators of potato area obtained from several sources to produce the published estimate in New Brunswick. One of these sources is the Objective Potato Yield Survey (OPYS), while another is the Agriculture Enumerative Survey (AES). These two surveys, although able to produce reliable potato estimates, each have their drawbacks. The OPYS is a list-based sample of farms, and thus depends on how accurate and up-to-date that list of farms is. The AES is a sample of list farms and area segments of land designed to produce estimates of various items of the provincial level. It does not produce reliable estimates of potato acreages for sub-provincial areas.

2.3 BENEFITS ACHIEVED THROUGH REMOTE SENSING

The procedure using remotely sensed data offers several advantages over the conventional approach. The image acquisition takes only a short time and hence the data are comparable and consistent. The same images can also be used alone or in combination with others to obtain information for a variety of applications such as land use, land use changes and forestry. Most important, remote sensing techniques can meet the accuracy and timeliness requirements at a low cost for even sub-provincial areas. (The example in Section 3.3 describes a method whereby remote sensing technology was used in conjunction with the AES to enable the accurate estimation of potato area in the St. John River Valley region of New Brunswick.) They also provide the possibility of reducing respondent burden, by reducing the number of farms which must be visited to obtain an adequate sample.

3. TECHNIQUES

3.1 PRINCIPLES

Remote sensing provides an interpreter with information to distinguish between surface features which, when viewed from high altitudes, have different characteristics.

The optimal method of image acquisition is dependent on both the characteristics of the sensor and the nature of the crop. To use remote sensing for identifying and measuring potato fields, it is necessary that potatoes differ in appearance from other crops and other land covers. Although there exist many methods of image acquisition and analysis, the discussion in this manual will be restricted to that one actually employed in the potato area estimation project. This project employed digital Landsat satellite imagery. Although data from other platforms[4] is also useful, their cost removed them from consideration here. The Landsat imagery is preferred because of the short time required for image acquisition and interpretation and the low cost to the user. The method was chosen on the basis of its proven capabilities from studies of other crops and its general availability in Canada.

3.1.1 Landsat Satellites

The U.S. National Aeronautics and Space Administration (NASA) Landsat IV satellite, orbits the earth in a circular, sun synchronous, near polar orbit at an altitude of approximately 705 km (438 mi.). Orbiting the earth once every 99 minutes, Landsat completes 14½ orbits of the earth per day, providing complete coverage of the earth every 16 days (Landsat Data Users Notes (23), July 1982, USGS). Coverage was provided every 18 days with Landsats I to III.

Landsat continuously receives and stores data using a Multispectral Scanner (MSS) which records light reflected from the ground in 4 spectral bands (band 4: 0.5 - 0.6 micrometers - green; band 5: 0.6 - 0.7 micrometers - red; band 6: 0.7 - 0.8 micrometers - near-infrared; band 7: 0.8 - 1.1 micrometers - near-infrared). In Canada, data received by the satellite are transmitted in real time to a ground receiving station at Prince Albert, Saskatchewan or Shoe Cove, Newfoundland, where they are recorded on magnetic tape. The data are then processed to produce standard products such as photographs and Computer Compatible Tapes (CCT's) each covering an area of 34,225 square kilometres (10,000 square miles). CCT's can be further processed through the Digital Image Correction System (DICS) (Guertin et al, 1979). DICS digital data are geometrically corrected with 50m pixels; each DICS image covers four 1:50,000 NTS topographic maps. The CCT's or DICS images can be analyzed by using digital methods, so long as the feature of interest is different from other materials in the area in at least one of the four spectral bands. If a difference does exist, that material is said to have a unique spectral signature which can be used for its identification.

The spectral reflectance characteristics of any given feature are related to its physical appearance and composition. The reflectance of a vegetation cover is a function of: the plant structure; the type of visible soil background; the surface adherents of the plant; the Void Area Index (an indication of the ratio of solid

---

[4]The definitions of all underlined words may be found in the glossary.

matter and air spaces within the leaf); the age/maturity of the crop; the Leaf Area Index (a measurement of the area occupied by the leaves relative to the soil background); and the height of the plants. These features affect the amount of light which is absorbed, transmitted and reflected by the plant and the type of soil background visible in each spectral band.

Potatoes, because of their complete ground cover and dense leaf canopy in mid-July/mid-August, are characteristically higher in near infrared reflectance than other crops in the St. John River Valley. High infrared reflectance is a general characteristic of healthy vegetation, the amount of reflectance varying according to the plant type and vigour. The existence of such a distinction must be verified for the area under study, particularly if the confusion crops differ from those in the upper St. John River Valley (Appendix A).

The techniques described in this manual will permit the reader to undertake a potato area estimation project for his own particular region. Such a project should include a pilot test to ensure smooth operation and adequate accuracy of the results. There is a possibility that due to very different terrain conditions or different confusion crops in the area under study, the accuracies for any given project may not equal those achieved here. A pilot project would be particularly important where extremely high levels of accuracy are required. In many cases existing imagery could be used to minimize data acquisition costs[5].

3.2 METHODOLOGY

Although Landsat data may be preferred it must be remembered during project planning that due to the time constraints the resource specialist must be prepared to use traditional sampling procedures on the ground if Landsat data is not available due to cloud cover, or satellite or ground processing systems failure. Aerial photography may also be used, as outlined in Ryerson et al, 1980. Since Landsat passes over a given area only once every 16 days (an eight day coverage cycle with two satellites), the inability to use the satellite data from three successive passes would jeopardize the entire program if a backup was not planned. The initial steps to be taken with respect to the planning of the program are the same for both options.

- -

[5]The likely cost of such a project given the use of existing Landsat images (which now cost $1500 for rush delivery per Digital Image Correction System (DICS) CCT) would include one man-week of field work, 4 to 8 hours of image analysis time and assorted other costs as is discussed in Section 3.3.1

3.2.1 Pre-Acquisition Phase[6]

Prior to the acquisition of the imagery, the resource specialist must first determine the geographical boundaries of the region he wishes to examine and the time at which the imagery must be acquired. This is late July to early August for potato area estimation in Canada. It must then be determined at what time the satellite path will cover the area under study and which images will cover the entire area[7].

Arrangements must then be made for the near real time acquisition of the DICS data from The Canada Centre for Remote Sensing in Prince Albert, Saskatchewan. Acceptable satellite data have been available for the St. John Valley every year but one since 1972. At the same time, arrangements to book time on the CIAS or an equivalent system must be made.

To ensure accurate analysis, the collection of ground data is crucial[8]. Ground data are essential both for the interpretation of remote sensing data and for the verification of the interpretation. Without ground data, it is very difficult to identify the crop type of individual fields or blocks of fields on a Landsat image. Information which must be collected includes field boundaries, crop types and other salient features in representative areas of the region being studied. Such information can be obtained through the analysis of existing aerial photography supplemented by field work. Ground data acquisition procedures developed and used by Statistics Canada are given in Appendix B.

Some potato fields identified at the time of ground data collection are used as training fields or sites, i.e. they serve as a basis for spectral signature generation. At least one field should be chosen for each geographically distinct area to derive the greatest possible accuracy in crop identification. In theory, training sites on Landsat data must have minimum dimensions of 3 to 4 pixels (180 to 240 m) in an east-west direction by 6 pixels (480 m) in a north-south direction to provide areas large enough for training and

---

[6] A step-by-step outline of the methodology is given in Figure 1.

[7] This information may be obtained from the User Assistance and Marketing Unit, Canada Centre for Remote Sensing, 717 Belfast Road, Ottawa, Ontario. K1A OY7. 613-995-1210.

[8] The major points of ground data collection are covered in the two following paragraphs. For the specific ground data procedures used here, see Appendix B.

to account for possible differences in the six Landsat detectors. (On DICS data corrections to the data are supposed to remove detector differences.) In practice, however, much larger fields are required than theory dictates. The training fields must be easily identifiable on both airborne and satellite imagery by means of their shape or proximity to a recognizable landmark.

For detailed field analysis (for pilot studies only) significant variations such as crop colour, crop height, crop mix, crop vigour, percent ground cover, weeds, soil colour and soil moisture, should be recorded on a sketch of the field.

3.2.2 Data Acquisition

Apart from the decision to rely on the satellite data or revert to traditional methods there is little for the user to do beyond letting the appropriate contracts. The data will be collected and processed into Landsat DICS products or CCT's, if available, and delivered to the user.

