Arizona Vegetation Resource Inventory (AVRI) Accuracy Assessment: Final Report

By John Szajgin, Lawrence R. Pettinger, David S. Linden and Donald O. Ohlen

---

Geological Survey Open File Report No. 82-814

1982

CONTENTS

ILLUSTRATIONS

TABLES

APPENDIX

ARIZONA VEGETATION RESOURCE INVENTORY (AVRI)

ACCURACY ASSESSMENT: FINAL REPORT

By John Szajgin[1/], Lawrence R. Pettinger[2/], David S. Linden[1/],

and Donald O. Ohlen[1/]

ABSTRACT

A quantitative accuracy assessment was performed for the vegetation classification map produced as part of the Arizona Vegetation Resource Inventory (AVRI) project. This project was a cooperative effort between the Bureau of Land Management (BLM) and the Earth Resources Observation Systems (EROS) Data Center. The objective of the accuracy assessment was to estimate (with a precision of ± 10 percent at the 90 percent confidence level) the commission error in each of the eight level II hierarchical vegetation cover types. A stratified two-phase (double) cluster sample was used. Phase I consisted of 160 photointerpreted plots representing clusters of Landsat pixels, and phase II consisted of ground data collection at 80 of the phase I cluster sites. Ground data were used to refine the phase I error estimates by means of a linear regression model. The classified image was stratified by assigning each 15-pixel cluster to the stratum corresponding to the dominant cover type within each cluster. This method is known as stratified plurality sampling.

Overall error was estimated to be 36 percent with a standard error of 2 percent. Estimated error for individual vegetation classes ranged from a low of 10 percent ± 6 percent for evergreen woodland to 81 percent ± 7 percent for cropland and pasture. Total cost of the accuracy assessment was $106,950 for the one-million-hectare study area.

The combination of the stratified plurality sampling (SPS) method of sample allocation with double sampling provided the desired estimates within the required precision levels. The overall accuracy results confirmed that highly accurate digital classification of vegetation is difficult to perform in semiarid environments, due largely to the sparse vegetation cover. Nevertheless, these techniques show promise for providing more accurate information than is presently available for many BLM-administered lands.

[1/] Technicolor Government Services, Inc. Work performed under U.S. Geological Survey contract no. 14-08-0001-20129.

[2/] U.S. Geological Survey

INTRODUCTION

The U. S. Geological Survey's Earth Resources Observation Systems (EROS) Data Center has been cooperating with the Bureau of Land Management (BLM) to evaluate the utility of digital Landsat and terrain data for wildland resource mapping and inventory. The areas studied are typical of the lands administered by BLM. One EROS-BLM cooperative project, the Arizona Vegetation Resource Inventory (AVRI), was conducted in a one-million-hectare area in northwest Arizona (figure 1) (Rohde and Miller, 1981). The objective of the study was to evaluate the utility of Landsat digital data, digital elevation data, and stratified-cluster sampling for mapping vegetation in the arid and semiarid southwestern United States.

Project results included (1) vegetation maps from clustering of Landsat multispectral scanner (MSS) data and a detailed description of the vegetative composition (at level IV in a hierarchical classification framework) of each spectral class, (2) a more general cover type map (level II) produced by grouping the spectral classes and using photointerpretation and digital elevation data, (3) a digital data base for use in making decisions about resource management alternatives, and (4) map overlays which identify areas of high potential for specific management activities. The original project did not provide resources to assess the accuracy of the level II vegetation map. These resources were provided later, however, and permitted this accuracy assessment to be performed.

The EROS Data Center and BLM agreed to conduct a quantitative accuracy assessment of the AVRI level II vegetation classification map (figure 2) which would demonstrate whether Landsat data could be used for cover type mapping in semiarid environments.

Three types of classification error were considered: errors of commission, errors of omission, and overall error. A commission error occurs when a Landsat picture element (pixel) is classified as a particular cover type but is later found to be some other cover type when field checked. An omission error occurs when a pixel, known to be a specific cover type after field checking, is misclassified. For example, a pixel classified as Great Basin desert shrub but known to be evergreen woodland would be a commission error for the Great Basin desert shrub cover type and an omission error for the evergreen woodland cover type. Overall error would be the proportion of the total number of pixels which are incorrectly classified--without regard to cover type.

OBJECTIVES

The objective of the accuracy assessment was to estimate the commission error in each of the eight level II vegetation cover types with a precision of $\pm$ 10 percent at the 90 percent confidence level. That is, the estimated value would have to be within 10 percent of the true value in 9 samples out of 10. Although the sampling frame was designed to estimate commission error, estimates of omission and overall error were also computed. However, no precision levels were specified for these estimates because it would have been impossible to control the sample allocation for these estimates.

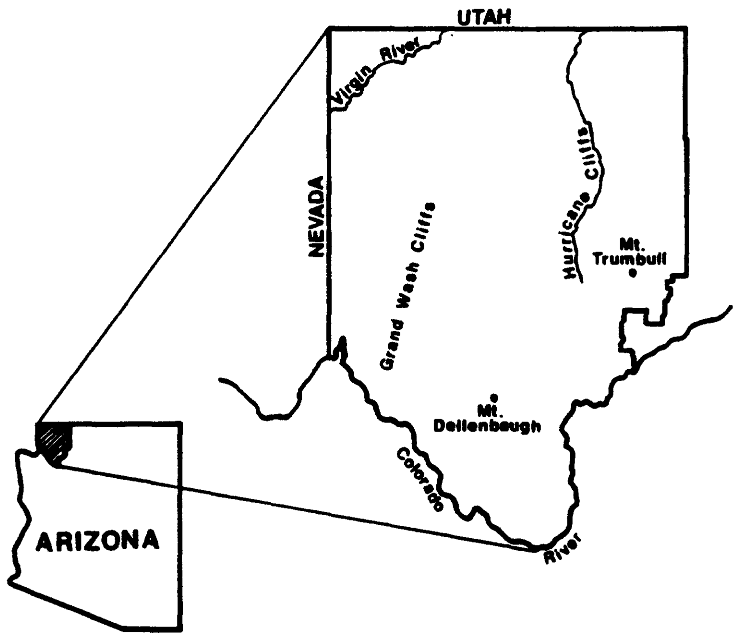


Figure 1.--Map showing location of the AVRI study area in northwestern Arizona.

