NASA Reference Publication 1015

Procedures for Gathering Ground Truth Information for a Supervised Approach to a Computer-Implemented Land Cover Classification of Landsat-Acquired Multispectral Scanner Data

Armond T. Joyce

JANUARY 1978

NASA

NASA Reference Publication 1015

Procedures for Gathering Ground Truth Information for a Supervised Approach to a Computer-Implemented Land Cover Classification of Landsat-Acquired Multispectral Scanner Data

Armond T. Joyce
Lyndon B. Johnson Space Center
Houston, Texas

NASA
National Aeronautics
and Space Administration

**Scientific and Technical
Information Office**

1978

CONTENTS

Section	Page

FIGURES

PROCEDURES FOR GATHERING GROUND TRUTH INFORMATION FOR A SUPERVISED APPROACH TO A COMPUTER-IMPLEMENTED LAND COVER CLASSIFICATION OF LANDSAT-ACQUIRED MULTISPECTRAL SCANNER DATA

By Armond T. Joyce*
Lyndon B. Johnson Space Center

SUMMARY

The accurate classification of land cover in the processing of satellite-acquired, multispectral scanner data requires the proper selection and assessment of training sample sites by field personnel. This report, one of a series by the NASA Earth Resources Laboratory in connection with a natural resource inventory project in Mississippi, describes the criteria involved in the selection of training sample sites, the orientation and training of field personnel, and the ground truth data forms and procedures for their use. The basic theory involved in the computer processing of land cover data is briefly presented. Although ground truth data for this project were acquired by local field personnel, suggested options include a ground truth field team devoted exclusively to that task. Experience from the project indicates that a field agent can assemble ground truth data for approximately six training sample sites per day, with operational costs estimated at approximately $0.06 for each square kilometer for the initial classification.

INTRODUCTION

This report provides the procedures for gathering ground truth information for a "supervised" approach to a computer-implemented land cover classification of Landsat-acquired multispectral scanner (MSS) data. The NASA Earth Resources Laboratory (ERL) has drawn heavily on experience gained during an applications system verification and transfer (ASVT) project in Mississippi, but an attempt is made to address alternatives to and deviations in procedures that may be appropriate for other situations.

*NASA Earth Resources Laboratory, 1010 Gause Blvd., Slidell, La. 70458.

The first step in planning a ground truth operation is to determine and define the major land cover categories that relate to the anticipated application. To ensure that the land cover categories are compatible with the data acquisition and processing technique, those factors that influence reflected energy, as measured by the Landsat MSS, must be addressed. The variations anticipated in each category must be listed so that training sample sites can be established to represent each source of variation. Appendix A shows a typical listing, established for Mississippi. It is important that each major land cover type in the listing be specifically defined and that the training sample site criteria be established. Appendix B provides definitions and criteria for the category terms used in appendix A and throughout this report.

As an aid to the reader, where necessary the original units of measure have been converted to the equivalent value in the Système International d'Unités (SI). The SI units are written first, and the original units are written parenthetically thereafter.

BACKGROUND AND BASIC THEORY

Although it is not essential for field personnel gathering ground truth information to understand all processes in the use of Landsat digital data for resource inventory, the quality of their work will be enhanced by an understanding of the basic principles involved. Therefore, some basic theory and principles are briefly addressed in this report; references 1 to 6 provide additional details.

After the acquisition of computer-compatible tapes (CCT's) that contain the raw data acquired by the Landsat MSS, the first step in data processing during the ASVT project involved the use of a module of six computer programs, developed at the ERL, named PATREC (Pattern Recognition Analysis). The basic function of the PATREC computer programs is to effect a computer-implemented classification of each data cell[1] (which represents 0.45 hectare (1.1 acres) on the Earth's surface) for which data have been acquired. The classification process results in the identification of each of these areas as a particular land cover category (e.g., pine forest, soybean field, sand beach, etc.) that the computer has been programed to recognize.