3.2.3 Analysis

3.2.3.1 Ground Segment Location

For most of the work on potato area estimation done in 1980 and 1981, segment location took place in the spring <u>before</u> the late summer analysis - although in 1982 part of the location work was done <u>after</u> the current year satellite data had been delivered. When current year satellite data are used for segment location, storage and overlay of segments on classification results becomes easier since the current year ground data are used to help locate the segments' boundaries. Ground segment location is done using for support a combination of 1:250,000 NTS topographic maps for general location, 1:50,000 NTS maps for more specific localization and aerial photographs with segment boundaries for fine detail.

The whole DICS scene is loaded as a <u>CIAS UNIDSK</u> file to disk and then it is displayed on a 512 x 512 display. From this, 9 primary 512 x 512 pixel subscenes are located and those containing agriculture are noted, (see figure 2.). These subscenes containing agriculture will co-incide with those eventually to be analyzed. In addition, 256 x 256 pixel subscenes are selected and co-ordinates recorded for later display as those containing ground segments. Subscenes used for segment location should not overlap and should be set-up as closely as possible to the four quadrants of the primary subscene. At this stage great care should be taken to ensure that the subscenes used for segment location fall within only one 512 x 512 primary subscene, except where the actual segment is split between primary subscenes.

Once a 256 x 256 subscene for segment location is displayed as 512 x 512 pixels (i.e. a two times enlargement) the Polygon Cursor Program of the CIAS is used to bound each segment. An experienced individual can bound and store a segment on old data in under fifteen to twenty minutes. Using new data this is reduced significantly since boundaries need not be drawn if no fields of the crop of interest are noted on the ground or classified on the image, since there is nothing of interest to measure. It should be noted that each ground segment in a given primary, or 512 x 512 pixel, subscene should be assigned a different theme colour of those available from three to eight. If more than five segments are located in one primary subscene, different themes should be used for those close to each other. Once all segments in one subscene are located, channel 5, which contains the theme information, should be stored as a UNIDSK file using the segment identifier of the form "XXXXXXXXX.NTS". "XXXXXXXXX" is the Statistics Canada nine digit identifier and "NTS" the 1:250,000 Map sheet identifier. For each 256 x 256 subscene used for segment location the DICS pixel co-ordinates should be recorded.

3.2.3.2 Landsat Analysis

DICS CCT's are analyzed using the CIAS or a similar digital analysis system. A primary subscene of the data (655 square kilometres, 512 pixels square) is displayed on a video screen using bands 4, 5 and 7 shown in blue, green and red respectively to simulate a colour infrared rendition. However, analysis of the data makes use of the digital values for all four bands.

Once the 512 x 512 primary subscene is displayed, the ground segment locations previously stored must be returned. To do this the appropriate UNIDSK file is identified. To avoid erasing the image display five channels are called for: the first four are zeroed, while to the last is assigned Channel 1 of the file. Since the file is 256 x 256 pixels and the original DICS scene location of the subscene is known, it is a simple (but labour intensive) task to calculate how to load the subscene so that it falls in its proper location.[9] <u>However, once the segments are all located on one primary DICS scene they only need be edited for future years</u>.

With ground segment maps showing current potato fields and their areas in hand, the next step is to select appropriately large fields within the segments for training. In the case of potatoes in the St. John Valley, this has required beginning in the area around Grand Falls and working north and west to areas which historically have fewer potatoes and inadequate training fields.

To identify potato fields with the image analysis system, large potato fields within segments are located on the image by the analyst and the analysis system must be trained to recognize them. To do this, a training field is isolated on the screen and the digital values for the four spectral bands within the site are used by the CIAS to generate the spectral signature of potatoes. The entire subscene is then examined by the system, pixel by pixel. Each pixel whose spectral signature falls within the range of values designated as "potatoes", is classified as such. When this process is completed, the pixels can be counted and the area determined, each pixel being 0.618 acres or .25 hectares in size. This process is known as the <u>supervised one-dimensional training and classification</u>. The classification and area measurement take only seconds to perform using the CIAS, although training can take considerably longer.

- -

[9]Software has recently been developed to record co-ordinates for more automatic overlay of segment data on the new satellite data. To date this step has been done manually and consequently it is both time consuming and subject to human errors in calculating overlays.

After the initial classification, the results are modified by retraining on pixels which the user identifies visually as being potatoes but which were missed in the first classification. These additional training pixels are identified from the ground data as well as from the pixels' appearance compared to classified pixels. The errors in the initial classification occur most often with boundary pixels whose spectral signatures are the result of the combined reflectances of two different crops; and within fields where minor local variations, such as an area of different soil moisture or varying topography, lead to different spectral signatures.

To perform both the original training, its evaluation and any subsequent retraining, a window or enlargement function on the CIAS is used to selectively magnify the area of any portion of the display by a factor of sixteen. This permits visual inspection of each pixel in the enlarged area to determine if it was correctly classified. If not, it may be selected for retraining. Extreme caution must be exercised in selecting pixels for retraining to avoid misclassification since confusion with other crop types might result in further misclassification.

Retraining and signature modification may be repeated if required to improve the classification accuracy. This retraining should be conducted on other fields sampled in the ground survey which have different characteristics such as different types of potatoes, different soil types, elevations, moisture content and so on to ensure classification of the maximum proportion of the potato crop.

If possible, training should be done on the centres of known potato fields to avoid the problem of accidentally retraining on a non-representative pixel. Pixels on the borders of fields which appear to be a mixture of potatoes (predominant) and other crops may also be selected on the basis of subjective judgement tempered with some training data. After classification has been performed, the histograms, which are displayed for each spectral band, should be inspected to locate any samples isolated from the majority. In so doing, any pixels representative of another land use accidentally entered into the training data may be identified and removed through "histogram chopping". By altering the limits of the histogram, the range of values comprising the spectral signature will be narrowed. Since potatoes are typically the brightest reds on the colour television display (indicating the highest near-infrared reflectance values), the histogram limits are usually modified at the lower end.

To ascertain whether the retraining and reclassification of the pixels has increased the accuracy of the classification, the area of the class potatoes can be measured within any segment. By measuring potato areain pixels where the acreage of potatoes is known, the accuracy of the classification can be determined. Despite the subjective choice of pixels for retraining, accuracies are generally increased by reclassification.

Once an acceptable classification is obtained for one subscene of an image and the area of crop in the whole subscene and in each segment is recorded, the next subscene is called back and displayed. Given the nature of the CIAS, the last classification parameters, or signature file, will automatically be displayed as the classification or alarm for the next scene. So long as the regions are similar agriculturally, the data are from the same imaging date and the same atmospheric conditions prevail, the automatic classification on the next scene will yield a suitable classification which may or may not require minor modification. In the New Brunswick work to date, it was found necessary to modify the classification beyond the original training no more than two times, for each imaging date. Where different atmospheric conditions prevail, or where there is a different date from one scene to the next, additional training is required.

The only output used for potato area estimation is generally the theme area measurement by segment and subscene since graphic output is not required. However, maps can be generated to show spatial distribution for each subscene if desired. Upon completion of the analysis, various other outputs from the CIAS could have been selected. These are: 1) a photograph of the video Landsat image; 2) a binary (single theme) map; 3) a scaled binary map; and 4) an alphanumeric line printer map at 1:50,000 or smaller showing only a sample of the pixels classified for up to 8 themes.

3.2.3.3. The Regression/Ratio Estimator

The procedure used here to generate potato area estimates is based on the assumption that the crop area measured by the satellite for the whole region will relate to the "true" crop area in the region the same way as crop areas measured by the satellite for segments (chosen at random from the region) related to the reported or true areas for those segments. The estimator method, developed and applied in a similar context first in the United States (Hanuschak et al, 1979) has proved to generate more accurate results than either satellite or survey data alone. A modification of the U.S. approach has been developed to take into account the different nature of the Canadian sample of segments. (See Appendix D for an explanation of the Canadian regression approach.)

As input into either the regression or ratio approach, the satellite based crop area estimates are compared to the segment crop areas compiled by ground enumeration. Either a simple ratio or a linear regression is used to generate a factor used to change the raw satellite estimate. (See Appredix D for a full explanation of the statistical methodology.)

3.3 EXAMPLE[10]

In this section, an example will be discussed with reference to the estimation of potato area using the method discussed. The example is an actual test which was done in the development phase of the satellite crop area estimation work at CCRS and Statistics Canada.

