ASSESSING THE ACCURACY OF THE LANDSAT-BASED CROP ACREAGE ESTIMATES

E. M. Hsu, Lockheed Electronics Company, Inc.*
A. G. Houston, National Aeronautics and Space Administration

ABSTRACT

The Landsat-based crop acreage estimates are derived by computer-aided analysis of images of Landsat-acquired segment areas. These segment images are interpreted by the image analyst to obtain training information and are then classified to obtain crop acreages for the areas.

This paper presents an overview of the fundamental approach to acreage estimation and discusses in some detail the LACIE classification procedure. The investigations of the error sources in the Landsat-based acreage estimation are described, as well as the study of the dependence of the errors on various causative factors. Some of the results of this accuracy assessment are also presented.

1. INTRODUCTION

The Landsat is a land observatory satellite with an onboard multispectral scanner, which records reflected radiance values in four channels. To facilitate the application of Landsat data, digital values are represented as colors and are presented in imagery form. The Landsat-based acreage estimates are derived by a complex computer-aided analysis of these Landsat images.

Various classification techniques have been developed and discussed in the many articles related to remote sensing and pattern recognition. However, the procedure for which this accuracy assessment is conducted, particularly the training sample determination, has been uniquely developed in the Large Area Crop Inventory Experiment (LACIE), and only LACIE numerical examples are used in discussing the results.

The LACIE is an interagency experiment in the use of Landsat and meteorological data to identify and inventory crop acreage, yield, and production. The participating agencies include the U.S. Department of Agriculture (USDA), the National Aeronautics and Space Administration (NASA), and the National Oceanic and Atmospheric Administration (NOAA).

This paper presents an overview of the fundamental approach to acreage estimation and discusses in some detail the LACIE classification procedure. The investigations of the error sources in the Landsat-based acreage estimation are described, as well as the study of the dependence of the errors on various causative factors. Some of the results of this accuracy assessment are also presented.

2. FUNDAMENTAL APPROACH TO ACREAGE ESTIMATION

The Landsat-based crop acreage estimates are derived from computer-aided analyses of Landsat-acquired images of segment areas. A segment is a 5 by 6 nautical mile area that is used as a sampling unit on which estimates of crop acreage for a large area are based. Each segment image consists of 22 932 pixels. (The term "pixel" refers to the basic resolution element in the image, which corresponds to a 1.1-acre area on the ground.)

The use of Landsat data to perform crop acreage estimation depends on the recognition of pixels by the analyst and the classifier. The analyst selects training samples by interpreting the Landsat image through the aid of ancillary data such as cropping practice data, soil data, crop calendars, and historical crop percentages for the political regions. The classifier uses the computer to process the training information, assigning the pixels to the various crops. The classification is treated as a stratification of a scene into potential crop classes. This stratification is then used to perform a stratified areal estimate. The crop acreage estimation flow for a segment is shown in Figure 1.

3. CLASSIFICATION PROCEDURE

When the crop signature (characteristic reflectance of a crop) has been determined by an analyst, training samples are defined for computer classification of a segment. The following sections describe the approach to defining training samples that has been developed and implemented in LACIE.

3.1 Determination of Training Samples

The image display of a segment has 117 row spacings (lines) and 196 column spacings (pixels). A total of 209 pixels, which coincide with a grid spacing every 10*th* row and every 10*th* column of a segment, are selected as candidates for image interpretation. Two independently preselected random subsamples of these 209 pixels are specified to be interpreted (labeled) by the analyst, as follows:

1. The type-1 subsample consists of a minimum of 30 pixels. These labeled pixels are used to initiate the clustering process and to help in determining crop labels for clusters.

2. The type-2 subsample consists of approximately 60 pixels. These pixels are used to obtain a stratified areal estimate (see Section 3.3).

When the type-1 and type-2 subsamples are selected and labeled, clustering (computer processing) is performed to group the pixels according to a certain distance measure (in particular, the ℓ_1 metric). In LACIE, the radiance value of each pixel in a segment is compared with the value of each of the type-1 starting pixels, which are used

*This material was developed under NASA Contract NAS 9-15200 and prepared for the Earth Observations Division, National Aeronautics and Space Administration, Lyndon B. Johnson Space Center, Houston, Texas.

to initiate clustering. Each pixel is then assigned to the group of the closest starting pixel. After all the pixels are clustered, the mean and standard deviation of each cluster are computed, and the cluster map is generated by the computer. The mean value of each cluster is again distance-compared to the radiance value of each of the type-1 pixels. The cluster is given a crop label according to the label of the closest type-1 pixel. All the pixels in a given cluster are treated as observations from a crop class (or subclass) and are used to estimate class (or subclass) statistics. These statistics are then used as estimators of the classification parameters (means and covariances).