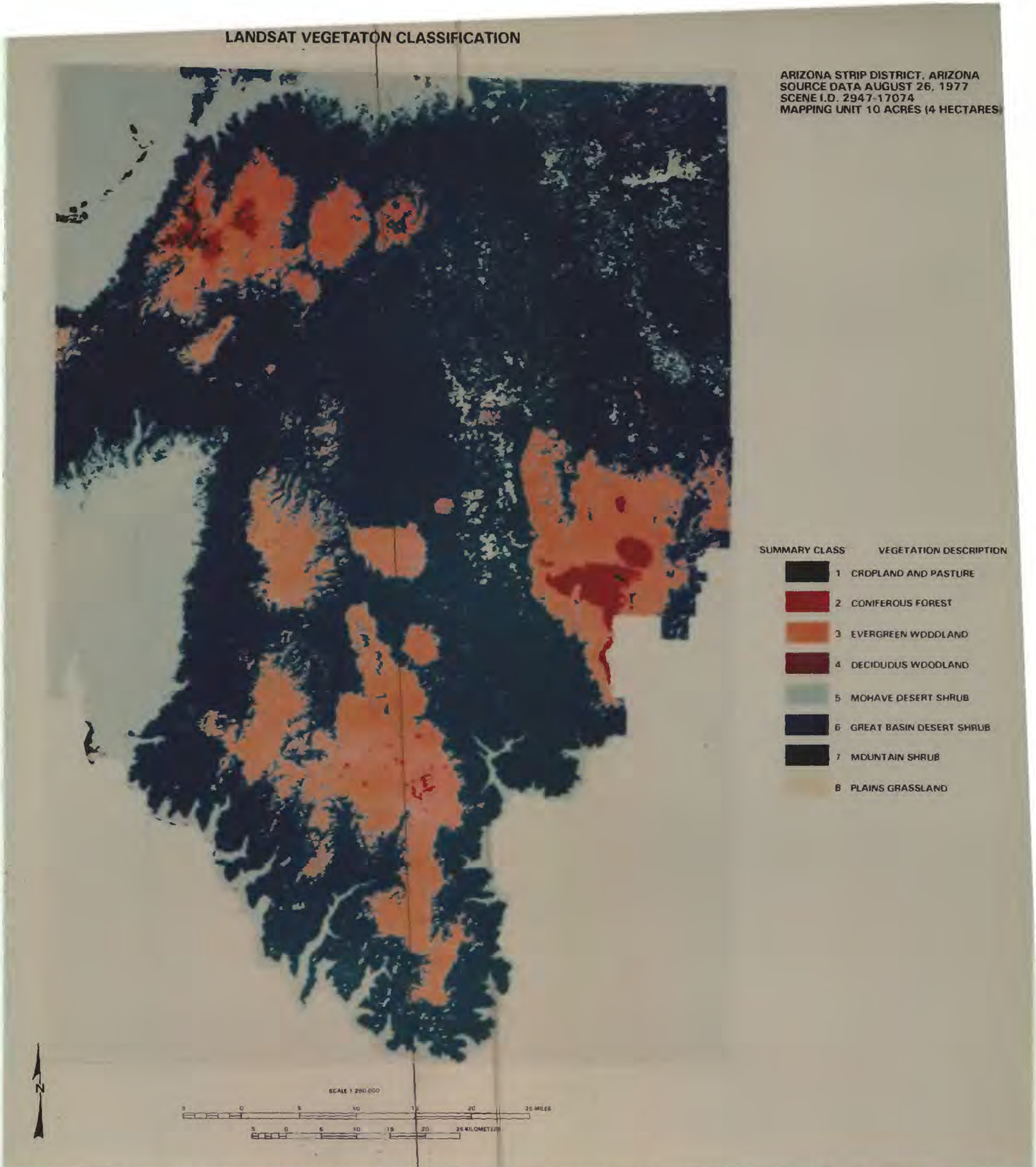


Figure 2.—Landsat-derived vegetation map of the AVRI study area. (Original map reproduced in color.)

This report presents the accuracy assessment results and discusses sample design and allocation, photointerpretation, and field data collection methodology, so that these procedures might be evaluated for use in future projects.

OVERVIEW OF AVRI CLASSIFICATION

Digital Landsat data acquired on August 26, 1977, for the study area were geometrically corrected and registered to a 50-meter square Universal Transverse Mercator (UTM) grid. The project boundary and BLM administrative units were plotted onto 1:250,000-scale maps, digitized, and registered to the same UTM grid. A mask of the project area was applied to the Landsat data so that only those Landsat pixels within the study area would be classified.

A controlled clustering technique, similar to that described by Fleming and others (1975), and an unsupervised clustering technique (Rohde, 1978) were used to separate a sample of the Landsat data into 76 spectrally distinct classes. Based on the spectral response for each of the four Landsat MSS bands, a maximum-likelihood classifier assigned each pixel in the study area to one of the 76 spectral classes. Each spectral class was then displayed systematically on the color monitor of the interactive analysis system. Color-infrared aerial photographs and vegetation maps of selected areas were used to aid in assigning each class to one of eight vegetation cover types (seven wildland vegetation types and cropland/pasture). Photographs of these cover types appear in figure 3.

Large-scale (1:6,000) natural-color aerial photographs over a sample of the classified Landsat pixels were interpreted to provide a more quantitative description of the mapped cover types. A total of 119 8- by 8-pixel sample clusters was allocated proportionately, based on strata (cover type) size. A tabular comparison of Landsat classification with the photointerpretation data showed differences between spectral classes and mapped cover types. Digital elevation data were then used with a modified parallelepiped classifier (Fleming and others, 1975) to reassign pixels to more appropriate cover types based on the known elevation range of each cover type. BLM field office personnel assisted in developing decision rules based on elevation for each of the 76 spectral classes. All pixels in the study area were reclassified using this procedure. An estimate of the improvement in percent correct classification was made using the photointerpretation data from the sample clusters described above. Overall correct classification improved from 54 percent without elevation data to 73 percent with elevation data (table 1).

Final products included color-coded maps (figure 2), area summaries by cover type (table 2), and statistical summaries of cover type composition.

Cropland and pasture

Plains grassland

Mohave desert shrub

Great Basin desert shrub

Mountain shrub

Evergreen woodland

Deciduous woodland

Coniferous forest

Figure 3.—Photographs of the eight vegetative cover types. (Original photographs reproduced in color.)

Table 1.--Comparison of accuracy estimates for cover type classification produced with and without elevation data

[Estimates were made from photointerpretation data used to label spectral clusters]

Cover Type	Without elevation data: Percent correct	Without elevation data: Standard error (percent)	With elevation data: Percent correct	With elevation data: Standard error (percent)
Cropland and pasture	15	17	19	23
Coniferous forest	72	7	81	12
Evergreen woodland	77	6	81	5
Deciduous woodland	4	4	70	23
Mohave desert shrub	74	6	96	2
Great Basin desert shrub	57	4	68	4
Mountain shrub	38	14	58	15
Grassland	1	1	2	1
OVERALL	54	5	73	5

Table 2.--Area summaries and comparison of the proportion of level II cover types from Landsat classification with corresponding stratified plurality sampling strata

Cover type	Area (ha)	Percent of area from Landsat classification	Percent of area from SPS strata
Cropland	1,333	<1	<1
Coniferous forest	9,073	1	1
Evergreen woodland	171,464	17	17
Deciduous woodland	817	<1	<1
Mohave desert shrub	165,131	16	17
Great Basin desert shrub	610,559	61	63
Mountain shrub	8,330	1	<1
Plains grassland	33,449	3	1
Total	1,000,156		

SAMPLE DESIGN AND ALLOCATION

The sample design used to assess accuracy was a stratified two-phase (double) cluster sample. Stratification was used to reduce variation and to increase precision of the desired estimates. Cluster samples were chosen because they are more efficient to use in wildland areas where the greatest cost in performing field work is associated with traveling to the sample sites. Once a pixel site is located on the ground, it is more efficient to collect data from a number of adjacent pixel sites than to locate and visit widely scattered individual sites (further discussion of the advantages of cluster sampling is given by Cochran, 1977). Phase I of the sampling consisted of the interpretation of cover types from large-scale aerial photographs of the sample cluster areas, and phase II consisted of ground data collection in half of the phase I clusters. The ground data were used to refine the phase I error estimates.

The classified image of the study area was stratified according to the eight cover types shown on the level II map. If the sample units had been defined as single pixels, then each stratum would have consisted of all the pixels labeled as the cover type corresponding to that stratum. Because clusters of pixels were sampled, and many clusters contain two or more cover types, a method was needed to assign each cluster to a single stratum so that the resulting set of strata would have a high degree of association with the cover types derived from Landsat classification. To do this, each cluster was assigned to the stratum corresponding to the dominant cover type within the cluster. This method was therefore called stratified plurality sampling (SPS), and was previously described by Linden and Szajgin (1981). By using this method of stratification, each sample cluster was assigned to the stratum for which it could provide the most information.

Several considerations contributed to the selection of a 3- by 5-pixel (15-pixel) cluster size. The primary concern was to ensure that the entire pixel cluster could be imaged on the large-scale aerial photographs. If a larger cluster size had been specified, the risk of getting only partial photographic coverage would have increased significantly. Previous experience in similar environments had shown that two clusters containing approximately 15 pixels each could be located and field-checked per day. Since field personnel and helicopter flight time were limited, a larger cluster-size requirement would have reduced the total number of clusters that could be visited. Finally, earlier work with larger clusters (8- by 8- and 10- by 10-pixel blocks) had also shown that there were diminished benefits.