The computer programs that form the PATREC module relate to the "supervised" technique, and the classifier algorithm is based on maximum likelihood ratio calculation and Bayesian decision rules. (See refs. 3 and 4 for additional theory and details.) The supervised technique requires that the location of a number of sites on which the land cover is known (e.g., a soybean field) be established in the Landsat data. The

[1]A data cell is also referred to as a pixel, a data element, or a resolution cell in other literature, and relates to the instantaneous field-of-view of the multispectral scanner.

area selected to contain a uniform, homogeneous land cover (e.g., a soybean field that is uniform with respect to planting date, density, vigor, etc.) are called "training sample sites" because, in a simplistic sense, they are used as references to "train" the computer to "recognize" the same land cover elsewhere. It is the office and field activities associated with establishing the true ground cover composition of these training sample sites that are encompassed by the ground truth information gathering operation.

FACTORS IN SELECTING TRAINING SAMPLE SITES

A ground truth operations plan must address four basic factors in the selection of training sample sites: the categorization of land surface features, the size and shape of training sites, the number and distribution of sites, and the homogeneity and uniformity of the land cover.

Categorization of Land Surface Features

There are three general categories of land surface features that significantly affect the reflected energy measured by the MSS. The features are those that relate to the vegetation cover, those that apply to land surfaces that do not support vegetation of any significance, and those that pertain to the topography of the land surface.

Vegetation cover.- Various elements of the vegetation cover influence the radiation measured by the MSS. These include plant species or species association, plant age and vigor, plant density, and understory vegetation.

Plant species or species association: There are many characteristics unique to each plant species that affect the intensity and wavelength of reflected energy. Some of these are the size of the plant cells, the thickness of cell walls and intercellular air spaces, the leaf arrangement on the stem, the pigments present, the thickness and shape of the leaf.

Some vegetation types such as agricultural crops, planted grasses, some forest plantations, and orchards are likely to consist of a single species; however, naturally occurring vegetation is usually a mixture of various species. Consequently, natural vegetation cover types are defined and named for the predominant species. For example, a forest may be termed a pine forest if 75 percent or more of the surface area of the tree crowns in the upper canopy were of pine trees. Therefore, a training sample site selected to represent a pine forest, as defined in this example, could include some hardwood trees provided they were uniformly intermingled with the pine trees and their crowns did not cover more than 25 percent of the total surface area.

It is also possible to define and name a vegetation cover type consisting of two or three species that grow in association with one another but which together are predominant. For example, a forest cover may be called oak-hickory if oak and hickory grew together in an intermingled manner and together constituted 75 percent or more of the surface area covered by tree crowns. Therefore, a training sample site established to represent the oak-hickory cover could include other species but should meet the criteria and be uniform with respect to how the various species were intermingled. Although forest vegetation was referred to in the previous examples, the same statements concerning vegetation-type criteria apply to marsh (nonforested wetlands) vegetation and brushland (multistemmed, woody shrubs) vegetation.

Plant age and vigor: Although a given plant species is likely to be uniquely different in plant characteristics from another plant species, it is possible that some of those characteristics can change with plant age or plant vigor. For example, a young, vigorously growing plant may have less leaf water content or may have cells with thinner cell walls than an older, slow-growing plant of the same species. Consequently, the plant's age/vigor is the second variable that must be addressed, especially in the case of perennial vegetation. Of these two parameters, vigor (the rate at which the plant is growing) is the more important and is referred to in conjunction with age only because of the general correlation between vigor and age.

In forested areas, there is likely to be a gradient from very young, vigorously growing forest stands to mature, stagnant, or even decadent forest stands as well as some all-aged stands. It is important that training sample sites be established to encompass all age/vigor variations in each vegetation type to be addressed; however, for practical purposes, it is recommended that the age/vigor categories be fairly broad. For example, such categories may include the following.

1. Young forest stands that are on good sites with respect to soil and rainfall and, therefore, are growing at a fast rate, usually characterized by the profuse flushes of terminal branch growth during the spring

2. Forest stands that have a moderate rate of growth because of site conditions and/or age

3. Forest stands in which growth has slowed down appreciably because the trees are mature, near maturity, stagnated, or on poor site conditions

As an example, in the case of Mississippi pine forests, the first category is likely to include plantations that are generally in the 1- to 10-year age bracket and are not yet of commercial size; the second category would consist mainly of pulpwood-size trees, generally in the 11- to 30-year age bracket, but possibly would include some natural regeneration on poorer sites that was not yet of commercial size; and the last category would include all other pine forest. In all three cases, the forest stand encompassed by each training sample site should be uniform in the criteria established for age and vigor and for vegetation cover.