3.2 Bayes Decision Model

The classification procedure of LACIE is a Bayes procedure (Anderson 1958) with the assumptions that the loss of each class due to misclassification is equal and that the underlying distributions are normal.

Let pixel X be the random variable (or random vector) having probability density functions $f_W(x)$ and $f_N(x)$ in populations π_W and π_N, respectively, where x is the observation of X and is drawn from the mixed population $\{\pi_W, \pi_N\}$. The Bayes procedure is to minimize the expected loss due to the costs of misclassification. For a given observed x, if

$$q_W f_W(x) \ell(N/W) \geq q_N f_N(x) \ell(W/N) \tag{3.1}$$

π_W is chosen; otherwise, π_N is chosen; where q_N is the *a priori* probability that x belongs to π_N, q_W is the *a priori* probability that x belongs to π_W, $\ell(N/W)$ is the loss due to misclassifying x from π_W into π_N, and $\ell(W/N)$ is the loss due to misclassifying x from π_N into π_W.

If $\ell(N/W) = \ell(W/N)$, as is assumed in LACIE, the decision model expressed in (3.1) becomes

$$q_W f_W(x) \geq q_N f_N(x) \tag{3.2}$$

If π_W and π_N have subpopulations π_{W_i} (i = 1, 2, ···, k) and π_{N_j} (j = 1, 2, ···, m), respectively, the decision model in (3.2) is

$$\sum_i q_{W_i} f_{W_i}(x) \geq \sum_j q_{N_j} f_{N_j}(x) \tag{3.3}$$

where q_{W_i} and f_{W_i} are, respectively, the *a priori* probability and the probability density function that correspond to π_{W_i}; similarly q_{N_j} and f_{N_j} correspond to π_{N_j}. In the LACIE application, π_{W_i} (i = 1, 2, ···, k) and π_{N_j} (j = 1, 2, ···, m) represent the subpopulations (subclasses) for small grains and nonsmall grains, respectively. The subclass statistics are generated in the clustering process as described in Section 3.1.

When the computer classification is made using the decision model in (3.3) for all the pixels in a segment, the small-grain proportion estimate $\hat{p}_W$ and the nonsmall-grain proportion estimate $\hat{p}_N$ are calculated by

$$\hat{p}_W = \frac{N_W}{N} \tag{3.4}$$

and

$$\hat{p}_N = \frac{N_N}{N} \tag{3.5}$$

where N is the total number of pixels in a segment (i.e., 22 932), N_W is the number of pixels classified as small grains, and N_N is the number of pixels classified as nonsmall grains. Consequently, the acreage estimate for small grains, $\hat{A}_W$, can be approximated by

$$\hat{A}_W = 1.1 \text{ acres} \times N_W \tag{3.6}$$

3.3 Stratified Areal Estimation

Since experience has shown that there is bias in $\hat{p}_W$ (the computer-classified areal estimate), the stratification technique is applied after the classification. The stratified areal estimate for small grains is obtained by combining the results of the type-2 sample with the direct computer classification results, using the following formula:

$$\hat{p} = \lambda_W \hat{p}_W + \lambda_N \hat{p}_N \tag{3.7}$$

where λ_W is the number of type-2 pixels called small grains by the analyst and classified small grains by the computer divided by the number of type-2 pixels classified by the computer as small grains, and λ_N is the number of type-2 pixels called small grains by the analyst and classified as nonsmall grains by the computer divided by the number of type-2 pixels classified by the computer as nonsmall grains.

It is clear that the stratification is not made according to the true ground-observed information. That is, the bias correction factors λ_W and λ_N will not correct the mislabeling of pixels by the analyst.

4. FIELD DATA ACQUISITION

In order to assess the accuracy of a Landsat-based crop acreage estimate at the segment level, the field observation data for the segment must be acquired. The segments which are designated for field data acquisition are chosen at random from the sample segments. The identity of these selected segments is withheld from the analysts so that these segments can be treated the same as the other segments. Then, the high-resolution color-infrared aerial photography over these segments is acquired before harvesting of the crop of primary interest. Simultaneously, field teams collect ground observation data for these segments. The field team labels all the fields on a pre-prepared field overlay according to the ground-observed crop types and makes a general appraisal of crop conditions in the segment. The application of ground-observed data is described in Sections 5 and 6. Figure 2 shows the field data acquisition flow.

