1981 SIGMAN

Presented at NASA Workshop on Registration and Rectification on November 17, 1981

81-15

USDA REGISTRATION AND RECTIFICATION REQUIREMENTS

Rich Allen
Statistical Reporting Service
U.S. Department of Agriculture

I am glad to have this opportunity to discuss some of the requirements of the United States Department of Agriculture (USDA) for accuracy of aerospace acquired data, and specifically, requirements for registration and rectification of remotely-sensed data acquired by space vehicles, such as Landsat, the Shuttle, and so on. The views presented are for the most part my observations and opinions from work completed to date or underway in the Department of Agriculture since only limited documentation exists on the accuracy required for registration and rectification as such.

Definition of Terms

It is perhaps wise to first explain the terms to be used in the presentation. One of the organizers of this conference indicated that registration refers to the accuracy of location of identical points on repeated acquisitions of data while rectification refers to accuracy of matching remotely-sensed data with corresponding points on the ground. Many people in the USDA use the expression "registration" for matching both scene-to-scene and scene-to-map and use the term "rectification" for indicating accuracy of maps or aerial photographs. I will try to utilize the announced definitions of this conference, although I may often use the broader term of positional accuracy to refer to the USDA requirements.

Remotely-Sensed Data: An Important Source of Information

In order to carry out its assigned missions, the Department of Agriculture must have information relating to observations of soils, crops, and other physical features. For example, we identify and map soil types, monitor conservation practices and conservation structures (construction of new conservation measures), and verify farmers' compliance with planting restrictions under various farm programs. All of these types of information programs have intense, detailed data needs. Aerial photography has been utilized to provide much of the needed data for these programs. An aerospace sensor is not now foreseen as being capable of providing comparable data required in such domestic programs. I emphasize the term "domestic". However, we can characterize the positional accuracy needs of these information programs, and, in so doing, may provide general guidance for aerospace sensor development programs.

If an aerospace sensor could be developed, budgets permitting, to satisfy some of the needs above, that sensor would have to have spatial resolution similar to that of aerial photography to be judged useful to USDA or, better put, to be competitive with aerial photography. By the same analogy, the USDA would need the same rectification as with aerial photography to utilize that spatial resolution.

Satellite Data Uniquely Suited for Specific Programs

However, in this paper I want to concentrate on four other types of information programs, ones for which satellite data are being routinely used or at least currently investigated for possible use later. In these, because of the unique features of the requirements, satellite data are "competitive" with other information sources. These four information programs are:

Foreign crop forecasting
Domestic crop acreage estimation
Forestry information applications
Rangeland condition evaluations

I believe there are similarities and contrasts among these four programs which will illustrate the range of USDA requirements for positional accuracy. These examples are not considered an exhaustive list of USDA's present interests for satellite data.

Landsat and Metsat Data Used for Foreign Crop Condition Assessment

The foreign crop assessment approach that I wish to discuss is that of the Foreign Agriculture Service's Foreign Commodity Condition Assessment Division (FCCAD). FCCAD is processing large numbers of satellite scenes from both Landsat and meterological satellites in its current efforts to monitor areal extent of anomolies and production of major crops in selected foreign areas. For the most part, FCCAD does not have access to ground collected data. Thus, it depends on sequential reviews of satellite data as a procedure for monitoring and evaluating change.

Presently FCCAD does not perform any registration on the data received from Landsat in the form of High Density Computer Compatible Tapes. Each scene is reduced to subsets or samples of full resolution data by subsampling rows and pixels to obtain a data set better suited to FCCAD's processing capabilities. This reduced data set can then be displayed on image processing equipment as raw data or by transformation to vegetative indexes, which measure the "greenness" of plants, and are indicators of plant vigor and stress. The Vegetative Index approach is an effort to standardize responses and allow meaningful comparisons within seasons and across seasons.

Much of the present analysis of FCCAD involves calculation of average values for a 25-mile square grid. Without registration procedures there may be some shifting of points included in a particular grid cell from acquisition to acquisition but the shifting should be minor compared to the size of the grid cells.

The approach of not performing any additional registration of satellite acquired data is not an optimum one in FCCAD's point of view but, in light of equipment and personnel resources and the volume of data being handled, it is not feasible to devote time to registration activities at this time. If each satellite scene obtained had consistently "high" positional accuracy to the ground and thus to corresponding scenes of other data, FCCAD would be interested in a major refinement of its multitemporal analysis procedure. A much better measure of changes in conditions or crop area could be obtained if the data for specific sampling segments could be matched throughout a season and across seasons. Based on the spatial resolution of the current Landsat Multispectral Scanner (MSS) this might mean use of a segment of approximately 15-20 square miles in size. All data, including vegetative indices or other transformations, would be calculated and stored for these segments.