In the case of annual vegetation, such as an agricultural crop for which all fields in a given region are likely to have planted within a 2- to 3-week time span, the difference in age is not strongly correlated with vigor. Therefore, for agricultural crops, vigor is addressed with respect to whether a particular crop is growing at the normal rate as opposed to growing under a stress caused by insect/disease infestations, inadequate moisture, or a lack of nutrients in the soil. If, at a particular time for which a classification is to be made, stress conditions are known to exist, training sample sites for a particular crop should be established both in fields in which plant growth is normal and in fields that are under stress. Because of likely differences in vigor, training sample sites for the same crop should be established for both irrigated and nonirrigated conditions, if both irrigation and dryland practices are intermingled for a given crop. Also, if there are other differences in practices, such as a particular grass species or species mix being grazed in one case and grown for hay in another case, separate training sample sites should be established for each.

Plant density: A span of several weeks in the planting period for a particular agricultural crop could cause variation in plant density, the third principal variable in vegetative cover. Plant density is an important factor because the MSS is taking a measurement for a 0.45-hectare (1.1 acre) area. Consequently, if, in the case of row crops, both plants and exposed soil between the rows are visible from above, the measurement involves an integration of the energy reflected from plants and from exposed soil. If a wide difference in planting dates for a particular crop indicates that a significant variation in density could exist, then training sample sites should be established for each density category. For practical purposes, such density categories should be fairly broad; e.g., 40 to 60 percent, 60 to 80 percent, and 80 to 100 percent. If the crop vegetation covers less than 40 percent of the surface, it is not likely that the specific crop could be identified.

Vegetation density is also a factor in forest and brush vegetation, but it is recommended that the variation in density be addressed with two categories rather than three, as recommended for agricultural crops. For example, training sample sites established to address density variation in a pine forest may be categorized as sparse (20 to 65 percent of the surface covered by tree crowns) or dense (65 to 100 percent of the surface covered by tree crowns).

Understory vegetation: The previous example leads to the fourth factor that must be considered in the major variations in vegetation that influence reflected energy, that of the understory vegetation. In the example of a sparse pine forest, there may be one condition in which a native grass would be visible from above through the gaps between the scattered trees; in another condition, the understory may be a brush species or species association rather than grass. In this example, it would be desirable to establish a training sample site for each of the two conditions. For Landsat data acquired during the winter season, when deciduous forests are leafless, it is also desirable to establish a training sample site for dense, deciduous forest that has an evergreen understory species and another site for dense, deciduous forest that has a deciduous understory species. Separate training sample sites would also be appropriate for a leafless deciduous forest flooded with water and one without flooding.

Nonvegetated land cover.- Land surfaces that are essentially devoid of vegetation are those on which soil has been temporarily exposed and those of inert materials that do not support any, or support very little, vegetation.

Temporarily exposed soil: During the spring, most cultivated areas are in some stage of soil preparation. Consequently, criteria should be established for the condition of the exposed soil rather than for the anticipated crop. The three main variables to consider are the physical state of the surface, soil moisture, and soil type. As a minimum, training sample sites should represent the extremes, should they exist, and the various combinations of these three variables. For example, the extremes for the state of the surface would be a rough surface that may have resulted after plowing and a smooth surface that may have resulted after harrowing and/or planting. Soil moisture criteria should reflect the dry conditions of some fields and the very wet or waterlogged conditions that may exist in other fields. Soil type extremes could range from the light-colored, sandy soils of some fields to dark-colored, clay soils of other fields. In the case of land used for cultivated crops, there may also be ground cover conditions other than green, growing vegetation or exposed soil that must be addressed at certain times of the year. One common condition is the stubble resulting from harvesting operations that may be widespread in early fall. Even though there should be no significant difference in energy reflected from dead stalks and debris from different crops, training samples should be established to represent various possible stubble conditions. For example, sites having stubble left after corn has been cut for silage (in which there is a low volume of stalk material left and considerable bare soil exposed) should be established separately from sites having stubble left from the harvesting of small grain.