5. RESULTS OF COMPARISONS BETWEEN PROPORTION ESTIMATES AND GROUND-OBSERVED PROPORTIONS

When the field data become available, the ground-observed crop proportions can be obtained through the use of computer digitization or manual planimetry. A study comparing the LACIE small-grain proportion estimates obtained from the stratified areal estimator and the corresponding ground-observed proportions was conducted for both winter and spring small grains based on the ground-observed segments. Ninety-one ground-observed segments for winter small grains were randomly selected from the states that are the major producers of winter small grains; namely, Colorado, Kansas, Nebraska, Oklahoma, Texas, Montana, and South Dakota. A total of 53 ground-observed segments for spring small grains were randomly selected from the U.S. northern Great Plains states, which consist of Minnesota, Montana, North Dakota, and South Dakota. The ground-observed proportions were compared with both the early-season and late-season proportion estimates for each type of the small grains. Only 77 early-season proportion estimates were available for the 91 winter small-grain ground-observed segments; 31 were available for the 53 spring small-grain ground-observed segments.

Table 1 and Figure 3 show that on the average the proportions were significantly underestimated for both winter and spring small grains at the 10-percent level. However, the bias due to classification, $\bar{D}$, decreased from -9.8 percent in the early season to -2.4 percent in the late season for the winter small grains and decreased from -3.8 percent to -3.3 percent for the spring small grains. The primary reason for this large underestimation during the early season is that the small-grain signatures are not well-developed at this time, and as a result, some small-grain training pixels were mislabeled as other crops by the analyst. When a significant improvement in detectable small-grain signatures is noted in the later season or some backup acquisitions become available for multitemporal interpretation, the LACIE estimates begin to approach the ground-observed proportions.

The plots in Figure 3 also indicate an overall trend toward a negative value of $\hat{p}_W - p_W$ as p_W increases. In other words, LACIE tends to underestimate the true small-grain proportion when that proportion is large.

In another interesting study, a comparison is made between small-grain areal estimation errors that resulted from the computer classification estimation, stratified areal estimation, and random sample estimation from the type-2 pixels. The mean square errors for these three types of areal estimates are tabulated in Table 2 for the States of North Dakota and Montana. This table shows that the mean square error of the stratified areal estimate is much smaller than that of the computer classification estimate. However, the mean square error of the stratified areal estimate is very close to that of the random sample estimate from the type-2 labeled pixels.

6. RESULTS OF LABELING EVALUATION

The accuracy of the stratified areal estimate is critically dependent on correct labeling of the type-2 pixels by the analyst. The stratified random sampling technique provides an optimal estimate provided that the analyst can correctly identify the type-2 pixels. Thus, this labeling evaluation study concentrates on the type-2 pixels.

Each segment is evaluated by checking for consistency and discrepancy between analyst and ground-observed labels. The confusion matrix is calculated from type-2 labeled pixels for a set of ground-observed segments. This matrix is used for quantitative examination since it contains the probabilities of correct labeling and mislabeling for both small grains and nonsmall grains. Specifically, the confusion matrix M is defined by

$$M = \begin{bmatrix} p(W/W) & p(N/W) \\ p(W/N) & p(N/N) \end{bmatrix}$$

where p(W/W) is the probability of correctly labeling small-grain pixels as small grains, p(N/N) is the probability of correctly labeling nonsmall-grain pixels as nonsmall grains, p(W/N) is the probability of mislabeling nonsmall-grain pixels as small grains, and p(N/W) is the probability of mislabeling small-grain pixels as nonsmall grains. When the areal estimate is made for small grains, p(W/N) is called the commission error, and p(N/W) is the omission error.

The following results were obtained from the type-2 samples of 18 ground-observed segments in North Dakota for crop year 1976-77.

$$M = \begin{bmatrix} \frac{341}{455} = 0.750 & \frac{114}{455} = 0.250 \\ \frac{30}{563} = 0.053 & \frac{533}{563} = 0.947 \end{bmatrix}$$

This matrix shows that 75 percent of the type-2 small-grain pixels and 94.7 percent of the type-2 nonsmall-grain pixels were correctly labeled by the analyst. In other words, in the State of North Dakota, the labeling error for small grains is larger than that for nonsmall grains. The causative factors for labeling errors in North Dakota are investigated by evaluating such data as the acquisition history, individual field size, pixel location, and signature quality. Table 3 presents the error sources and the percents of commission error and omission error caused by each error source.

7. MAJOR ERROR SOURCES IN ACREAGE ESTIMATION

The major error sources in the acreage estimates at the segment level have been identified through the study in various experiments. These major error sources include:

1. Missing key acquisitions, which generally cause the analyst to be unable to separate the signatures of competing crops and thus lead to mislabeled training data.

2. Abnormal signature development, which is often caused by such problems as late planting, drought, crop rotation, disease, and soil variability.

3. Inadequacy of the Landsat scanner in resolving small fields.

ACKNOWLEDGEMENTS

Thanks are given to J. Clinton and E. Cooper for providing information on labeling evaluation.

REFERENCES

Anderson, T. W. (1958), *An Introduction to Multivariate Statistical Analysis*, New York: John Wiley and Sons, Inc.

Duda, R. O., and Hart, P. E. (1973), *Pattern Classification and Scene Analysis*, New York: John Wiley and Sons, Inc.