The aim of this project was to estimate the 1980 and 1981 potato acreage for the St. John River Valley - an area of less than 7,000 square kilometres. Ground data were acquired by Statistics Canada in the area. Field work and post analysis of high level aerial photography were used to verify the ground data collected in 1980, but such supporting data were not

- -

10This example is based on work performed in 1980, 1981 and then used operationally again in 1982. The previous manual (Ryerson et al 1980) reported on methods which used aerial photographs and Landsat data from the early-mid 1970's.

required in 1981. In 1980 the area of each potato field identified in a segment was then measured from the aerial photographs through the use of a computer driven graphic digitizer. Such detail was only required to permit proper testing and evaluation of the method over a large area. In the operational applications in 1981 and 1982, there was less intensive field work and no aerial photography was required. A sampling of a number of different fields exhibiting different topographic features, potato varieties, soil conditions, moisture conditions, etc. would be sufficient. The required sample was obtained from the Agriculture Enumerative Survey (AES).

CIAS analysis was done to develop a signature file. This file was then applied to the entire potato belt, with some modifications as we moved south subscene by subscene (512x512 Landsat pixels). Since the New Brunswick potato growing area in the St. John Valley is covered by a number of subscenes, it was not possible to simultaneously display the entire area, pixel by pixel, on the video screen. Thus, the coordinates of each subscene were recorded and a mosaic was developed.

The measurements derived using the regression approach yielded an estimate which was within two percentage points of the estimate subsequently published by Statistics Canada.

Problems of pixel misclassification or omission were primarily the result of small field sizes, of pixels on field boundaries, or of some potato fields which did not have a solid crop canopy. Additionally, confusion occasionally was found between potatoes and clover, hay/pasture, corn, weeds, or the pea/forest boundary. This confusion resulted from the similar reflectance values, particularly in the case of pixels which included harvested peas and forest. There were also some problems with hay fields which included regularly spaced stone piles. The high near IR reflectance of the stones, combined with the green colour of hay, moved the pixels to a reflectance similar to that of potatoes in the four Landsat MSS bands. The use of the regression approach is expected to handle the problem of confusion crops. For a more detailed description of the confusion crops in the St. John Valley potato belt, see Appendix C.

The time required to perform the analysis of the four DICS CCT's, which cover the New Brunswick potato belt in the upper St. John River Valley, was in the vicinity of six to eight hours exclusive of data storage (DICS CCT to disk), drawing segment boundaries, and the training of Statistics Canada staff. Location and subsequent display of the sixteen primary (512 x 512) pixel subscenes from four tapes required one hour; training and subsequent classification modification took one hour; the location of the US border and deletion of the area outside Canada was completed in fifteen minutes; while the remainder of the time was used to overlay segments on the classification results and calculate area of crop within each. With accurate pre-bounding of segments and by following the guidelines given here, an experienced individual could reduce the segment manipulation and area calculation to two to four hours. Using software recently developed, the time could be reduced to one hour for verification of the areas calculated automatically.

In practice the total cost of the New Brunswick potato mapping program will depend upon the cost of DICS data, on commercial rates for image analysis, and on the speed with which other systems could perform the tasks relative to the six to eight hours here. Image analysis would likely cost between $600 and $1200; these costs may be lower than commercial rates and are subject to change. Data costs (using DICS MSS) would be $4500 to $6000, assuming a rush turnaround basis. (Based on the CCRS price list effective December 1, 1982). Using system corrected data (on which the pre-location of segments would not be done as easily on the CIAS) the analysis costs may rise and accuracy of the estimates may decrease, while the data costs would be reduced to $2000; at the same time, data delivery would be two days after the pass instead of ten to fourteen days.

For the 1983 crop year the entire procedure has been transfered to run on a DIPIX ARIES-II[11] with standard CCT's to reduce data costs and increase timeliness. Because of the roam and zoom capabilities of the ARIES display, system corrected data can be analyzed as efficiently and accurately as DICS data were on the CIAS.

4. LIMITATIONS

4.1 CRITICAL FACTORS

If not properly considered a number of critical factors could reduce the accuracy of the potato area determination. The key factors to be considered before image acquisition are the specific time frame required and the minimum field size needed for training as discussed in Section 3.2.1. Experience has shown that a minimum of four to ten well chosen fields are required for training in the whole St. John Valley. Most critical is the need to recognize that an alternative form of data acquisition must be planned given that the present satellite acquisition system cannot gaurantee a useful image every year.

For many other applications, cloud cover during a satellite pass will render procedures invalid for data from that pass. In this case, however, methods have been developed to impute or estimate under localized cloud or cloud shadows less than several kilometres on a side. The method of imputation of potatoes

---

[11]Mention of trade names is for information and is not necessarily an endorsation by the Government of Canada, nor is exclusion of any trade name necessarily a criticism by the Government of Canada.

under cloud is quite straight-forward. It is assumed that the percent area of crop under the clouds is similar to the area of crop in an adjacent "like-appearing" area in the same region. A simple formula is used to determine potato area under cloud, P_C:

$$P_C = \frac{P_M}{T_A - C_A} \times C_A$$

where: P_M = potato area measured in the cloud-free area; C_A = area in cloud and shadow in the region and T_A = total area in the region. In the experiment described in section 3.3 these areas were incorporated into the total satellite estimate. (See Appendix D). Sample segments containing cloud were not used in the analysis.

Assumptions which must be made regarding satellite image interpretation are primarily concerned with the use of an image analysis system. Since the spectral response of a given feature can change with variations in crop maturity, soil type, soil moisture, field size and orientation, weather conditions, and cultural factors such as farming patterns and so on, the resource specialist must be well-acquainted with the geography of the area under examination to enable him to determine whether the signature file of the crop can be extended beyond a known test area. Signatures should not be extended to areas of different soil or land class type as such an extension would increase the probability of serious errors in classification. In the case of potatoes in the St. John River Valley - New Brunswick's potato belt, the decision to extend the signature was made despite the occasional confusion of potatoes with other crops. Extension was justifiable because of the relative similarity of the terrain and the sufficiently unique spectral signature of potatoes compared to the confusion crops such as corn and beans which were of relatively low acreage. If signature extension is not possible, a separate training set should be used in each sub-area which has unique properties.

Boundary pixels present another problem for interpretation. Retraining on these pixels may increase the measured accuracy substantially, but care must be taken to avoid retraining on pixels that, despite their location, are not potatoes. One method of verifying the accuracies obtained after retraining on the CIAS is the use of the Polygon Cursor option on the training areas (Section 3.2.3). These training areas should be strategically selected in representative areas of each different soil/climatic region to account for as many spectral signature differences as possible.

4.2 POSSIBLE ALTERNATE REMOTE SENSING APPROACHES

If a sampling approach is used, one could consider the use of an airborne video recorder. Such systems offer low cost, fast turnaround and the output is readily amenable to quantitive measurement.

Normal Colour aerial photography and colour infrared film (CIR) at a scale of 1:60,000 could serve as the principal alternative form of data, assuming that the high cost and difficulties of incorporating them into Statistics Canada's procedures were not limiting factors here. Normal colour film has been evaluated and reported on for this general application in a previous manual. (Ryerson, et al, 1980). Colour infrared film is sensitive to reflectance in the near-infrared portion of the spectrum and is particularly useful for highlighting vegetation which appears bright red in a CIR photograph. CIR film distinguishes very well between potatoes and hay/pasture, a source of confusion with normal colour film. This is due to the very high infrared reflectance of potatoes which renders them in a deeper red than hay/pasture on the CIR film.

At the time the original method was developed in 1975, (Ryerson et al, 1980) colour infrared film was not used because of technical problems related to consistently obtaining proper exposure. These problems are now solved (Fleming, 1978) and although colour infrared film has not been tested for this application, it might prove to be a reasonable alternative to normal colour. However, it should be noted that the cost of image acquisition using CIR film is 30-50% higher than that of normal colour film.

4.3 NON-FEASIBLE ALTERNATIVE REMOTE SENSING TECHNIQUES

Techniques which have been considered and rejected as possible alternative approaches to the problem of image acquisition for potato area estimation include black and white photography, lower altitude colour photos, and higher altitude colour photos.

In the case of black and white film, the resulting imagery does not have sufficient tonal distinction between fields except for the very largest scaled images. Low altitude colour images were easily interpreted as a result of high resolution, however they provided more detail than was necessary. Moreover, the large photographic scale reduced the overall perspective of the area in question. Hence, these images do not warrant the increased manpower and financial on cost of acquiring, handling and interpreting them. High altitude colour photos cover a very large area yielding the advantages of increased overall perspective and decreased overall cost. However, the scale is so small that the required detail is either not visible or only discernable to a very experienced interpreter.

In terms of alternate methods of analysis for potato area estimation, non-digital analysis methods are not recommended for the Landsat data, nor, due to the spectral and spatial resolution of the photographs, are any of the electro-optical devices (such as colour additive viewers and density slicers) recommended for interpretation of the Landsat photographic data. The large number of discrete gray levels which make-up the

Landsat digital image cannot be adequately distinguished by either of these methods from standard Landsat photographic products. Hence, confusion is likely to result between crops with similar canopies and infrared reflectance values, unless special purpose enhancements are developed.