A 3- by 5-pixel shape was used, and the photographic survey flight lines were oriented parallel to the longer cluster axis to increase the likelihood of achieving complete photographic coverage. This rectangular cluster shape also encompassed areas of greater ecological diversity than would a square cluster shape.

To accomplish the stratification, the classified image was first divided into a population of mutually exclusive 3- by 5-pixel clusters (figure 4) from which the sample was drawn. Each cluster was assigned to the stratum corresponding to the dominant cover type within the cluster (figure 5).

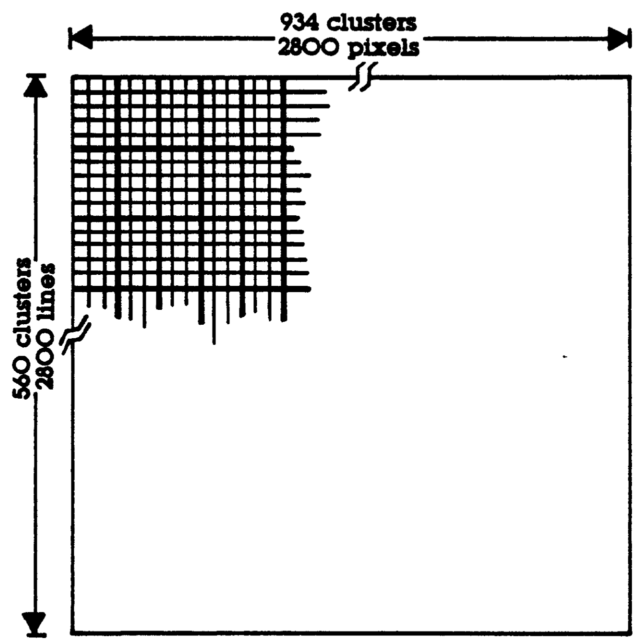


Figure 4.—Sketch of the study area image subdivided into a matrix of 3- by 5-pixel clusters. Phase I samples were selected from this population of clusters.

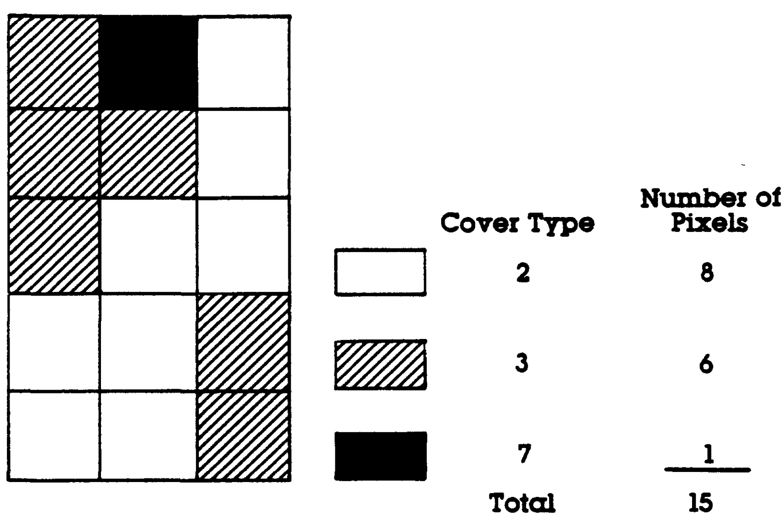


Figure 5.--Example showing the allocation of a sample cluster to a cover type stratum. This cluster was assigned to stratum 2 because the majority of its pixels were classified as cover type 2.

This approach provided the necessary control for allocating the sample clusters. Table 2 shows the close association between areal extent of cover type and SPS strata.

Figure 6 shows the number of pixels by cover type expected within a randomly selected sample cluster compared to that actually achieved by using the SPS method. The effectiveness of the SPS method for assuring adequate representation of all cover types is apparent. Although a particular cover type corresponding to the stratum under allocation was guaranteed to have the plurality in any selected cluster due to the definition of the strata, all pixels in a selected cluster were used in the estimation process.

Considering the desired level of precision and confidence, previous experience in accuracy assessment, and available resources, a sample of 160 3- by 5-pixel clusters was allocated equally to the eight strata in phase I (Linden and Szajgin, 1981). Clusters were allocated equally within the predefined strata because the relative variability of individual cover type accuracy was not known. Since the photointerpretation data collected to label the spectral clusters in the AVRI project were based on a stratification using a slightly different set of cover type classes, and because the estimates from these data were believed to be somewhat biased, that sample was not used to assess the within-cover-type variability. Accuracy was estimated by comparing photointerpretation classification of the sample photographs for each of the clusters with the corresponding digital Landsat classification.

Some ground data were necessary to assure a refined estimate of the accuracy of the digital classification. In phase II, ground data for a subsample of 80 clusters (allocated equally to the eight strata) were used to refine the phase I estimate. A least squares method was used to develop a regression relationship between the photointerpretation accuracy estimates and ground data accuracy estimates. This regression relationship was used in a double sample estimation procedure (Cochran, 1977; Hansen and others, 1953; and Scheaffer and others, 1979) to refine the desired estimates of accuracy[1/].

ACQUISITION OF AERIAL PHOTOGRAPHS

Natural-color aerial photographs were acquired for the phase I evaluation under contract in early to midsummer of 1980, at a scale of 1:2,000. This season was chosen because it would be best for discriminating the vegetative communities. Photographs of the allocated sample units were obtained during a 6-week period in September and October of 1980. Although this was later than specified, project staff judged that the photographs were satisfactory for making level II cover type interpretations.

[1/] After the project was completed, photointerpretation results were compared with ground data to assess whether a two-phase sample was justified. The photointerpretation results were only 70 percent accurate, indicating that ground data were, in fact, needed to refine the photointerpretation estimates.

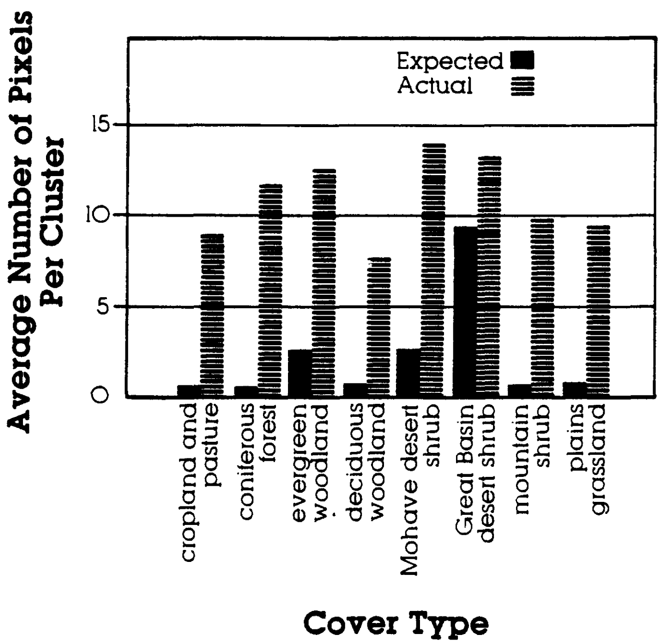


Figure 6.—Comparison of average pixel counts for sample clusters expected from a simple random sample with the actual average pixel counts achieved using the stratified plurality sampling method. Using a simple random sample, the expected average number of pixels per class would be proportional to the class size (see table 2). Using the stratified plurality sampling method, the expected number of pixels for each class is more uniform because the cluster composition is more homogeneous within a stratum.