Based on the size of the segment visualized it would be reasonable to employ this sampling approach if satellite data were rectified or registered to a mean accuracy of plus or minus one pixel. This error in location from acquisition to acquisition would be small enough compared to the size of the segment to be effective.

If a sensor with finer spatial resolution was available, the FCCAD's accuracy requirement would likely remain at one pixel. Its approach might become one of utilizing somewhat smaller segment sizes but with data registered within one pixel to maintain the same relative accuracy.

Use of Landsat Data for Domestic Crop Acreage Assessment

Research into the use of Landsat MSS data by the Statistical Reporting Service (SRS) for the improvement of acreage estimates of major domestic crops provides a marked contrast to the procedures for foreign areas. Unlike the Foreign Agricultural Service which often has no ground data, SRS has available a sample of ground observation data which by itself provides an accurate estimate of acreage.

One of the major inputs to crop acreage estimates in the United States is the June Enumerative Survey (JES) conducted each year by SRS in the 48 coterminous states. The JES is a probability area frame survey in which each state has been divided into land use strata based on percentage of cultivation and type of land use. Each stratum is divided into sampling units, with size depending on the stratum. In the Midwest, intensive agricultural sampling units or segments are typically one square mile in size. Segment sizes in rangeland strata are much larger and sizes in residential or commercial strata are much smaller.

The JES survey provides estimates of major crop acreages with coefficients of variation ranging from 2.5 percent to 6.0 percent in major states. At the U.S. level these coefficients of variation are as small as 1.0 - 2.0 percent.

Thus, the SRS research effort is an attempt to improve the precision for relatively "good" acreage estimates. This is in direct contrast to the foreign crop forecasting problem. Therefore, the demands for positional accuracy are much higher.

SRS uses the JES data for training of classification algorithms, for testing of classification results, and for estimation of acreage through a regression estimator. The JES data is closely edited on a field-by-field basis. Random fields of each cover of interest are selected for training. The SRS approach extracts interior pixels of training fields for labelling so it is important to ensure accuracy of field locations. The training fields for a scene are selected from JES segments located throughout the scene. Typically, about 30 segments are available in a Landsat scene.

Once an analyst is satisfied with the clustering relationships for a scene, each segment in the scene and the entire scene (within boundaries of geographic counties and excluding cloud covered areas) is classified. The classification of segments provides the correlation results and regression parameters. The classification of the entire scene provides an adjustment for any differences in crop acreage relationship between the sample of segments and the entire population of possible segments within the scene.

The measure of effectiveness of using the satellite data which is most meaningful to SRS is relative efficiency. Relative efficiency is defined as the variance of the direct expansion (ground) estimate divided by the variance of the regression estimator. This number indicates what amount of additional ground data would be needed to give an estimate with the same precision as the regression estimator. Relative efficiencies in practice have ranged from 2.0 to 4.0 or higher in most situations.

The research approach that SRS uses for full state estimates (having a revised acreage estimate by the end of the estimation season) was first utilized in 1978 for the State of Iowa and has been utilized in 1980 for Iowa and Kansas and in 1981 for those two States plus Missouri and Oklahoma. Based on those experiences and other research SRS is projecting an approach of utilizing unitemporal satellite data for estimates rather than a multitemporal approach. While multitemporal data have been used and would give better discrimination for some classification problems, it is rare to obtain cloud free images during the growing season in the Midwest or the Great Plains. Some counties in a scene are usually lost due to clouds in the growing season acquisition date and if other counties are lost in a spring scene, the resultant multitemporal data set might not have enough segments remaining for proper training and estimation. SRS does have multitemporal procedures on line and will use them for land cover estimates and for crop research in areas such as California.

Two Stages of Registration Insure Positional Accuracy of Landsat MSS Data

Since the SRS approach depends on the use of very specific data sets of fields for training it is important to match as exactly as possible the satellite and ground locations. Presently SRS uses two stages of "registration" to insure adequate positional accuracy of Landsat MSS data. The first effort is called "global registration". A sample of points is selected across the scene from transparency data products. The corresponding points on U.S. Geological Survey base maps are located and digitized. A mathematical transformation is calculated to adjust all pixels to predicted longitude/latitude locations. This process usually ends up with rectification within ± 1 to 2 pixels across the scene. This operation is now performed by the Remote Sensing Branch Support Staff and most scenes can be "registered" in about 3 hours, even with editing of outlier points.