Inert materials: The general category of inert materials includes beaches, sand bars, mud flats, rock outcropping, extractive areas (e.g., gravel pits), asphalt, concrete, etc. Except for the topographical configuration, there is little variation within each of these land cover types; however, their basic characteristics may relate to different degrees of reflectivity in the four MSS-measured wavelengths. For example, concrete is highly reflective, whereas asphalt has low reflectivity. Consequently, training sample sites should be established to represent each of these inert materials that may be present. In some cases, inert materials may not exist in a pure enough form over an area large enough to serve as a training sample site, and, therefore, a site may be established to represent a particular complex. For example, a site that contains a heterogeneous mixture (concrete streets, gravel parking lots, metal roofs, etc.) would be appropriate in an area where these materials only exist in such a mixture. In the urban environment, some training sample sites may be termed "high density" to reflect a criteria requiring a pure or mixed form of inert materials with no vegetation intermingled; whereas, others may be termed "low density" to reflect a criteria permitting up to 35 percent of the total surface to be covered by isolated patches of vegetation (no larger than 31 meters (100 feet) in maximum dimension) with the remainder encompassed by pure or mixed forms of inert materials. High density may typify large urban commercial centers or industrial sites; low density may typify suburban residential areas with scattered trees partially overtopping the streets and houses.

Topography.- The topography (slope and aspect) is also a factor in establishing uniform, homogeneous training sample sites if there is pronounced topographical variation. It is recommended that slope categories be established for 0 to 10 percent slopes, 10 to 30 percent slopes, 30 to 50 percent slopes, and 50 percent or greater slopes. Aspect is usually not considered in the training sample site criteria unless slopes are 30 percent or greater, in which case aspect is categorized according to the four cardinal directions. Most steep slope conditions are in forested areas and are likely to be automatically categorized as to aspect in the course of applying criteria for defining the species or species association cover type. For example, in the western United States, a pronounced slope with a south aspect may support ponderosa pine, whereas a pronounced slope with a north aspect may support larch-Douglas fir.

Size and Shape of Training Sites

Size of site.- There are two principal factors that relate to the size of the training sample site. One factor concerns the facility with which a training sample site can be located in the land surface image displayed on a cathode ray tube (CRT). The other factor relates to the number of data cells (0.45-hectare (1.1 acre) areas) required to develop valid statistics from the MSS measurements.

Experience at ERL indicates that it is most desirable to establish training sample sites that are approximately 16 hectares (40 acres) in size. A site of this size can usually be located in the land surface image displayed on the CRT without difficulty and will encompass around 30 data cells, thereby providing a sample large enough to develop valid statistics. It is not recommended that training sample sites for natural vegetation be larger than approximately 65 hectares (160 acres) because of the difficulty in finding a site of that size that does not violate the criteria for uniformity. Conversely, the smaller the site, the higher the probability that it cannot be located in the CRT display. Also, the smaller the site, the lower the efficiency in developing valid statistics. For example, two 8-hectare (20 acre) sites with 15 data cells each could eventually be grouped in data processing to equal the 30 data cells encompassed by one 16-hectare (40 acre) site, but this would require twice the effort in field work and data analysis. Consequently, it is not recommended that training sample sites smaller than 16 hectares (40 acres) be established unless the particular land cover type in question does not exist except on areas smaller than that. In any event, it is recommended that 4 hectares (10 acres) be the absolute minimum size for reasons of efficiency, statistical validity, and probability of locational accuracy in the Landsat data.

In view of this recommended restriction, it can be seen that land cover types that only occur on areas smaller than 4 hectares (10 acres) should be precluded from the list of land cover types that can be addressed with Landsat data. However, once valid statistics have been derived for a training sample and the spectral signature has been developed for a particular land cover type within a Landsat scene, the classification will be performed for each individual 0.45-hectare (1.1 acre) data cell in the Landsat scene (provided that cloud-free conditions permit processing the four corresponding CCT's as a data set). Therefore, even though there may be some large areas in the 185- by 185-kilometer (115 by 115 statute mile) scene within which a particular land cover type only occurs in units between 0.45 and 4 hectares (1.1 and 10 acres) in size, these units may be classified accurately through use of training sample sites of adequate size found elsewhere in the Landsat scene.