Fieldwork for this project was scheduled for April of 1980 when personnel and logistical support were available. Since the large-scale color photographs would not be available for use as navigation aids during this field effort, another aerial survey contract was let. Black-and-white photographs at a scale of 1:3,000 were acquired quickly for field location and percent cover estimation. These photographs were acquired in spring of 1980 and were received in time to be used in the field. This scale proved adequate to discern details within individual pixels and still provide a sufficiently wide field of view to be useful in locating each cluster.

A third type of aerial photograph was used in the study. Color-infrared aerial photographs (scale 1:24,000) had already been acquired in 1976 for use in BLM resource management activities. These photographs were also used in the project for general reference and location purposes.

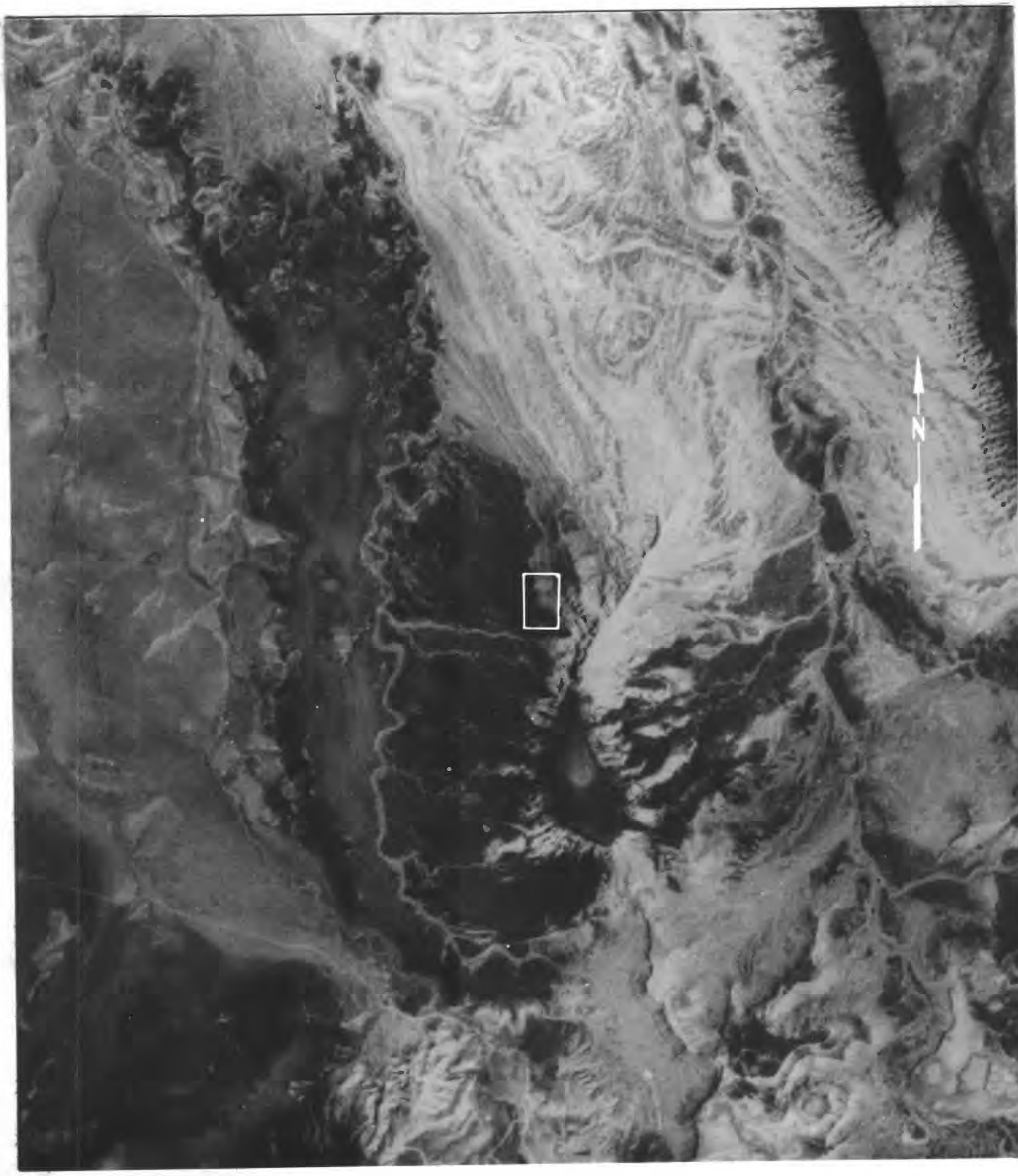
PLOTTING OF SAMPLE CLUSTERS

Field data accuracy depends upon precise onsite location of sample clusters. Since accurate location of sample clusters has historically been a problem, new procedures for plotting and locating sample clusters were developed and tested during this project. Because photogrammetric control was limited and because inherent distortions due to relief displacement occur in aerial photographs, the selected sample clusters could not be plotted by machine directly onto the large-scale aerial photographs. The following steps were used to accomplish the plotting:

1. Using the known coordinates of the corners of the sample cluster in the UTM-registered image, outlines of the clusters were plotted onto overlays of orthophotoquads (figure 7) using a Calcomp flatbed plotter.

2. The four corners of a cluster were visually transferred to an acetate overlay placed over the appropriate 1:24,000-scale color-infrared BLM resource photographs (figure 8). Since the scale of the resource photographs approximated the scale of the orthophotoquads, this transfer was relatively easy to make. The plotted boundary of clusters often did not have parallel sides because features on the photographs had inherent image displacement due to relief.

3. Cluster boundaries were transferred directly onto the 1:2,000-scale color aerial photographs (figure 9) by comparing the image detail with the resource photographs. Boundaries of the 80 clusters to be field-checked were also transferred to the 1:3,000-scale black-and-white aerial photographs (figure 10). These photographs were originally acquired for use in field positioning. They were also used in intermediate plotting between the 1:24,000-scale and the large-scale color aerial photographs because the slightly smaller scale and black-and-white tones were easier to compare to the resource photographs than were the large-scale color photographs.



METERS
0 240 480

Figure 7.--Portion of the Yellowhorse Flat 7.5-minute orthophotoquad in northwestern Arizona showing the plotted outline of sample cluster 129.

Figure 8.--Portion of color-infrared BLM resource aerial photograph (date: October 3, 1976; ID no. AZAS 7-15-77) showing the outline of sample cluster 129. (Original photograph reproduced in color.)