5. SUMMARY

As indicated at the outset, the purpose of this manual was to introduce the technique developed for potato acreage estimation with the expectation that it could be used in the organization of individual projects. In the interest of brevity and ease of comprehension, detailed technical information which is not directly related to the functioning of the technique has been omitted. Those seeking further information on remote sensing should consult the references listed.

Although the technology is continually changing, the methods introduced in this manual are the most cost-effective techniques currently available for accurate potato acreage determination. The accuracies achieved using this program have met the standards set by Statistics Canada and New Brunswick Department of Agriculture. If these accuracies are satisfactory to the resource specialist, and the confusion crops in the region to be studied are similar to those encountered in this program, this method is a practical solution to the problem.

REFERENCES

American Society of Photogrammetry. (1968) Manual of Colour Aerial Photography. 550 p. American Society of Photogrammetry, Falls Church, VA.

American Society of Photogrammetry. (1968) Manual of Photographic Interpretation. 868 p. American Society of Photogrammetry, Washington, D.C.

Canadian Journal of Remote Sensing. 1975 - Half-yearly, Canadian Aeronautics and Space Institute, Ottawa, Ontario.

Committee on Remote Sensing for Agricultural Purposes, Agriculture Board, National Research Council (1970) Remote Sensing with Special Reference to Agriculture and Forestry. 424 p. National Academy of Sciences, Washington, D.C.

Estes, J.E. and L.W. Senger (eds) (1979) Remote Sensing: Techniques for Environmental Analysis. 339 p. J. Wiley and Sons, New York, N.Y.

Fourth Canadian Symposium on Remote Sensing Quebec, Quebec, May 16 - 18, 1977. Canadian Aeronautic and Space Institute, Ottawa, Ontario.

Fifth Canadian Symposium on Remote Sensing Victoria, B.C., August 28 - 31, 1978. Canadian Aeronautic and Space Institute, Ottawa, Ontario.

Fleming, J. (1978) Exploring the Variability of Aerochrome Infrared Film. Photogrammetric Engineering and Remote Sensing. 44(5), 601-605.

Gagnon, H. (1974) La Photo Aérienne: Son Interprétation dans les Etudes de l'Environnement et de l'Aménagement du Territoire. 278 p. Editions HRW Ltee., Anjou, Quebec.

General Electric Co. (1976) LANDSAT Data Users Handbook. LANDSAT Investigation Support, Code 902.6, NASA Goddard Spaceflight Center, Greenbelt, MD.

Girard, C.M. and M.C. Girard (1975) Applications de la Télédétection à l'Etude de la Biosphère. 186 p. Masson Co., Paris.

Goodenough, D. (1975) Image 100 Classification Methods for ERTS Scanner Data. Canadian Journal of Remote Sensing, Vol. 2, No. 1, May 1976, 18-29.

Goodenough, D. (1979) The Image Analysis System (CIAS) at the Canada Centre for Remote Sensing Canadian Journal of Remote Sensing, Vol. 5, No. 1: 3-17.

Guertin, F.E.; T.J. Butlin; and R.G. Jones (1979) La Correction Géométrique des Images Landsat au Centre Canadien de Télédétection. Canadian Journal of Remote Sensing, Vol. 5, No. 2: 118-127.

Hanuschak, G. R. Sigman, M. Craig, M. Ozga, R. Luebbe, P. Cook, D. Kleweno, and C. Miller (1979) "Crop Area Estimates with Landsat: Transition from Research and Development to Timely Results" Proc. of the Fifth Machine Processing of Remotely Sensed Data Symposium, Lafayette, Indiana, pp. 86-96.

Harper, D. (1975) Eye in the Sky. 164 p. Multiscience Publications Ltée., Montreal, Quebec.

Heffernan, J. (1979) Landsat MSS and RBV. 25 p. User Assistance and Marketing Unit, Canada Centre for Remote Sensing, Ottawa, Ontario.

Information Bulletin Airborne Operations (3rd ed.) (1977) 112 p. and Appendices, Airborne Operations Section, Data Acquisition Division, Canada Centre for Remote Sensing, Ottawa, Ontario.

The LACIE (Large Crop Inventory Experiment) Symposium - Proceedings of Plenary Session, Lyndon B. Johnson Space Center, Houston, Texas, October, 1978. National Aeronautics and Space Administration.

Mosher, P.N.; R.A. Ryerson and W.M. Strome (1978) New Brunswick Potato Area Estimation from Landsat. Twelfth International Symposium on Remote Sensing of the Environment, Manila, Philippines. pp. 1415-1419.

Photogrammetric Engineering and Remote Sensing 1934 - Monthly, American Society of Photogrammetry, Falls Church, VA.

Reeves, R.A.; (ed.) (1975) Manual of Remote Sensing. 2V. 2144p. American Society of Photogrammetry , Falls Church, VA.

Ryerson, R. A., J.L. Tambay, R.J. Brown, L.A. Murphy and B. McLaughlin (1981) " A Timely and Accurate Potato Acreage from Landsat: Results of a Demonstration" Proc. of the Fifteenth International Symposium on Remote Sensing of Environment Vol. II. 587-597.

Ryerson, R.A., P. Mosher and J. Harvie (1980) "Potato Area Estimation Using Remote Sensing Methods" CCRS Users Manual 80-2, CCRS, EMR, Ottawa, 38p.

Ryerson, R.A. (1977a) Ground Data Collection in Support of Digital Analysis of Landsat Data. Remote Sensing in Canada, No. 4, Feb. 1977, Canada Centre for Remote Sensing, Ottawa. pp. 2-4.

Ryerson, R.A. (1977) New Brunswick Potato Acreage Inventory Project. 23 p. Canada Centre for Remote Sensing, Ottawa.

Ryerson, R.A. and W.S. Dahms (1975) Remote Sensing of New Brunswick Potatoes: A Report for the New Brunswick Department of Agriculture. 17 p. Canada Centre for Remote Sensing, Ottawa.

Ryerson, R.A., P. Mosher, V.R. Wallen and N.E. Stewart (1978) Three Tests of Agricultural Remote Sensing for Crop Inventory in Eastern Canada: Results, Problems and Prospects. Proceedings of the Fifth Canadian Symposium on Remote Sensing, Victoria, B.C., pp. 441-454. (Also published in the Canadian Journal of Remote Sensing Vol. 5, No. 1, May 1979).

Ryerson, R.A. and H.A. Wood (1971) Air Photo Analysis of Beef and Dairy Farming, Photogrammetric Engineering, 37(2), 151-169.

Schepis, E.L. (1968) Time Lapse Remote Sensing in Agriculture. Photogrammetric Engineering, 3, 34(11), 1166-1179.

Sixth Canadian Symposium on Remote Sensing, Halifax, Nova Scotia, May 21-23, 1980. Canadian Aeronautics and Space Institute, Ottawa, Ontario.

Seventh Canadian Symposium on Remote Sensing, Winnipeg, Manitoba, Sept. 8-11, 1981. Canadian Aeronautics and Space Institute, Ottawa, Ontario.

Shlien, S. and D.G. Goodenough (1973) Automatic Interpretation of ERTS-A Imagery Using the Maximum Likelihood Decision Rule. CCRS Research Report 73-3, Canada Centre for Remote Sensing, Ottawa.

Statistics Canada, 1972-1975. Farm-Cash-Receipts. Catalogue 21-201, Ottawa, Canada.

Swain, P.H. and S.M. Davis (eds.) (1978) Remote Sensing - The Quantitative Approach 396 p. McGraw-Hill Inc., New York.

Symposium on Remote Sensing and Photointerpretation, Banff, Alberta, October 7-11, 1974. (International Society for Photogrammetry, Commission VII). Canadian Institute of Surveying, Ottawa, Ontario.

REFERENCES (Cont'd)

Third Canadian Symposium on Remote Sensing, Edmonton, Alberta, September 22-24, 1975. Canadian Aeronautic and Space Institute, Ottawa, Ontario.

Wendereth, S. and E. Yost (1972) Multispectral Photography for Earth Resources. 183 p. C.W. Post Center, Science Engineering Research Group, Long Island University, Greenvale, N.Y.

Wood, H.A. (1967) Accummulated Data Used in the Air Photo-Interpretation of Agricultural Land Use. Proceedings of the Second Seminar on Air Photo Interpretation in the Development of Canada, Ottawa. Published in 1968, Queen's Printer, Ottawa.