Shape of site.- The shape of training sample sites is not crucial; however, in some cases, a square or rectangular training sample site may prove easier to locate in the Landsat data display. As will be further elaborated later in this report, locating sites in the CRT display of the land surface is often facilitated by visually projecting lines from prominent, easily identified surface features to two or more of the sides of a square or rectangular site.

Number and Distribution of Training Sites

From a theoretical point of view, only one training sample is needed to develop statistics and, subsequently, to perform a computer-implemented classification of the land cover feature that the particular training sample site was established to represent, provided that the training sample statistically represents the land cover type to be classified. However, for several reasons, it is recommended that an attempt be made to establish at least three training sample sites for each land cover feature.

First, it is possible that a training sample site could be lost either because its location cannot be established in the data or, infrequently, because its location coincides with scan line dropout (an electronic or transmission failure during which no measurements are recorded for all or part of the data cells on a particular scan line).

Second, it may be necessary to discard a training sample because the statistics, once derived from the Landsat data, indicate that it is not a uniform, homogeneous land cover type from a spectral viewpoint. Such statistics may have resulted from human error during training sample site location either as delineated on field maps or located in the Landsat data, or they may be due to the basic nature of the particular land cover type. In any event, unless the problem can be corrected, the training sample should be discarded. Consequently, if only one training sample site had been established, data processing would have been interrupted by the need to redefine the boundary or by another field trip to establish a new site.

Third, the analysis of the statistics is easier if the statistics from three or more sites established to represent the same land cover condition can be compared. For example, if the mean and the standard deviation calculated for one training sample are significantly different from those of the others, the analyst may discard that sample because those remaining better represent the particular land cover condition or he may carry that sample as a separate spectral subclass. However, if the analyst were dealing with only one training sample, he would have no basis of comparison and would have to accept the one sample as being representative; or, if he were dealing with two for which the statistics were substantially different, he would not know which of the two was more representative.

Finally, even though the statistics from three or more training samples established to represent the same land cover condition should not be substantially different, the means and standard deviations are usually not exactly the same. Consequently, the analyst may wish to group the three or more samples to create new statistics to develop a spectral signature that encompasses more variation in the particular land cover type. This approach is illustrated in figure 1, in which the three dashed-line ellipses represent statistically defined areas that encompass the measurements from all the data cells in each sample as they cluster around the means in the center of each ellipse. The solid-line ellipse constructed around the three dashed-line ellipses represents a hypothetical situation resulting from a grouping of the statistics of the three individual samples.[2] Consequently, if the measurements for an unknown data cell fixed its location in the shaded area of figure 1 during classification, that data cell would be classified as pertaining to the particular land cover type; whereas, it would have been left uncategorized had each of the three training samples been carried as separate classes. In this hypothetical case, it can be seen that three grouped training samples would have resulted in a more accurate classification than one or two samples either held separately or grouped, if the grouped statistics more correctly estimated the true statistical population for the ground cover condition.

A ground truth operation is usually oriented to a particular Landsat scene, which measures 185 by 185 kilometers (115 by 115 statute miles), an area covering approximately 3.4 million hectares (8.5 million acres). The number of training sample sites needed within a particular scene varies with the number of land cover types to be classified within the scene and with the variations within each land cover type. As indicated by the listing in appendix A, there may be up to 11 major land cover categories in a State as large and varied as Mississippi, and it may be necessary to establish 80 or more training samples in order to address all variable conditions within these categories for all seasons of the year. However, ERL experience indicates that, within the area