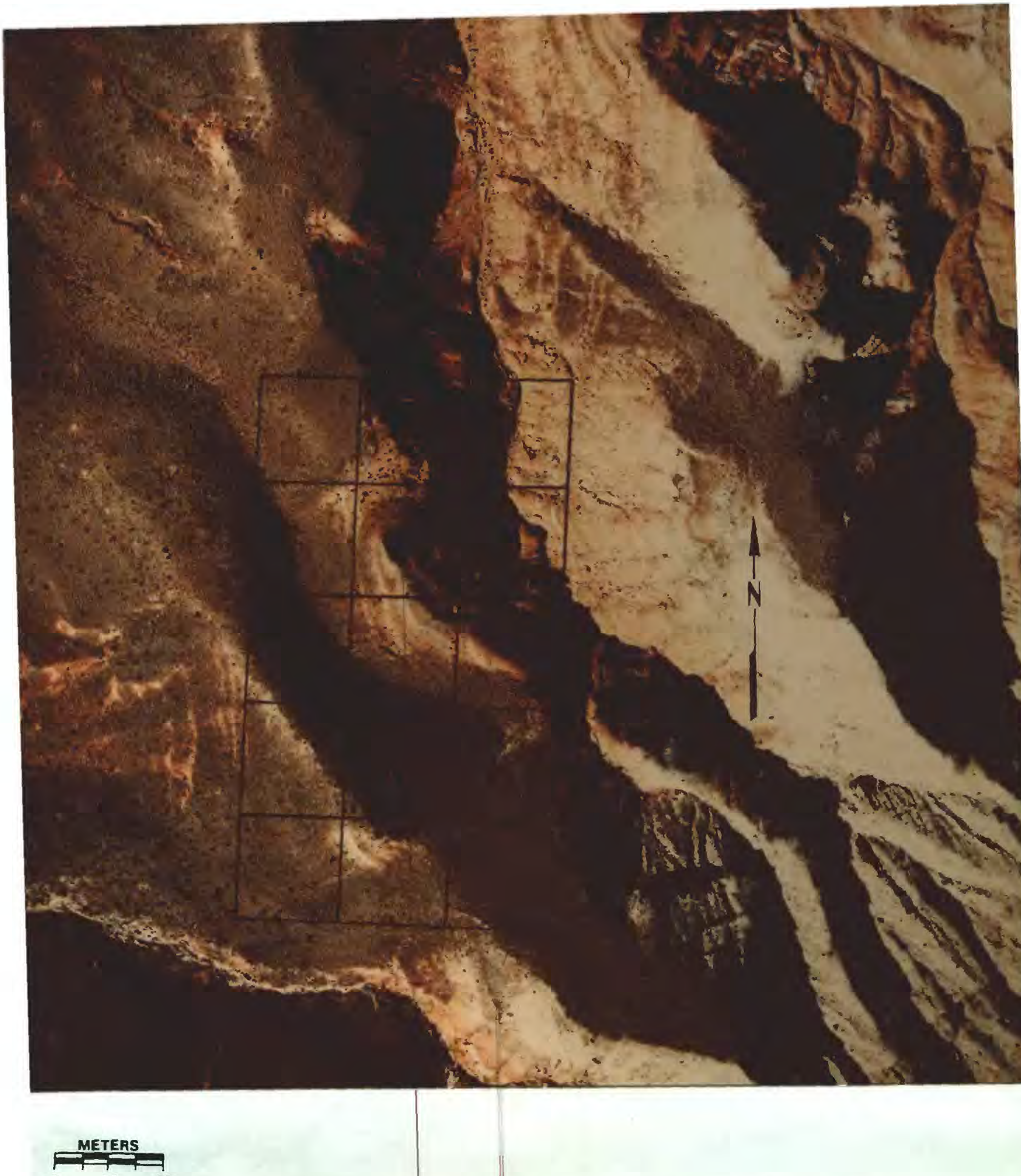


Figure 9.--Portion of natural-color aerial photograph (date: October 9, 1980; ID no. 129-03) with outline of sample cluster 129. These photographs were used for the Phase I interpretation. (Original photograph reproduced in color.)

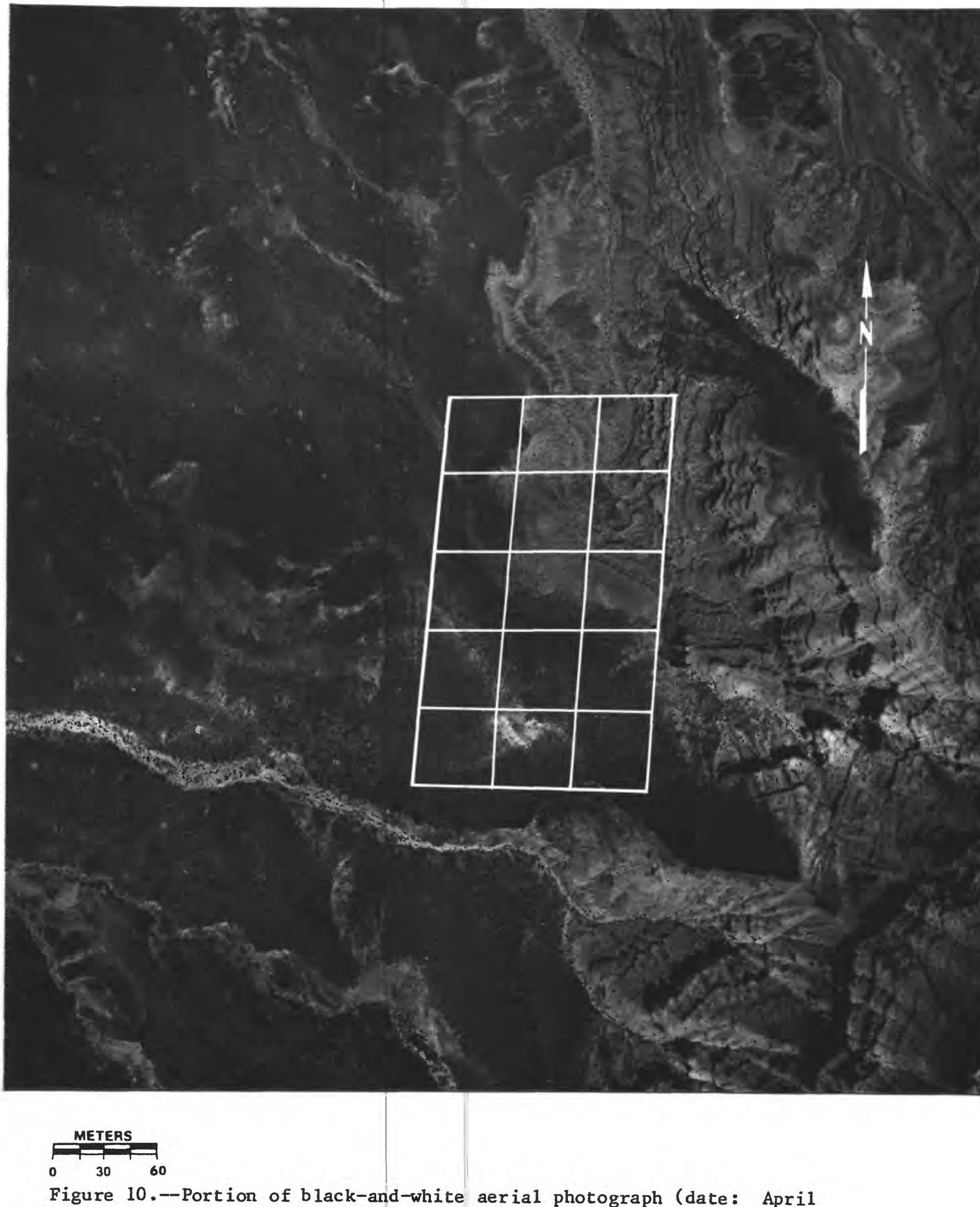


Figure 10.--Portion of black-and-white aerial photograph (date: April 27, 1980; ID no. 8000316-SH 129-2) with the outline of sample unit 129. These photographs were used primarily in the field to locate corners of individual pixel sites.

4. The 3- by 5-pixel grid was established within each cluster by using proportional dividers. This approach did not account for all photographic displacement, but it provided acceptable representation of Landsat pixel locations.

Errors in cluster locations can be attributed largely to the error associated with the geometric registration of the Landsat data to the UTM grid. The mean residual errors between predicted and observed locations of control points were ±50 m, or about 1 resampled UTM grid pixel (Rohde and Miller, 1981). The aerial photographs were of good quality and there were adequate image features to permit transfer of cluster corner locations. Therefore, the only significant plotting errors (estimated at perhaps 10-15 m) resulted from using proportional dividers to locate individual pixels in areas of high relief. Even in these cases, errors associated with the geometric registration were much greater.

PHASE I PHOTOINTERPRETATION PROCEDURES

Two interpreters were needed to complete the work within the time available. Each individual had a natural resources background, experience in interpretation of medium- and large-scale aerial photographs, and had performed fieldwork in the AVRI project study area. Each individual interpreted half of the sample clusters, which were selected on a random basis.

Once the 1:2,000-scale natural-color aerial photographs (figure 9) were obtained, each interpreter reviewed his field notes and studied training examples prepared from color aerial photographs of selected sample units. Ground identifications made during fieldwork were used to confirm the interpretations made during the training phase. Image characteristics such as size, shape, texture, pattern, location, association, and height were used to identify the cover types.