GLOSSARY

A number of the definitions included in this glossary are drawn from the Manual of Remote Sensing (Reeves, Anson and Landen, 1975) and Remote Sensing: The Quantitative Approach (Swain and Davis, 1978). The remainder is original work by members of the CCRS staff.

Alphanumeric Line Printer Map: An unscaled output option of the CIAS which shows only a sample of the pixels classified for up to eight themes. (Also see Scaled Alphanumeric Line Printer Map.)

Binary Map: An output option of the CIAS, printed by an electrostatic printer, showing each pixel of a display which has been classified as one theme.

CIAS: CCRS Image Analysis System - an interactive, digital image analysis system built and operated by the Canada Centre for Remote Sensing.

Colour Additive Viewer: An electronic viewing instrument which uses black and white multiband imagery acquired simultaneously, developed, registered and combined using different colour filters. It provides a high resolution colour composite image and by varying colour, density and hue, an interpreter can often detect subtleties undetected by the unaided human eye.

Colour Infrared Film: A type of film which is sensitive to a range of frequencies from ultraviolet through the near-infrared. In practice, the ultraviolet, violet and blue wavelengths tend to be screened out through filtration. CIR film is particularly useful for the examination of vegetation which, in this type of image, has a pink to red colour.

Computer Compatible Tapes: Digital Landsat data are available in the form of CCT's. Tapes are high speed quality .5" polyester-base magnetic tapes. One Canadian MSS CCT contains one Landsat MSS scene of 34,225 square kilometres.

Computer-Driven Graphic Digitizer: A device which reads a curve, or any irregular shape as a series of discrete points and converts them into a series of x-y coordinates. It can then compute the enclosed area.

Density Slicer: Electronic enhancement device - it assigns specific colour or monochromatic hues to photographic density intervals.

Digital Image Correction System (DICS): A system developed by CCRS which produces precision processed subscenes of Landsat MSS imagery compatible with NTS (National Topographic System) maps. The CCT's which are produced subsequent to geometric correction, rotation, resampling and framing of the image, cover an area of 0.5° latitude (58 km) by 1.0° longitude (20-85 km) with 50 metre square pixels aligned on the UTM (Universal Transverse Mercator) grid. This corresponds to 4 map sheets in the 1:50,000 series or one quarter of a map sheet in the 1:250,000 series. All products of a particular location are based on the same reference image and are automatically registered.

Dot Planimeter: A regularly/randomly spaced grid of dots which can be counted to serve as a sample of an area within a boundary of interest.

Electronic Planimeter: A manually operated device which calculates electronically the area enclosed by lines drawn on an aerial photograph or map.

Focal Length: That distance measured along the optical axis from the optical centre (rear nodal point) of the lens to the plane at which a very distant object is brought into focus.

Graphics Tablet: An option developed for the CIAS which permits overlays of map features onto an image on the CIAS video display thereby permitting the transfer of information from the map to the image.

Gray Levels: Varying shades of gray on a photograph, representing the spatial distribution of light intensity as reflected from the original camera target. Gray levels from a photo are quantified by a densitometer which measures the light attenuation of the film (at any one spot) thereby providing a light transmittance estimate, often converted to units of photographic density.

Ground Data: Supporting data collected on the ground, and information derived therefrom, as an aid to the interpretation of remotely recorded surveys, such as airborne imagery etc. Generally this should be performed concurrently with the airborne surveys. Data as to weather, soils and vegetation types and conditions are typical. The term "ground truth" is jargon for ground data.

Histogram: The graphical display of a set of data which shows the frequency of occurence (along the vertical axis) of individual measurements or values (along the horizontal axis); a frequency distribution.

Large Scale: An aerial photograph with a scale of 1:500 to 1:10,000 or a map with a scale greater than approximately 1:50,000.

Multispectral Scanner (MSS): A line scanning device which uses an oscillating mirror to continuously scan perpendicular to the flight path. In the case of Landsat MSS, 6 lines are scanned simultaneously in each of the 4 spectral bands for each mirror sweep. This information is recorded on 24 optical units. This results in continuous strip imagery which is then transformed into framed images of approximately 185 km (115 mi.)/side.

One-Dimensional Training and Classification: The examination of pixels chosen for training and the recording of the range of reflectance values in each of the 4 spectral bands. Classification is performed by examining all of the pixels and separating them into 2 groups: those whose reflectances fall within the range of the target crop for all of the spectral bands - these being classified as such; and those whose reflectances fall outside the range for all spectral bands.

Pattern: In a photo image, the regularity and characteristic placement of tones or textures. Some descriptive adjectives for patterns are regular, irregular, random, concentric, radial and rectangular.

Pixel: An acronym for "picture element" which is the minimum resolution element for which data are recorded by a multispectral scanner. In the case of Landsat 1, 2, and 3, the unique area of an MSS pixel due to overlap is approximately 60 x 80 m. or 1.1127 ac.

Platform: A sensor carrying device such as a satellite, aircraft or balloon.

Polygon Cursor Program: An option on the CIAS whereby areas of any shape may be enclosed and then used in the classification and subsequent measurement of classes only in the selected polygon.

Real Time: Near instantaneous operation. This term can be used in two contexts: 1)Real time data acquisition - data are collected by the satellite and are then transmitted and processed into a CCT immediately; 2) Real time classification modification - an attribute of the CIAS whereby modifications made to the classifications displayed on the screen are performed and displayed instantaneously.

Recognition Element: The smallest object which can be recognized by photointerpretation.

Resolution: The ability of an entire remote sensor system to render a sharply defined image.

Scale: The ratio of a distance on a photograph or map to its corresponding distance on the ground. The scale of photograph varies from point to point because of displacements caused by tilt and relief, but is usually taken as f/H where f is the principal distance (focal length) of the camera and H is the height of the camera above mean ground elevation. Scale may be expressed as a "unit equivalent" (e.g. 1 mm = 25 m), a "representative fraction" (e.g. 1/25,000) or a "ratio" (e.g. 1:25,000).

Scaled Alphanumeric Line Printer Map: An output option of the CIAS which shows only a sample of the pixels classified for up to eight themes at scales of 1:50,000 or larger.

Signature File: A collection of the reflectance values in the four spectral bands used by Landsat which are used in pixel identification by the CIAS.

Small Scale: An aerial photograph with a scale smaller than 1:40,000 or a map with a scale less than 1:1,000,000.

Spectral Signature: Quantitative measurement of the reflectance of an object at one or several wavelength intervals.

Sun Synchronous: An earth satellite orbit in which the orbital plane is near polar and the altitude such that each satellite ground trace repeats its earth coverage at the same local sun time.

Supervised Classification: A form of digital image analysis whereby the spectral signature of a known feature is examined and a thematic map is created by identifying each pixel having the same spectral signature or a similar spectral signature to that of the known object..

Texture: In a photo image, the frequency of change and arrangement of tones. Some descriptive adjectives for textures are fine, medium or coarse, and stippled or mottled.

Theme: Individual classes of spectral signatures which can be stored on the CIAS and be displayed by colour codes. Up to 8 classes can be displayed at the same time to yield a thematic map of classes overlain on the Landsat data display.

Theme Area Measurement: An output on the CIAS whereby the area belonging to each theme is printed out in table form.

Tone: Different shades of gray visible in photographic images, usually as a result of differences in reflected energy.

Training Fields: Sites of a known land use chosen from ground data which are used to recognize and classify other pixels of the same land use through the use of spectral signatures.

Unsupervised Classification: A form of digital image analysis whereby pixels having similar spectral signatures are grouped together and identified as being the same although what they actually represent is unknown.

Window: A small section of the Landsat image displayed on the CIAS video screen, can be isolated and magnified, to permit visual examination of individual pixels.

APPENDIX A

Crop Areas and Field Sizes 1975[1]

	Total Area in Hectares (acres)*	Number of Fields	Average Field Size Hectares (acres)
Potatoes	3559 (8794)	405	8.8 (22.2)
Peas	452 (1117)	27	16.7 (41.4)
Bare	186 (460)	33	5.6 (13.9)
Hay	915 (2260)	180	5.1 (12.6)
Grain	868 (2145)	154	5.6 (13.9)
Pasture	289 (715)	52	5.6 (13.8)
Broccoli	174 (430)	20	8.7 (21.5)
Weeds	408 (1007)	124	3.3 (8.1)
Brussel Sprouts	2 (6)	3	.8 (2.0)
Corn	17 (43)	3	5.8 (14.3)
Buckwheat	40 (100)	12	3.4 (8.3)

*rounded to the nearest acre

NOTE: Potato area calculated using a digitizer, all other areas calculated using a dot planimeter. The original measurements were in acres.