The second stage of "registration" is called "local segment shifting" by SRS. This step is necessary because of scanner anomalies. A gray scale of a window enclosing the predicted segment is created. An overlay of the digitized segment and field boundaries (printed on transparent paper at satellite data scale and with boundaries adjusted for the path of the satellite) is then placed over the gray scale. The overlay is shifted as needed to properly line up segment and field boundaries. This row and column shift from predicted location is then utilized to correctly identify the most accurate location of each field in the segment.

This segment shifting processing results in rectification of data to the nearest one half pixel for each segment. No adjustment is made in the remainder of the scene from the first stage of rectification. Experience has shown that the segment shifting adjustments will normally vary across the scene in direction and size of shifts so the segment shifting results are not imposed on the global rectification stage. This indicates that although SRS needs ½ pixel accuracy for small scale classification of segments the accuracy of 1 pixel or so across the rest of the scene is sufficient. There are large numbers of pixels associated with each stratum within a scene and the 1 pixel potential error is not a critical percentage. SRS would replace either or both of these "registration" procedures if data received always were rectified to the appropriate accuracy.

The "40 meter Accuracy Requirement" Better Stated as "One-Half Pixel"

This might be the best point to discuss the USDA's "40 meter accuracy requirement". In the Domestic Crops and Land Cover project of the AgRISTARS research program SRS has specified 40 meters as the desired precision for both scene-to-scene and scene-to-map applications which utilize LANDSAT MSS data. Some members of the remote sensing community have interpreted this to be an absolute (ground distance) requirement. The 40 meters originated because the resolution of the MSS data as acquired was assumed to be 80 meters. Accuracy to a half pixel would be 40 meters. Regardless of effective pixel size due to processing techniques the 40 meters was regarded as the feasible positional accuracy for Landsat MSS. The 40 meters is not the important measure; the half pixel is the key.

SRS regards one half pixel to be the accuracy goal for any advanced resolution sensor. This is due to the emphasis on providing the most accurate data set for training on a field to field basis. A sensor with improved resolution such as the thematic mapper should result in improved classification results but only if a pure training and testing set can be insured.

Rectification and Registration in Forestry Applications

The U.S. Forest Service (USFS) has utilized aerial photography as a tool in providing a number of inventory and management needs. The USFS approach is similar to that of SRS in that ground observation data is or can be available. The USFS objectives are more demanding than those of SRS because a variety of types of information are desired. Instead of just estimating area of forest land there is a need for determination of forest types, measures of change, and some detailed estimates such as the annual increment of production. USFS is often interested in production estimates for land with considerable slope and varied terrain in contrast to the relatively level terrain for crop acreage estimates.

The USFS has a major research effort referred to as the Multiresource Inventory Methods Pilot Test (MIMPT) which is an advanced demonstration of the use of Landsat satellite technology to supplement current methods of conducting recurrent inventories over large land areas. The base program has established a large sample of sites which can be periodically monitored for change and new resource assessment data. Much of the necessary data can be obtained from aerial photography and thousands of aerial photos are utilized when new photography acquisitions become available. The basic reference link is provided by 7-1/2 minute USGS quadrangle maps.

Results from the first phases of the MIMPT indicated that computer interpreted Landsat data combined with data from a relatively few aerial photos can replace human interpretation of thousands of photos to estimate land use acreages. When the Landsat data are included in a geographic information system with other data sources such as topographic and soils data, it is possible to derive variables needed in multiresource surveys such as sedimentation, disturbances and other spatial dependencies such as public use, utilities, or transportation infrastructures.

Because of the combination of Landsat data with other data in a geographic information system and the necessity to extract specific features from ground sites for training and evaluation, very precise rectification of imagery is needed. The desired accuracy of the USFS would be rectification of 95 percent of all pixels within ground sites to within ± 20 meters of true ground location. This very precise goal is because of the irregular shape of many features being observed and of other variables such as topographic information. In this specific application, the agency requirement is stated in terms of a physical measurement irrespective of pixel size, in contrast to FAS and SRS requirements.

Rangeland Applications: One Pixel Accuracy May be Sufficient

The last aerospace remote sensing application of the USDA that I would like to discuss is that of assessing rangeland condition. I will not attempt to discuss any programs of the Bureau of Land Management related to rangeland carrying capacity or other measures. Instead I want to focus on research interests of USDA, notably of the Soil Conservation Service (SCS). Although some initial work has been done, research is basically just getting underway, and conclusions on accuracy are tentative at best.

SCS is interested in such factors as conservation needs ot rangeland as well as range condition and biomass production. There is a great need for additional research into methods for using satellite data in proper estimation of range condition. Rangeland typically is quite variable across an area of any size. That is, there is not usually the homogeneity that would be expected in a field of a planted crop. Sampling procedures tend to obtain kinds and amounts of biomass for relatively small areas, often for points selected by the range conservationist on the ground which are "representative" of conditions observed.