Duplicate positive transparencies were interpreted while being viewed on a light table with an Old Delft Scanning Stereoscope[2/]. Available reference material included the classification key (table 3), field notes, ground photographs, and topographic maps. A random pattern dot grid (200 dots/in^2) was used to estimate percent cover if cover was a criterion for identification. Interpretations were recorded as digital codes on a data sheet.

PHASE II FIELD DATA COLLECTION PROCEDURES

Sample clusters were located using 1:250,000 topographic maps, 1:125,000 NASA color-infrared photographs, 1:24,000 color-infrared BLM resource photographs (figure 8), and the 1:3,000 black-and-white panchromatic photographs (figure 10). This sequence of photographs (with varying scale, area of coverage, and detail) offered considerable versatility. The flight navigator and crew chief confidently used these materials to quickly locate a corner of each sample cluster.

[2/]Any use of trade names and trademarks in this publication is for descriptive purposes only and does not constitute endorsement by the U.S. Geological Survey.

Table 3.--Level II classification key

1.	<5 percent vegetated	2
1.	≥5 percent vegetated	3
2.	≥50 percent of pixel is water	Water
2.	Otherwise	Barren
3.	≥50 percent ground cover is agricultural	Cropland and pasture
3.	Otherwise	4
4.	≥10 percent of ground cover is tree cover	5
4.	<5 percent of ground cover is shrub cover	Plains grassland
4.	Otherwise	7
5.	≥50 percent of tree cover is ponderosa pine or fir	Coniferous forest
5.	Otherwise	6
6.	≥50 percent of tree cover is pinyon pine-juniper	Evergreen woodland
6.	Otherwise	Deciduous woodland
7.	≥50 percent of shrub cover is Mohave desert shrub species	Mohave desert shrub
7.	≥50 percent of shrub cover is mountain shrub species	Mountain shrub
7.	Otherwise	Great Basin desert shrub

After a sample cluster corner was located using a Hughes 500D helicopter, it was determined whether the cover type in each pixel of the sample cluster could be classified without landing. If so, the field form (figure 11) was completed and note made that the sample cluster was classified from the air. If the entire sample cluster could not be classified from the air, landing was made as close as possible to one of the cluster corners. Using the 1:3,000 black-and-white photographs, the crew then located the remaining three corner points of the sample cluster, visited each pixel site within that cluster, and identified the level II cover types.

Previous experience revealed cover type identification could be accomplished most accurately from ground observation, while composition and percent cover could be estimated most accurately using aerial photographs (Linden and Szajgin, 1981). Therefore, this combination of observations was used during this project to classify each pixel. Canopy cover percentage was usually based on ocular estimates from the 1:3,000 black-and-white photographs. When necessary, a 200 dot/in^2 random dot grid was also used to estimate cover percentage.

A stereoscopic pair of 35-mm oblique photographs was acquired on the ground at each corner of the cluster, and a photograph of each pixel area was taken from the midpoint of one of the pixel edges. Each photograph was labeled and filed for later use in verifying field notes.

ESTIMATION PROCEDURES AND RESULTS

In phase I, each pixel was interpreted to level II of the AVRI classification framework. In phase II, half of the sample clusters allocated to each stratum were visited in the field, and each pixel was classified in level II of the framework. For each cover type stratum, a stratified two-phase accuracy estimate was made using the photointerpretation data for the 160 phase I sample clusters and ground data for a subsample of 80 phase II sample clusters. These estimates were calculated using the equations shown in the appendix.

In order to use the phase II information obtained from the field, linear regression analyses were used to determine the relationships between error percentage within a cluster from photointerpretation and error percentage from field data. The paired observations were the percent of Landsat pixels incorrectly classified as determined from the photographs against the percent of pixels incorrectly classified as determined from ground data. Regression models were developed for overall, commission, and omission errors. In each case the significance of strata and cover type in accurately predicting error percentage was investigated using indicator variables (Neter and Wasserman, 1974). No indicators for strata were found to be significant at the 90-percent level, which meant that there was no significant difference between regression models for the different strata. Therefore, one model could be used for all strata. For ease of estimation and interpretability, the simple linear model on a single independent variable (percentage of photointerpreted error) was utilized for all three models. The regression models used in the estimation procedure are summarized in table 4.

AVRI ACCURACY ASSESSMENT

FIELD DATA FORM

PSU #: ______________________ Date: ______________________

1:3000 Photo #: ________________ 1:24000 Photo #: ______________

Quad Sheet: __________________ Declination: ________________

Crew Chief: ____

Crew: ____

Visited By: Helicopter __________ Ground Trans. __________

Time Arrived: ______________ Departed: ______________

Field Photos Taken: Roll # Exposure #

Digital Code	Cover Type
1	Cropland and pasture
2	Coniferous forest
3	Evergreen woodland
4	Deciduous woodland
5	Mohave desert shrub
6	Great Basin desert shrub
7	Mountain shrub
8	Plains grassland

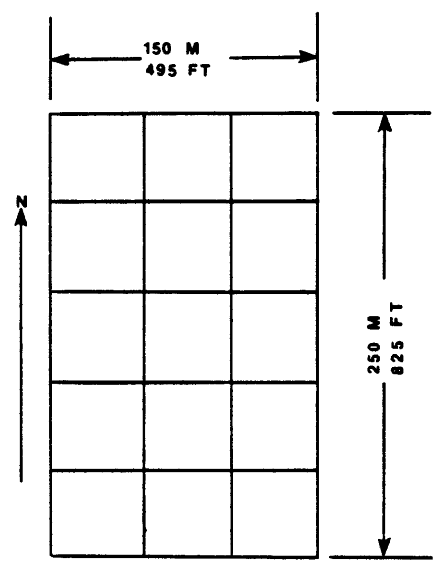


Figure 11.--Field data collection form.

Table 4.--Regression models and linear correlation coefficients (r) used to estimate commission, omission, and overall error

[See appendix for derivation of $\hat{p}_h$, the estimated commission error for cover type h]

Type of error	Regression model	r
Commission	$.111 + .786\ (\hat{p}_h)$	.79
Omission	$.310 + .463\ (\hat{p}_h)$	.42
Overall	$.088 + .832\ (\hat{p}_h)$	.86

These three regression models were used to refine the stratified estimates derived from the phase I data. The stratified mean error figure for each cover type was used in the regression model to generate a refined mean error estimate. The appropriate estimators are shown in the appendix. The final estimates are summarized in table 5.

COST SUMMARY

The total cost of the accuracy assessment was $106,950 (table 6). The most expensive single item was helicopter service ($37,500), and total labor cost was $32,000. Aerial survey and photographic costs were $23,000 which is less than 25 percent of the total project cost.

After the project was completed, costs and accuracy figures were reviewed to determine if the number of phase II ground plots might have been reduced. The optimum subsampling rate in a two-phase sample is given by:

$$\frac{n'}{n} = \sqrt{\frac{1 - r^2}{r^2} \times \frac{C_1}{C_2}}$$ (Hansen and others, 1953, p. 466)

where n' = phase II sample size (number of ground plots)

n = phase I sample size (number of photointerpreted plots)

r = correlation coefficient (from commision error estimate in this study = .79; r^2 = .62)

C_1 = cost of a phase I plot

C_2 = cost of a phase II plot

Since ground plots were approximately 10 times more expensive to survey than photointerpreted plots, $C_1/C_2 = 0.1$. Thus,

$$\frac{n'}{n} = \sqrt{\frac{1-.62}{.62}\ (.1)} = .25 = 1/4.$$

Therefore, the ratio of ground plots to photointerpreted plots could have been reduced by 50 percent.