[1]Ryerson, 1977.

Appendix B

Statistics Canada

Agricultural Enumerative
Survey Manual

Remote Sensing Potato Project
Interviewers Instructions For
Madawaska, Victoria and Carleton
Counties

New Brunswick

C. Bradshaw
Special Projects
Regional Operations Division
Statistics Canada

1. PURPOSE OF THE PROJECT

Since 1980, Interviewers in some areas of New Brunswick have been participants in a project carried out between Statistics Canada, and the Canada Centre for Remote Sensing. The project is part of the development of a faster, more efficient method of arriving at potato acreage estimates by the use of ground "reference" data supplied by Interviewers in the Agriculture Enumerative Survey (AES) and remote sensing data in the form of satellite photos.

In 1981, the project involved the three major potato producing counties of New Brunswick, which include Victoria, Madawaska, and Carleton. AES Interviewers who worked in these three counties last year were asked to outline fields of potatoes and record their acreages on special enlargements of the aerial photograph of the enumerated segments.

The enlarged aerial photos were used to "train" a computer to recognize potato areas using a sample of known potato fields. Then data obtained from the Landsat satellite, orbiting 565 miles (909 Kilometers) above the earth, were analyzed by the computer. By employing both the Landsat satellite data and a sample of known mapped potato fields (provided by the AES Interviewer), estimates of potato areas were created in a matter of hours for the whole St. John Valley to produce an estimate of the total potato area for the three counties.

As a result of the success of the project, the program will be carried out again this year in areas of New Brunswick, as well as in Prince Edward Island and parts of Ontario.

2. THE 1982 REMOTE SENSING PROJECT IN NEW BRUNSWICK

The 1982 Remote Sensing project in New Brunswick involves the same three counties. This year, however, not all segments in the area are involved in the plotting of fields for remote sensing. A total of 28 segments have been selected. The label on the segment folder for the 28 segments will have the word "REMOTE", printed under the segment identification.

In each of these "REMOTE" segment folders, you will find an enlarged copy of the aerial photo along with the other aerial photo and maps. In your survey supplies, you will be provided with two special markers for outlining the areas on the enlarged aerial photo.

The plotting of fields on the enlarged aerial photo for remote sensing segments will be done as you are enumerating the segment for the AES. Complete the procedures for remote sensing after your have completed the aerial photo and VCR procedure and have determined that the operator has an agricultural holding.

One of the difficulties of the Remote Sensing approach is that the satellite "confuses" potatoes with other crops such as field beans and corn. We refer to these crops as "confusion crops". As a result of this confusion, the satellite must be "trained" to identify these crops separately. To do this, we must know the exact location of potato fields and the "confusion crop" fields.

You must outline all potato, field bean and corn fields within the Remote Sensing segments on the enlarged aerial photo. In the case of potato fields, you must also note the size of the potato area in acres.

3. NEW PROCEDURE FOR SEGMENTS NOT DESIGNATED FOR REMOTE SENSING

For all other segments in the counties of Madawaska, Victoria and Carleton which were not designated for the Remote Sensing Project, you will find a label on the segment aerial photo.

Example of label on aerial photo:

CROSS/REF TO SPEC FARMS ONLY*

VCR NO.	POTATO AREA IN SEGMENT	VCR NO.	POTATO AREA IN SEGMENT
----	----------	---	----------
----	----------	---	----------

SPECIFY LAND UNITS IF NOT IN AREA

If a specified farm has land in one of these segments, you must determine the area of potatoes within the segment and record this on the label. In this case, although a segment farm questionnaire is not completed, we still require the area of potatoes within the segment from the specified farm operator. (More details about this procedure are provided in Section 5).

4. FIELD PROCEDURES FOR REMOTE SENSING SEGMENTS

Before you begin to enumerate a segment designated for Remote Sensing, familiarize yourself with the segment boundaries and the key crop fields within the segment so that you may help the farm operators outline their potato, field bean and corn fields, if necessary.

When you contact the farm operator, complete the aerial photo and VCR procedures. Then, outline the farm operator's holding on the enlarged aerial photo with the special green marker and enter the VCR number within the outlined holding. This is the same procedure you follow for the smaller aerial photo.

If the holding is non-agricultural, mark "N/A" inside the boundaries of the holding as you did on the smaller aerial photo. Thank the respondent and leave.

If the holding is agricultural, proceed as follows:

A. Ask the farm operator if he has any potato, field bean or corn fields within the segment. YOU MUST ACCOUNT FOR ALL OF THESE CROPS WITHIN EACH HOLDING ON THE ENLARGED AERIAL PHOTO.

(If the respondent has none of these crops, proceed to complete the AES questionnaire.)

Outline each potato, field bean and corn field using the special red marker. Mark "P" inside the areas representing potato fields, "FB" for field beans, and "C" for corn fields. (A label with all field codes is attached to the enlarged aerial photo.) Outline these fields as accurately as possible.

In some cases, fields may not be visible on the enlarged aerial photo because the photo was taken in 1976. Try to outline the exact location of the fields on the aerial photo even though some of the landmarks have changed. For example, a wooded area on the aerial photo may have been cleared and is now used for growing crops.

B. For each potato field, ask the farm operator to estimate the size of the field, and record the acreage beside the "P" on the enlarged aerial photo.

C. Check that the acreage given by the respondent for each field is consistent with that given for other fields. For example, does a field given as forty acres seem twice as large as a field given as twenty acres?

A square, representing 10 acres has been placed on a corner of the enlarged aerial photo. Use the 10 acre square as a guide to judge the relative field sizes. If the field sizes do not appear consistent, check:

- that the farmer has assigned the acreage to the correct field;

- that the farmer is referring to only the part of the field inside the segment;

- that the field has not been enlarged or altered since the photo was taken;

- that the farmer is not referring to a group of fields.

Field boundaries should be adjusted with the red marker to account for any discrepancies.

D. Outline only the potato, field bean and corn fields which are one acre or more in size. FIELDS LESS THAN ONE ACRE ARE NOT TO BE MARKED.

E. If the farm operator has potato fields on his holding, record the following information in the box on the right hand corner of the enlarged aerial photo:

- In column (1), enter the VCR number of the holding;
- In column (2), record the number of potato fields mapped on the farm operators holding;

VCR NO. (1)	# OF POTATO FIELDS (2)	TOTAL ACRES OF POTATOES MAPPED IN THE SEGMENT (3)	VERIFY (2) & (3) WITH RESPONDENT (4)
002	1		
SPECFIY LAND UNITS IF NOT IN ACRES			

- Add the farm operator's potato field acreages that were recorded on the enlarged aerial photo, and enter the total potato acreage in column (3)

- Column (4) asks you to verify with the respondent, the number of potato fields and the total potato acreage within the segment. If these are not correct because several fields of less than one acre were not outlined, provide a brief comment. For additional comments, use the back of the enlarged aerial photo. Always identify the holding by referring to the VCR number.

VCR NO. (1)	# OF POTATO FIELDS (2)	TOTAL ACRES OF POTATOES MAPPED IN THE SEGMENT (3)	VERIFY (2) & (3) WITH RESPONDENT (4)
002	1	8	
003	2	37	
004	2	9	½ acre seed plot not mapped
SPECFIY LAND UNITS IF NOT IN ACRES			

NOTE: If the respondent does not have potato fields, there will be no entry in this box.

F. Comment on the back of the enlarged aerial photo if you encounter any ususual circumstances, for example, diseased potato fields, vegetable interplanting, flooded fields, etc. Refer to the VCR number in all cases.

G. After completing the Remote Sensing procedures proceed to complete the remaining AES procedures.

5. SPECIAL PROCEDURES

A. Refusals and No Contacts

If you encounter any land inside a Remote Sensing segment which belongs to a respondent who refuses to participate in the survey or a respondent who you are unable to contact, mark the boundaries of their holdings on the enlarged aerial photo. Place the intials "R" (Refusal) or "NC" (No Contact) as appropriate, inside the boundaries. Only if you are positive that there are no potatoes in any of the "R" or "NC" fields in the segment, indicate this by a note on the back of the enlarged aerial photo. In these cases, make sure you identify the land by referring to the VCR number.

B. ZZZZ Farms and YYYY Farms (Two Large Holdings treated separately and not identified here for reasons of confidentiality). ZZZZZ and YYYY are excluded from our survey because they are being enumerated by another method. If you should encounter any land inside the segment farmed by YYYY or ZZZZ, mark the boundary of their holdings and place an "X" inside the boundaries.