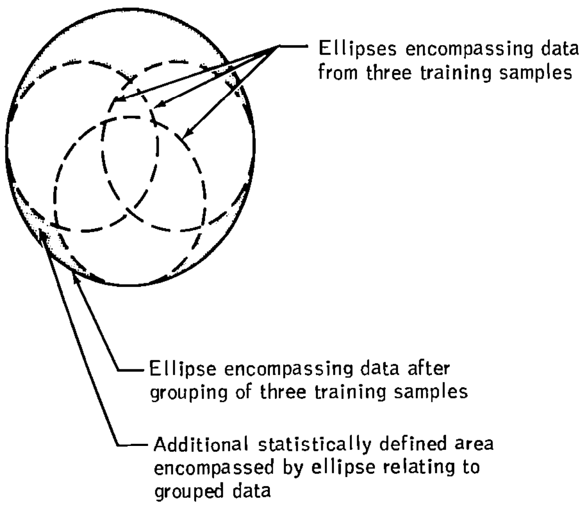


Figure 1.- Statistical grouping of data from three training sites to develop statistics that encompass more variation in a land cover condition.

---

[2]The particular drawing of the solid-line ellipse in figure 1 was designed to illustrate the concept of grouping data and should not be taken to imply that the ellipse resulting from the grouping is always tangent to the individual sample boundaries. In fact, its position will depend on the confidence interval defined.

encompassed by a particular Landsat scene during the season for which ground truth is being gathered, there are likely to be 8 to 10 major land cover types for which around 30 to 50 training sample sites must be established to address the various land cover conditions. Consequently, since three or more training sample sites should be established for each land cover condition, 90 to 150 training sample sites may be established in the 3.4-million-hectare (8.5 million acre) area encompassed by each Landsat scene. Using the upper extreme of this example and assuming that each training sample site is 16 hectares (40 acres) in size, it is evident that the area required for the 150 training sample sites amounts to less than one-thousandth of the 3.4 million hectares (8.5 million acres).

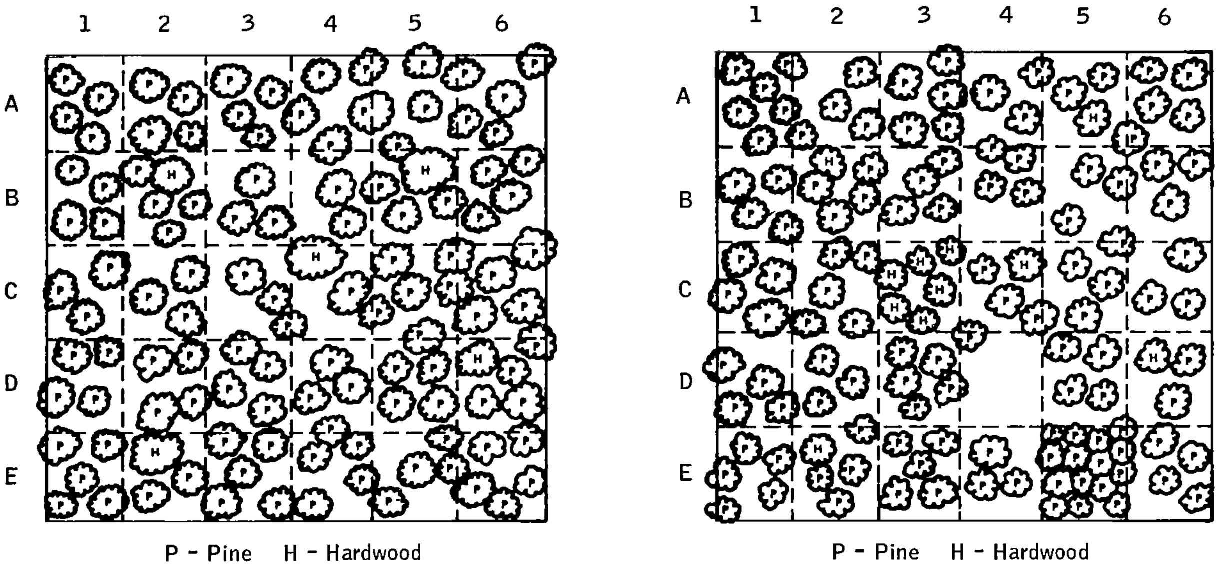
One set of training sample sites would normally be established for the area encompassed by each Landsat scene because the four CCT's relating to a scene are usually processed as a set. However, through an approach referred to as signature extension, it may be possible to process more than the four CCT's from a Landsat scene as a set, thereby reducing the number of training sample sites per scene. This possibility could arise if two or three cloud-free scenes (8 to 12 CCT's) of data have been acquired on a particular pass under fairly uniform atmospheric conditions. This situation is most often encountered when the passage of a strong cold weather front precedes a Landsat pass by 1 or 2 days. However, it is recommended that ground truth gathering activities be planned for the area encompassed by each scene, and that the concept of signature extension be considered only in respect to data processing efficiency after Landsat-acquired data have been assessed for quality.