One approach which might be helpful in the evaluation of range condition would be to utilize an image as a stratification device and to collect biomass data for each stratum present in the image. However, while range condition changes with longer term use or abuse, range biomass changes quickly with rapid response to rainfall and drought. By the time an image is available, interpreted, and in the hands of a person on the ground, the conditions may have changed drastically. Thus, some type of sampling approach is needed which collects data at or near the time of satellite data acquisition to be related by cluster analysis or some other type of estimation based on the satellite data.

One approach which might be applicable to the SCS information needs is to collect data over some type of grid pattern. Collection of data for enough grid points and collection of biomass (or other information) for a large enough unit at each point to minimize within sampling unit variation would allow estimation through a poststratification approach with the strata based on the satellite imagery. Satellite data has not proven to be easily integrated into SCS's conventional range site and condition surveys, nor to the point sample used in SCS national inventories but it is an approach of great current interest.

Since there will not usually be "field" boundaries in rangeland areas it will not be possible to match data as precisely as in the SRS crop acreage approach or the Forest Service approach. Rectification of one pixel accuracy or registration at one pixel accuray if repetitive coverages are interpreted should be adequate in light with the type of ground data available. As with most other USDA applications registration or rectification needs for new sensors would continue to be in terms of one pixel accuracy rather than a specification in terms of absolute meters.

Sampling vs. Whole Frame Studies

One topic of interest to individuals concerned with registration and rectification is differences in requirements for sampling approaches versus whole frame studies. From my viewpoint of USDA requirements I feel that the answer is contained in the examples cited. For estimation purposes when whole frame estimates are created in the SRS approach or when FAS desires to make estimates for areas without corresponding ground data positional accuracy of one pixel is sufficient. However, when sampling approaches are used such as the SRS matching of ground data with corresponding satellite data for training or the USFS matching of various ground data items with satellite data, positional accuracy is more critical and one-half pixel or greater accuracy is needed. These accuracy requirements are relative to the purposes for which the data are used and should be thought of as related to the resolution of each sensor and not normally as absolute measures of ground distances.

Need for Standardized Concepts

As a closing note I would like to comment on observations from the communications during the past two years and particularly from the past few weeks in preparing for this workshop. Disagreement in definition of terms was already addressed in the opening of the paper. Even within the use of a particular definition I think there is considerable misunderstanding.

Although the term "RMS" error as in "40 meters RMS" is widely used, I would question if all people using it have the same concept of how to calculate RMS error and I feel from discussions many users are misunderstanding what is meant by a certain result. At least as used by the Statistical Reporting Service RMS error is a confidence interval statement that two-thirds of all pixels will be within plus or minus that distance from "true" location where true location may come from another data scene or from map locations. Since so much emphasis has been placed on this RMS concept some data users may believe that "all" data are within this RMS bound. SRS has adopted a procedure of calculating a second accuracy measure, the R-90 criterion, which is the radius of a circle containing 90 percent of the deviations of pixels from "true" locations. This is felt to convey more information about the accuracy of registration or rectification than just the RMS calculation. In the Forest Service example above, the data users are interested in knowing how precisely 95 percent of all points are registered which is essentially a distance equal to twice the RMS measure.

If this workshop can publicize procedures for properly measuring or calculating positional accuracy and can educate users as to proper interpretation of accuracy statements, it will have accomplished a great deal.

REFERENCES

Since little is documented in U.S. Department of Agriculture publications on registration and rectification requirements an alternate approach is to provide a list of references for more information about applications of the various Agencies. Below are listed individuals who can be contacted for the remote sensing approachs mentioned in this paper.

Foreign Agriculture Service - Bobby Spiers, Chief, Analysis Branch, Foreign Commodity Condition Assessment Division, Foreign Agriculture Service, 1050 Bay Area Blvd., Houston, Texas 77058.

U.S. Forest Service - Edgar Chapman, Cartographer, Nationwide Forestry Applications Program, U.S. Forest Service, 1050 Bay Area Blvd., Houston, Texas 77058.

Soil Conservation Service - Bill Hance, Soil Conservationist, Inventory and Monitoring Staff, Soil Conservation Service, 12th & Independence Avenue, S.W., 5241 South Building, Washington, D.C. 20250.

Statistical Reporting Service - Rich Allen, Chief, Remote Sensing Branch, Remote Sensing Branch, Statistical Reporting Service, 12th & Independence Avenue, S.W., Room 4832 South Building, Washington, D.C. 20250.