Fukunaga, K. (1972), *Introduction to Statistical Pattern Recognition*, Academic Press.

National Aeronautics and Space Administration, Lyndon B. Johnson Space Center (1978), *Large Area Crop Inventory Experiment Phase I and Phase II Accuracy Assessment Final Report.*

TABLE 1. COMPARISON OF PROPORTION ESTIMATES AND GROUND-OBSERVED PROPORTIONS FOR CROP YEAR 1976-77

Type of small grains	Acquisition	M	$\hat{p}_W$	$\bar{p}_W$	$\bar{D}$	$S_{\bar{D}}$	90-percent confidence limits for μ_D
Winter	Early (prior to February)	77	15.6	25.3	-9.8	1.7	(-12.6, -7.1)*
	Late (prior to October)	91	22.2	24.6	-2.4	0.7	(-3.6, -1.2)*
Spring	Early (prior to August)	31	15.3	19.1	-3.8	1.5	(-6.3, -1.3)*
	Late (prior to October)	53	15.7	18.9	-3.3	0.9	(-4.8, -1.8)*

LEGEND:

M — Number of ground-observed segments available

$\hat{p}_W$ — Average of small-grain proportion estimates obtained from the stratified areal estimator

$\bar{p}_W$ — Average ground-observed small-grain proportion

$\bar{D}$ — $\hat{p}_W - \bar{p}_W$

$S_{\bar{D}}$ — Standard error of $\bar{D}$

μ_D — Population $\bar{D}$

* — μ_D not significantly different from zero at 10-percent level

TABLE 2. MEAN SQUARE ERRORS FOR COMPUTER CLASSIFICATION ESTIMATION, STRATIFIED AREAL ESTIMATION, AND RANDOM SAMPLE ESTIMATION FROM TYPE-2 PIXELS

State	Number of segments used	Computer classification estimate	Stratified areal estimate	Random sample estimate
North Dakota	22	235.4	168.2	170.5
Montana	12	88.5	55.4	42.2

TABLE 3. NORTH DAKOTA LABELING ERRORS

Error source	Omission error (percent)	Commission error (percent)
Insufficient acquisitions	13.2	33.3
Fields too narrow	11.4	3.3
Border pixel	28.9	23.3
Abnormal signature	21.1	16.7

FIGURE 1. LANDSAT-BASED CROP ACREAGE ESTIMATION FLOW

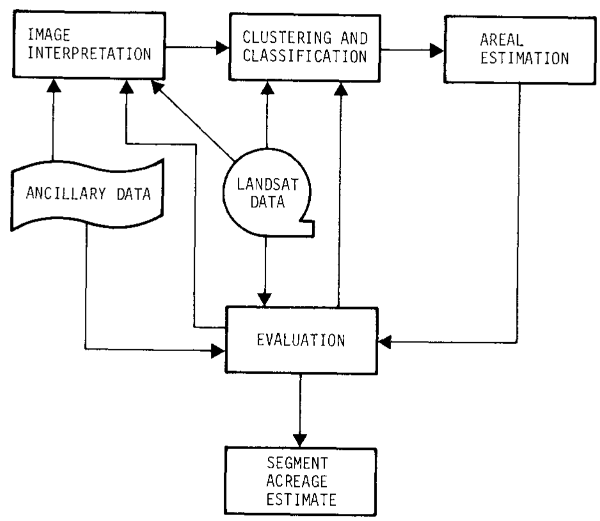

FIGURE 2. FIELD DATA ACQUISITION FLOW

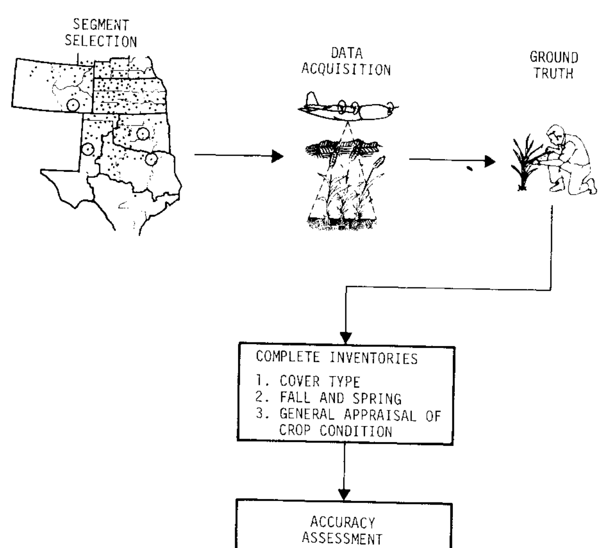

FIGURE 3. PLOT OF PROPORTION ESTIMATION ERRORS
VERSUS GROUND-OBSERVED PROPORTIONS