Homogeneity and Uniformity of Training Sites

The fundamental requirement of the computer programs used to perform a land cover classification is that the statistics derived from the MSS data conform to a normal distribution. These statistical parameters are used to establish an elliptically shaped decision boundary, which, in turn, is based on a normal distribution. If statistics that do not reflect a normal distribution are used to define the decision boundary, the classification will be degraded. Therefore, it is necessary for a training sample site to reflect a uniform and homogeneous vegetation/land cover condition. The uniformity/homogeneity specification is made for those vegetation/land cover variables that influence the reflected and/or radiant energy being measured by the MSS.

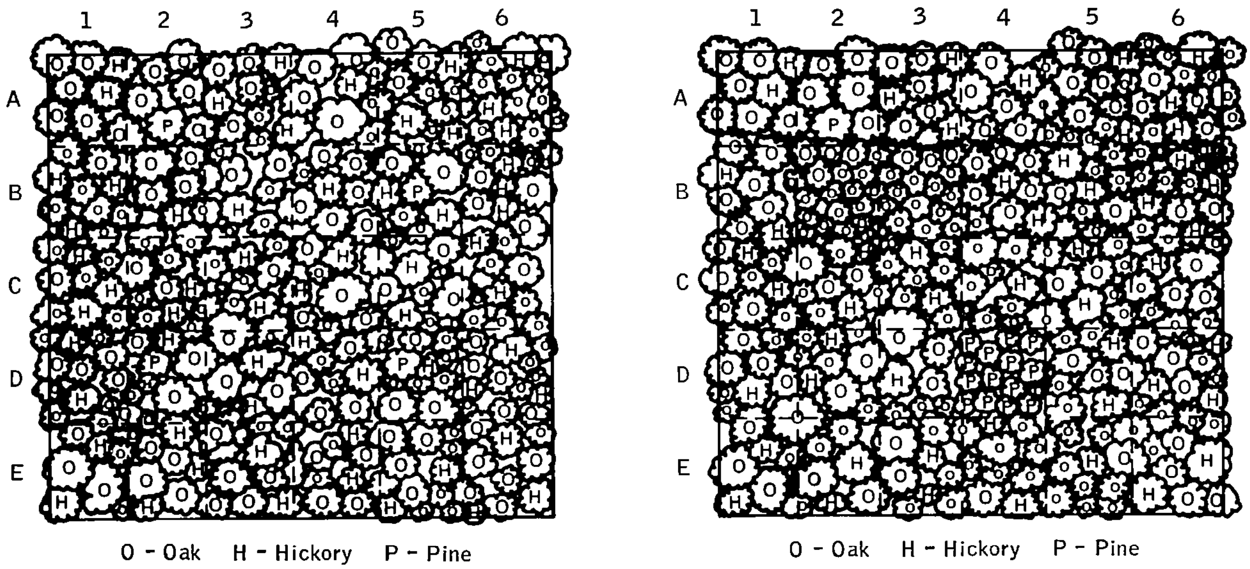
In figures 2 and 3, the large squares with solid lines represent training sample sites of approximately 16 hectares (40 acres). Within each large square, the dashed lines form small rectangular areas that represent the 0.45-hectare (1.1 acre) cells for which the MSS takes a measurement. The circles represent areas of tree crowns as viewed by the MSS.