C. Segment Farms on the Segment Screening List

If a segment farm in a remote sensing segment is on the Segment Screening List, you will complete the cover page, page 2 and the screening questionnaire on page 17. You must also complete the enlarged aerial photo and map the fields and potato acreages for the holding.

D. Cross Reference Situations in Remote Sensing Segments - Segment Farms, List Farms and Specified Farms.

The AES procedures for a cross reference in a segment to a farm in another segment or to a list farm state that you are not required to complete a full AES segment questionnaire the second time the farmer is contacted. In the case of land operated by a specified farmer found in a segment, only the specified farm questionnaire is completed.

If any of these cross reference situation occurs in one of your remote sensing segments, you are still required to complete the enlarged aerial photo, map the fields and potato acreages as outlined in these procedures. IT IS VERY IMPORTANT THAT ALL ENLARGED AERIAL PHOTOS IN THE REMOTE SENSING SEGMENTS ARE COMPLETED AS FULLY AND AS ACCURATELY AS POSSIBLE.

E. Cross Reference Situation for Specified Farms in Segments which do not have Remote Sensing.

In all of the segments in Madawaska, Carleton and Victoria counties which are not selected for the Remote Sensing procedures, you will find a label on the front of the aerial photo. A line on the label is to be completed for every specified farm with land inside the segment. This will ensure that we obtain potato areas inside the segment for all land inside the segment.

Enter the VCR number, and Potato area in the segment for every specified farm with land inside the segment.

If there is no land belonging to a specified farm inside your segment, place a note, "No Specified farms" on the label.

CROSS/REF TO SPEC FARMS ONLY

VCR NO.	POTATO AREA IN SEGMENT	VCR NO.	POTATO AREA IN SEGMENT
003	37		

SPECFIY LAND UNITS IF NOT IN ACRES

6. FINAL CHECK

Since the material you provide is the only source of data available to train the computer to map potatoes from the satellite, and since the acreages are critical to providing accurate numbers from the satellite, one final verification is necessary. Where there are no potatoes, ignore this step.

The final varification consists of two phases:

a) Verify the relative size of fields marked as potatoes. Do the sizes of the fields and acreages appear consistent? Note any inconsistencies on the back of the enlarged aerial photo.(Refer to VCR numbers in all cases.)

b) Check that potato fields have been identified on the enlarged aerial photo for all your remote sensing segments.

7. INFORMATION ON REMOTE SENSING

The use of Remote Sensing methods would enable Statistics Canada to produce accurate, timely data on crop acreages that would be very cost effective and result in reduced response burden for the respondent.

If any of your respondents have any questions on remote sensing or wish to get further information, instruct them to contact CCRS directly at the following address:

User Assistance Unit
Canada Centre for Remote Sensing
Canada Department of Energy, Mines and Resources
717, Belfast Road
Ottawa, Ontario
KlA OY7 (613) 995-1210

Note: On the original interviewers' instructions there were several illustrative items: (1) flow chart of information; (2) example of a segment photography, (3) a cartoon review of key points prepared by R. Dobbins, Remote Sensing Unit, Crops Section, Statistics Canada.

APPENDIX C

Confusion Crops

As indicated in the text, errors in classification do result both in visual and in digital interpretation. These errors may be errors of commission - in which another crop (a confusion crop) is identified as being a member of the target crop, or errors of omission - in which a member of the target crop is classified as being something else. In this section, the confusion crops will be further examined for both the visual and the digital forms of analysis. The discussion will be drawn from the results obtained in certain test projects in New Brunswick.

It can be seen from examining the results in Table C-1 for the Landsat/CIAS analysis, that of a total potato acreage of 970.9, 35 acres of other crops were incorrectly identified as potatoes, while 82.9 acres of potatoes were omitted entirely. Of these errors which did occur, most occurred as a result of the location of the pixels in small fields, on field boundaries or in fields which did not have a solid canopy. For example, 33 acres of the 82.9 acres missed were in eight fields of less than six acres each. Other crops which were present were broccoli, grain and grass as well as bare ground. None of these provided any difficulty in classification. Thus, the predominant problem in digital analysis appears to be size-related.

Whereas the major sources of confusion in digital analysis were peas/forest and hay, hay/pasture provided the most difficulty in visual interpretation. (See Table C-2.) As noted previously, this confusion is the result of similar closed crop canopies and similar colours. Consequently, 25% of all hay/pasture was identified as potatoes, while 10.8% of all potatoes were incorrectly called hay/pasture. Table C-2 provides additional detail on other confusion crops.

Table C-1

Results of Digital Image Analysis

	True Class				
	Potatoes	Hay	Corn	Weeds	Peas/Forest
Chosen Class					
Potatoes	888 (91)[1]	14(10)	3(100)	7(16)	11 (N/A)
Omitted Potatoes	82.9(9)				

[1]Acres (percent)

Table C-2

Confusion Matrix for Visual Air Photo Analysis and Classification

	True Class				
	Potatoes	Grain	Peas	Broccoli	Hay/Pasture
	(all values in Percent)				
Chosen Class					
Potatoes	80.7	6.8	0.0	0.0	25.0
Grain	2.3	66.0	0.0	0.0	3.1
Peas	1.8	2.2	57.9	0.0	0.0
Broccoli	4.4	4.5	42.1	87.5	6.2
Hay/Pasture	10.8	20.5	0.0	12.5	65.6
	100.0	100.0	100.0	100.0	100.0[1]

[1]Actual total of 99.9 because of rounding.

APPENDIX D

The Regression Approach

1. The Sample Design

The regression estimators used for potatoes in 1980 and 1981 were based on the design of the Agriculture Enumerative Survey (AES) as it provided the segments for which the satellite and ground data were collected. AES segments in New Brunswick were selected following a stratified two-stage replicated sample. Census-based Enumeration Areas (EA) were stratified according to certain characteristics, and two independent samples (replicates) of EAs were drawn from each stratum. This was the first-stage sample. Each selected EA was then subdivided into roughly equal areas called segments, and a sample was chosen from those segments. This was the second stage of the sample.

Being a multi-purpose survey, the AES was not well suited for the test. The main problem was that strata did not follow geographical regions. The number of segments available in each stratum depended on how many AES segments happened to be chosen within the region of interest. For test purposes, segments sampled outside the study area were ignored, and those inside the test region were treated as if they had been specifically sampled from within the test region of each stratum. The assumption was also made that all the EAs selected in a stratum were sampled independently, as it was not practical to work within the replicates. The design used for the estimation thus became a stratified two-stage sample drawn from within the test region.

In 1980, segments with missing data were deleted from the sample. This was also done for one outlier. Missing data were manually imputed for in 1981.

2. The Design Estimate

Let y_{hij} denote the potato area in the j^{th} selected segment of the i^{th} EA sampled in stratum h. The design estimator of Y, the total potato area in the test region, is

$$\hat{Y} = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{m_{hi}} y_{hij} \qquad (1)$$

where N_h = total number of EAs in stratum h (in the test region), h = 1, 2, ..., L;

n_h = sampled number of EAs in stratum h (in the test region);

M_{hi} = total number of segments in the i^{th} sampled EA of stratum h; and,

m_{hi} = sampled number of segments in the i^{th} sampled EA of stratum h.

The variance estimate of Y is

$$v(\hat{Y}) = \sum_{h=1}^{L} \frac{N_h^2}{n_h(n_h-1)} \sum_{i=1}^{n_h} (\hat{Y}_{hi} - \hat{\bar{Y}}_h)^2 , \qquad (2)$$

where $\hat{Y}_{hi} = \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{m_{hi}} y_{hij}$; and,

$$\hat{\bar{Y}}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} \hat{Y}_{hi} .$$

The regression estimator also requires using the same design to calculate an estimate of the (known) total potato area in the test region, as determined by the satellite classification. The known area is denoted by X, while its estimate is denoted by $\hat{X}$. If x_{hij} denotes the satellite classification potato area for the same segment as y_{hij}, the total satellite classification potato area in the test region is estimated by substituting x_{hij} for y_{hij} in (1). Thus

$$\hat{X} = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{m_{hi}} x_{hij}. \qquad (3)$$

The variance estimate of X is also obtained by substituting x_{hij} for y_{hij} in (2).