If this type of assessment were repeated in a more operational fashion, certain costs could be reduced. If the ratio of ground plots to photointerpreted plots could be reduced by 50 percent, then helicopter costs and associated labor costs could be reduced proportionately. If local BLM field personnel could collect the field data, then travel costs could be reduced significantly. Lastly, purchasing only the necessary large-scale color aerial photographs and not the black-and-white photographs would further reduce total costs.

Table 5.--Comparison of commission and omission errors from accuracy assessment with commission errors from spectral cluster labeling data

Cover Type	Commission error from accuracy assessment data				Commission error from photointerpretation data used for spectral clustering[1]			
	Omission error		Commission error		With elevation data		Without elevation data	
	Estimated error ($\hat{p}$)	Standard error (SE)[2]	Estimated error ($\hat{p}$)	Standard error (SE)[2]	Estimated error ($\hat{p}$)	Standard error (SE)[2]	Estimated error ($\hat{p}$)	Standard error (SE)[2]
Cropland and pasture	.46	.16	.81	.07	.81	.23	.85	.17
Coniferous forest	.22	.06	.48	.03	.19	.12	.28	.07
Evergreen woodland	.46	.04	.10	.06	.19	.05	.23	.06
Deciduous woodland	.75	.03	.64	.04	.30	.23	.96	.04
Mohave desert shrub	.16	.03	.21	.04	.30	.23	.96	.06
Great Basin desert shrub	.24	.04	.43	.03	.32	.04	.43	.04
Mountain shrub	.45	.07	.66	.04	.42	.15	.62	.14
Plains grassland	.76	.14	.80	.07	.98	.01	.99	.01
OVERALL ERROR			.36[3]		.27[4]		.46[4]	
STANDARD ERROR			.02		.05		.05	

[1]Miller and Others (1980).
[2]90 percent confidence interval.
[3]From commission and omission error estimates.
[4]From commission error estimates only.

Table 6.--Summary of accuracy assessment costs

Aerial survey contracts		
Black-and-white (1:3,000)	$4,000	
Natural-color (1:2,000)	7,500	
		$11,500
Photographic processing		
Duplicate transparencies	$5,500	
Paper prints	4,500	
Miscellaneous	1,500	
		$11,500
Computer processing		
Image processing	$1,500	
Statistical analysis	2,500	
Calcomp plots	500	
		$ 4,500
Field materials		
35-mm film	$ 600	
Miscellaneous	400	
		$ 1,000
Helicopter transportation (100 h @ $375/h)		$37,500
Personnel travel (air fare and per diem)		
BLM (Denver)	$4,200	
EROS Data Center	$4,750	
		$ 8,950
Labor (14 mo @ $10/hr)		$22,400
Administration and documentation (6 mo @ $10/h)		$ 9,600
	TOTAL	$106,950

DISCUSSION

Commission errors for all eight cover types (table 5) were estimated within the prescribed 10 percent precision level, thus achieving the stated objective. Overall error, which was estimated from both commission and omission error, was 36 percent (standard error of 2 percent). Note that error estimates vary, ranging from 10 percent for evergreen woodland to 81 percent for cropland and pasture. A contingency table comparing Landsat classifications with ground classifications (table 7) shows where confusion occurred between specific cover types. Note, for example, that only a few pixels determined by ground check to be evergreen woodland were confused with other classes, while cropland and pasture pixels were often confused with other cover types.

Although the classification accuracy is variable among the different classes, BLM field personnel on the Arizona Strip District (Ramey and others, 1981) have concluded that the digital data base has immediate application in land use planning, resource management, and environmental analysis. Evaluation of the existing Unit Resource Analysis vegetation map used routinely by BLM personnel for management activities reveals that it may be only 40 percent accurate when compared to ground data collected for the accuracy assessment. It should be noted that the Unit Resource Analysis map was not used as the basis for the sampling frame of the accuracy assessment. This evaluation suggests that, although the accuracy of the Landsat classification is not high, the Landsat-derived map is more accurate than the existing Unit Resource Analysis map.

When BLM personnel produce thematic map overlays from the digital data base for such applications as pinyon-juniper chaining or blackbrush burning, they often select specific spectral classes and combine them with particular parameter values from the elevation, slope, aspect and ownership files. Hence, BLM personnel judge the utility of individual spectral classes based partially on the quantitative spectral class descriptions produced during the AVRI project. As a result, the accuracy of the level II map is not directly applicable in determining the accuracy of other thematic maps produced for a particular application.

Accuracy figures from the preliminary photointerpretation data used to label the spectral clusters (Miller and others, 1980) are also presented in table 5. Only the "with elevation data" should be compared to the accuracy assessment results. Note that overall error from the preliminary data based only on commission errors is slightly less (27 ±5 percent) than overall error from the accuracy assessment data based on both commission and omission errors (36 ±2 percent). Since the preliminary data were used for spectral cluster labeling as well as for preliminary accuracy assessment, they can be considered to be similar to training data. When training data are also used to assess classification accuracy in a digital analysis project, the results are often biased toward higher accuracy than when an independent sample is used.

Table 7.--Contingency table of ground classification against Landsat classification

[Correctly classified pixels are tallied in the boxes along the diagonal.]

LANDSAT CLASSIFICATION	GROUND CLASSIFICATION: Cropland and pasture	Coniferous forest	Evergreen woodland	Deciduous woodland	Mohave desert shrub	Great Basin desert shrub	Mountain shrub	Plains grassland	Other (barren, water)	Total
Cropland and pasture	11			42	17	2		2	46	120
Coniferous forest		71	58	3		1	17		3	153
Evergreen woodland		10	172	5		6	6			199
Deciduous woodland	4	2		28	20		2		46	102
Mohave desert shrub				27	113	11			33	184
Great Basin desert shrub			95		21	150	3	6	1	276
Mountain shrub		13	50	3		5	33		1	105
Plains grassland			20			91		7		118
Total	15	96	395	108	171	266	61	15	130	1,257

The combination of the SPS method of allocation with double sampling provided the desired estimates within the required precision levels. The SPS method provided adequate control of the cluster sample, allowing control of the number of clusters allocated to each strata and the precision of the estimate for each strata. The SPS method also guaranteed that the number of pixels of a particular cover class within any cluster would be greater than that expected using simple random sampling, as figure 6 demonstrates. Furthermore, using the SPS method, strata are defined which are highly associated with the cover types of interest. Table 2 demonstrates this point.

The overall accuracy results confirm that digital classification of vegetation is difficult to perform in semiarid environments, due largely to the sparse vegetation cover. These results are similar to those achieved by Todd and Gehring (1980) who achieved accuracy values ranging from 61.7 $\pm$ 19.2 to 97.3 $\pm$ 1.4 percent for various pinyon-juniper and shrub types in the Lake Mead National Recreation Area. As suggested by BLM personnel, possibilities for improving future classifications include (1) incorporation of other ancillary data (especially soils data) and (2) slight modification of the classification framework.

REFERENCES

Barrett, J. D., and Nutt, M. E., 1979, Survey sampling in the environmental sciences: a computer approach: Wentworth, New Hampshire, COMPress, Inc., 319 p.

Cochran, W. G., 1977, Sampling techniques: New York, John Wiley and Sons, Inc., 428 p.

Fleming, M. D., Berkebile, S. S., and Hoffer, R. M., 1975, Computer aided analysis of LANDSAT-1 MSS data: a comparison of three approaches including a "modified clustering" approach: West Lafayette, Indiana, Purdue University, Laboratory for Applications of Remote Sensing, LARS Information Note 072475, 9 p.

Hansen, M. H., Hurwitz, W. N., and Madow, W. G., 1953, Sample survey methods and theory: New York, John Wiley and Sons, Inc., v. 1, 638 p.

Linden, D. S., and Szajgin, John, 1981, Verification of land cover maps from Landsat data, in Western Regional Remote Sensing Conference, Monterey, California, 1981, Proceedings: Moffett Field, California, National Aeronautics and Space Administration, Ames Research Center, p. 1-119 to 1-125.