3. The Regression Estimator

Given that the value of X is known, the regression estimator used in 1980 was

$$\hat{Y}_{Reg.} = \hat{Y} + \hat{B}\ (X - \hat{X}), \qquad (4)$$

$$\text{where } \hat{B} = cov(\hat{X},\hat{Y}) \ / \ v(\hat{X}) \text{; and,} \qquad (5)$$

$$cov(\hat{X},\hat{Y}) = \sum_{h=1}^{L} \frac{N_h^2}{n_h(n_h-1)} \sum_{i=1}^{n_h} \left(\hat{X}_{hi} - \hat{\bar{X}}_h\right)\left(\hat{Y}_{hi} - \hat{\bar{Y}}_h\right). \qquad (6)$$

The variance estimate of $\hat{Y}_{Reg.}$ was

$$v(\hat{Y}_{Reg.}) = v(\hat{Y}) - \frac{[cov(\hat{X},\hat{Y})]^2}{v(\hat{X})} \qquad (7)$$

In 1981, the total potato area of a few large operations was obtained separately[1]. Ground data were not to be collected for any segment land belonging to these operations. The regression estimator thus became

$$\hat{Y}_{Reg.} = \hat{Y}' + Y_o + \hat{B}' (X' - \hat{X}'), \tag{8}$$

where Y_o = the total potato area for these large operations, $\hat{Y}'$, $\hat{B}'$, $\hat{X}'$, and the variances and covariance were calculated as in 1980 except for the fact that segment data excluded any land belonging to these large operations. The known total satellite potato area in the test region was $X = X' + X_o$, where X_o denotes the total satellite potato area for the large operations. To obtain a value for X', the value x_{ohij} (the area of potatoes in the land belonging to the large operations, as given by satellite classification) was obtained in each sample segment. Using the design estimate (1), $\hat{X}_o$, an estimate of X_o, was calculated. Assuming

$$\frac{\hat{X}'}{X'} = \frac{\hat{X}' + \hat{X}_o}{X' + X_o} = \frac{\hat{X}' + \hat{X}_o}{X} \tag{9}$$

the value X' was obtained as

$$X' = \frac{\hat{X}'(X' + X_o)}{(\hat{X}' + \hat{X}_o)} = \frac{\hat{X}' \; X}{(\hat{X}' + \hat{X}_o)} \tag{10}$$

Finally, the variance of (8) was calculated as in 1980, excluding from segment data all land belonging to the large operations.

4. The Ratio Estimator

The ratio estimator used in 1980 was

$$\hat{Y}_{Ratio} = \frac{\hat{Y}}{\hat{X}} X, \tag{11}$$

with variance estimate

$$v(\hat{Y}_{Ratio}) = v(\hat{Y}) - 2 \frac{\hat{Y}}{\hat{X}} cov((\hat{X},\hat{Y}) + \frac{\hat{Y}^2}{\hat{X}^2} v(\hat{X}). \tag{12}$$

In 1980, the estimator was

$$\hat{Y}_{Ratio} = Y_o + \frac{\hat{Y}'}{\hat{X}'} X' , \tag{13}$$

[1]see Appendix B

using the same notation as before. The variance of the estimator was calculated as in (12), excluding from segment data all land belonging to the large operations.

APPENDIX E

Results from 1980, 1981 and 1982

The 1980, 1981 and 1982 estimates for the potato acreage in the potato belt of New Brunswick (the test region) were obtained using the sample design of the Agriculture Enumerative Survey. In 1980, 42 segments were used for the regression and ratio estimates. In 1981, 62 segments were used. In 1982, 59 segments were used. (see Appendix B.) The ground areas for the 1980 estimates were derived from high altitude aerial photographs of the sample segments. AES enumerators supplied the ground data in 1981 and 1982.

Table E-1 gives the satellite estimates of potatoes for the test region and for all of New Brunswick. The New Brunswick estimates are obtained by expanding the test region estimates by a constant. The Statistics Canada published estimates for the province are also given. Of interest is the fact that, although very similar in 1980, the regression and the ratio estimates were quite different in 1981 and 1982. In all three years, the regression estimate was within about two percent of the published estimate.

TABLE E-1

1980, 1981 and 1982 Estimates of the Potato Area in New Brunswick

	1980 Estimates (Acres)			1981 Estimates (Acres)			1982 Estimates (Acres)		
	Test Region	New Brunswick	C.V. (%)	Test Region	New Brunswick	C.V. (%)	Test Region	New Brunswick	C.V (%)
Ratio Estimate	49,504	51,524	5.5	55,628	57,898	5.4	55,264	57,964	5.2
Regression Estimate	49,115	51,119	5.5	51,717	53,827	5.5	52,533	55,233	5.2
Statistics Canada: Published Estimate	-	52,000	-	-	53,000[1]	-	-	54,000	-
1981 Census of Agriculture	-	-	-	-	53,793	-	-	-	-

Note: The coefficient of variation (c.v.) of an estimate is calculated as:

$$C.V.\% = \frac{\text{Variance of Estimate}}{\text{Estimate}} \times 100\%$$

[1] Revised after 1981 Census to 54,000.

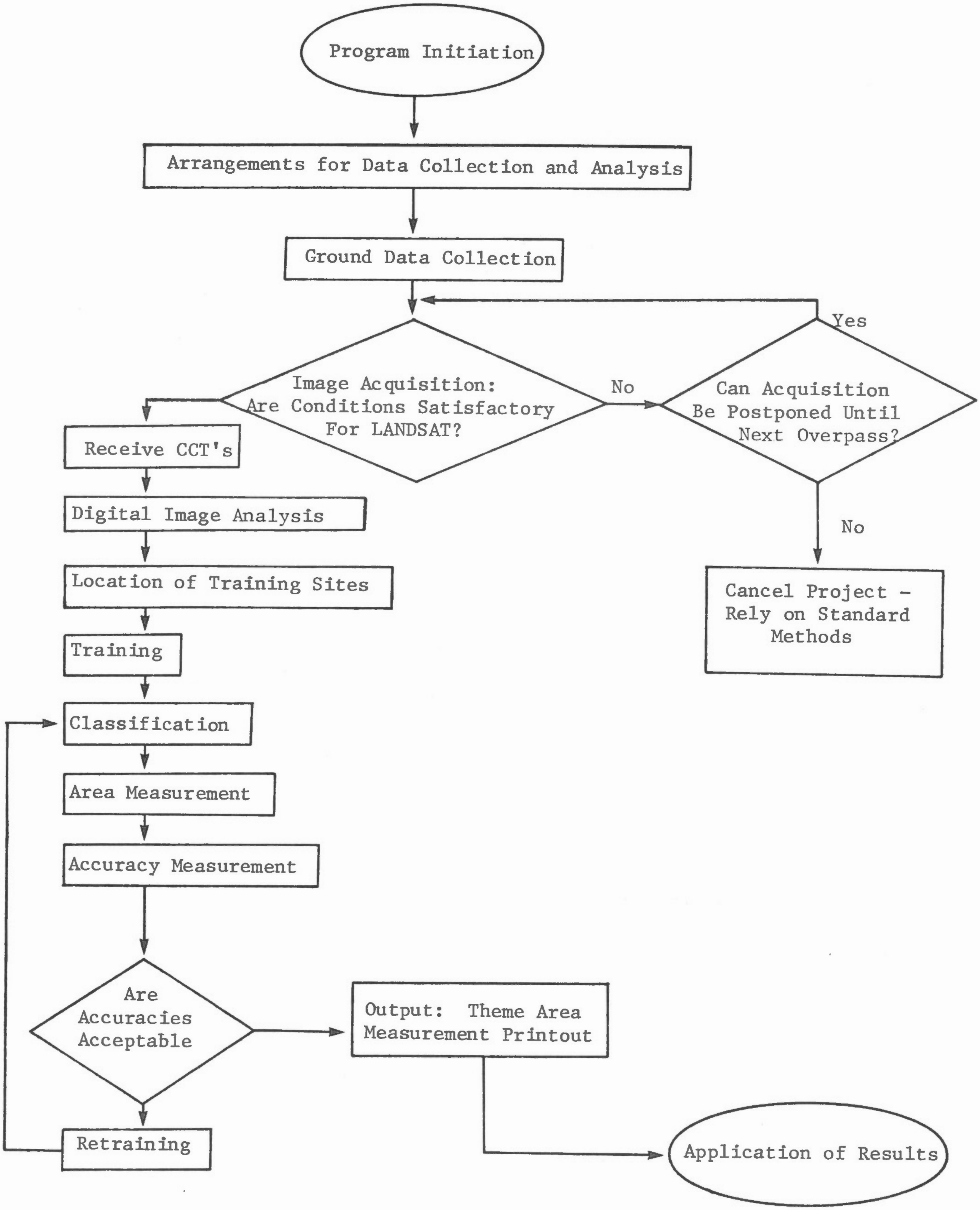

Figure 1: Potato Area Estimation - Procedural Flowchart