Mendenhall, William, Ott, Lyman, and Schaeffer, R. L., 1979, Elementary survey sampling: Belmont, California, Duxbury Press, 247 p.

Miller, W. A., Bonner, W. J., Rohde, W. G., and Schwartz, L. P., 1980, Digital Landsat and terrain data applied to an arid land resource inventory, in Arid Land Resource Inventories: Developing Costs-Efficient Methods, La Paz, Mexico, 1980, Proceedings: U.S. Department of Agriculture, U.S. Forest Service, General Technical Report WO-28, p. 589-591.

Neter, John, and Wasserman, William, 1974, Applied linear statistical models: Homewood, Illinois, Richard D. Irwin, Inc., 842 p.

Ramey, George, Moore, Ken, Walker, L. D., and Wolfe, R. C., 1981, Executive summary - Arizona vegetation resource inventory and mapping project final report: Phoenix, Arizona, Bureau of Land Management, Arizona State Office, 17 p.

Rohde, W. G., 1978, Image analysis techniques required for natural resource inventories, in National Computer Conference, Anaheim, California, 1978, Proceedings: Montvale, New Jersey, American Federation of processing Societies, p. 93-106.

Rohde, W. G., and Miller, W. A., 1981, AVRI (Arizona Vegetation Resources Inventory) final report: Sioux Falls, South Dakota, U.S. Geological Survey, EROS Data Center, EROS Office Internal Report to BLM-Denver, 200 p.

Todd, W. J., and Gehring, D. G., 1980, Vegetation and terrain mapping of Lake Mead National Recreation Area using Landsat digital data: Terre Haute, Indiana, Indiana State University, Department of Geography and Geology, Professional Paper No. 12, 24 p.

APPENDIX

ESTIMATORS FOR CLASSIFICATION ERROR BASED UPON TWO-PHASE STRATIFIED PLURALITY CLUSTER SAMPLING

Note: Estimators are used for both commission and omission errors. For commission errors, the term "cover type" refers to cover type as mapped. For omission errors, "cover type" means as classified on the ground. For overall error the cover type subscript is ignored.

---

1. Subscript definitions

h = the cover type of interest

L_h = the number of strata in which the cover type of interest occurred (was sampled)

i = the stratum where $i = 1, \ldots, L_h$

j = the cluster

---

2. Population size:

N_{hi} = total number of clusters in the i^{th} stratum that contain at least one pixel of cover type h

N_h = total number of clusters that contain at least one pixel of cover type h

$$N_h = \sum_{i=1}^{L_h} N_{hi}$$

N = total number of clusters in the population

---

3. Phase I

A. Sample size at phase I:

n_{hi} = number of clusters photo-sampled in the i^{th} stratum that contain at least one pixel of cover type h

n_h = number of clusters photo-sampled that contain at least one pixel of cover type h

$$= \sum_{i=1}^{L_h} n_{hi}$$

n = number of clusters photo-sampled

$$= \sum_{h=1}^{8} n_h$$

B. Within-cluster errors at phase I:

a_{hij} = number of pixels in the j^{th} cluster of the i^{th} stratum of cover type h which are incorrectly classified based on photo-interpretation

m_{hij} = number of pixels in the j^{th} cluster of the i^{th} stratum classified from Landsat as cover type h

x_{hij} = estimate of percent error of Landsat classification from photo interpretation in the j^{th} cluster of the i^{th} stratum for cover type h

$$= \frac{a_{hij}}{m_{hij}}$$

$\overline{X}$ = average percent error within a cluster at phase I

$$= \frac{\sum_{h=1}^{8} \sum_{i=1}^{L_h} \sum_{j=1}^{n} x_{hij}}{n}$$

C. Stratified estimate at phase I

To estimate individual class commission errors, let:

$\hat{p}_{hi}$ = the estimated commission error for cover type h within the i^{th} stratum based on photointerpretation only

$$= \frac{\sum_{j=1}^{n_{hi}} a_{hij}}{\sum_{j=1}^{n_{hi}} m_{hij}}$$

$\hat{p}_h$ = estimated commission error for cover type h based on photointerpretation only

$$\hat{p}_h = \sum_{i=1}^{8} \frac{N_{hi}}{N_h} \left(\hat{p}_{hi}\right)$$

---

4. Phase II

A. Sample size at phase II:

n'_{hi} = number of clusters ground-sampled in the i^{th} stratum that contain at least one pixel of cover type h

n'_h = number of clusters ground sampled that contain at least one pixel of cover type h

$$= \sum_{i=1}^{L_h} n'_{hi}$$

n' = number of clusters ground sampled

$$= \sum_{h=1}^{8} n'_h$$

B. Within-cluster errors at phase II:

a'_{hij} = numbers of pixels in j^{th} cluster of the i^{th} stratum of cover type h which are incorrectly classified based on ground observation

$m'_{hij} = m_{hij}$

y_{hij} = estimate of percent error of Landsat classification from ground observation in the j^{th} cluster of the i^{th} stratum for cover type h

$$= \frac{a'_{hij}}{m'_{hij}}$$

$\overline{Y}$ = average percent error within a cluster at phase II

$$\overline{Y} = \frac{\sum_{h=1}^{8} \sum_{i=1}^{L_h} \sum_{j=1}^{n'_{hi}} y_{hij}}{n'}$$

---

5. Regression coefficients:

The least squares regression coefficients are estimated by:

$$b_1 = \frac{\sum_{h=1}^{8} \sum_{i=1}^{L_h} \sum_{j=1}^{n'_{hi}} (x_{hij} - \overline{X})(y_{hij} - \overline{Y})}{\sum_{h=1}^{8} \sum_{i=1}^{L_h} \sum_{j=1}^{n'_{hi}} (x_{hij} - \overline{X})^2}$$ (Neter and Wasserman, 1974, p. 37)

$$b_0 = \overline{Y} - b_1\overline{X}$$

---

6. Double sampling estimates

A. Double sampling estimates:

$\hat{P}_h$ = double sample estimate of commission error for cover type h

$= b_0 + b_1 (\hat{p}_h)$

B. Variance of the double sampling estimates:

To estimate the variance of $\hat{p}_h$ let

S_x^2 = the variance of the estimate of percent error of Landsat classification from photointerpretation

$$S_x^2 = \frac{\sum_{h=1}^{8} \sum_{i=1}^{L_h} \sum_{j=1}^{n'_{hi}} (x_{hij} - \overline{X})^2}{n'}$$

S_y^2 = the variance of the estimate of percent error of Landsat classification from ground observation

$$S_y^2 = \frac{\sum_{h=1}^{8} \sum_{i=1}^{L_h} \sum_{j=1}^{n'_{hi}} (y_{hij} - \overline{Y})^2}{n'}$$

S_{xy} = the covariance of the estimates of percent error of Landsat classification from photointerpretation and ground observation

$$S_{xy} = \frac{\sum_{h=1}^{8} \sum_{i=1}^{L_h} \sum_{j=1}^{n'_{hi}} (x_{hij} - X)(y_{hij} - Y)}{n'}$$

The variance of $\hat{p}_h$ is then estimated by:

$$S^2_{(\hat{P}_h)} = \left[\frac{(n'-1) \left[S^2 y - \dfrac{S^2_{xy}}{S^2_x} \right]}{n' - 2} \right] \left[\frac{1}{n'} + \frac{(\overline{X} - \overline{Y})^2}{(n'-1)S^2_x} \quad \frac{(n-n')}{n} + \frac{S^2 y}{n} \right]$$

(Barrett and Nutt, 1979, p. 200)

C. Standard error of the double sampling estimates

Then the standard error of the double sampling estimate is estimated by:

$$S_{(\hat{P}_h)} = \sqrt{S^2(\hat{P}_h)}$$