Figure 2 shows two training sample sites established to represent a sparse (20 to 65 percent crown coverage) pine forest (90 percent or more pine) with a native grass ground cover apparent in the gaps between the trees. The area shown in figure 2(b) is inadequate as a training site for a uniform, homogeneous, sparse pine forest for several reasons. In cell 3C,



(a) Adequate site. (b) Inadequate site.

Figure 2.- Illustration of training sample site established to represent a sparse pine forest.



(a) Adequate site. (b) Inadequate site.

Figure 3.- Illustration of training sample site established to represent a dense oak-hickory forest.

a concentration of hardwood trees encompasses the entire cell; in cell 4D, the gap between trees is so large that the cell would reflect only the grass between the trees; and in cell 5E, the density of the pine trees is such that it exceeds the criteria for a sparse condition. Conversely, the site shown in figure 2(a) is adequate in reflecting a uniform, homogeneous condition. Pine trees are scattered throughout the site in such a manner that a sufficient number fall in each cell without exceeding the criteria for a sparse forest. Even though a few hardwood trees are present, they do not occur in concentrations, they do not exceed 10 percent of the area covered by tree crowns, and there are no large gaps in the canopy.

Figure 3 shows two training sample sites established to represent a dense oak-hickory forest (90 percent or more oak-hickory). The site in figure 3(b) is inadequate as a uniform, homogeneous site for two reasons. First, there is a concentration of pure oak in cells 2B and 3B; and, second, there is a concentration of pine in cell 4D. The site in figure 3(a) is an adequate representation of a uniform, homogeneous condition in that both oak and hickory trees occur in each cell in roughly the same proportion; and, even though a few pine trees are present, they are not concentrated and do not exceed 10 percent of the total area covered by tree crowns. With these examples, it is apparent that the main requirement for uniformity is that there be no condition within an area as large as 0.45 hectare (1.1 acres) that differs substantially from the criteria that defines the land cover condition to be represented by the training sample site. Although figures 2 and 3 used forest vegetation as an example, the same criteria should be applied to other types of vegetation. For example, in a marsh (nonforested wetlands) or a pasture grass characterized as a species association with two or more intermingled species, none of the species should occur singly over an area as large as a cell. It is also recommended that a species occupying less than 25 percent of the area not be included in the name of multispecies associations; therefore, the species would not be considered in applying uniformity criteria.

Agricultural crops are usually single species, but conditions of density and/or vigor may have a bearing on uniformity and homogeneity. For example, areas 0.4 hectare (1 acre) or larger in size having bare soil resulting from germination failure or having differences in vigor as a result of uneven fertilization or poor nutrient availability should not be permitted in a training sample site established to represent an otherwise healthy crop.

Topographic features should also be uniform in respect to broad categories, as suggested previously in this report. For example, in mountainous terrain, a training sample site should not be established so that part is on a north aspect slope and part is on a south aspect slope if such slopes are greater than 30 percent. Also, slopes should not exceed the limits that have been defined for a slope category (e.g., 30 to 50 percent).

OPERATIONS BEFORE IMPLEMENTING FIELD PROCEDURES

A number of operations can be initiated prior to the actual field activities that will lend efficiency to the process of gathering ground truth data. These include the preselection of training sites by use of aerial photographs, notation of prominent surface features for location of sites, concentration of sites wherever possible, acquisition of aerial photography, the use of prints or maps, and the assembling of ground truth packages.

Site Preselection By Use of Aerial Photography

After the major land cover categories have been determined and the variations in conditions have been listed, the next step consists of using available aerial photography to preselect training sites according to predefined criteria for each land cover condition. Although preselecting training sample sites through aerial photointerpretation is not essential, it can be used to gain efficiency in ground truth activities. Any type of available aerial photography is adequate if it is not too old (within the last 5 years, under most conditions of land use change). However, ERL experience shows that preselection is most efficient with color infrared positive transparencies in roll form at scales of 1:60 000 to 1:120 000 when interpreted under magnification. This efficiency is gained because large areas can be viewed on a single frame, the resolution at these scales is compatible with locating the training sample in Landsat data, and the logistical planning of field work is often facilitated when using one print that shows the road network for a large area. However, scales in the range of 1:15 000 to 1:30 000 can also be used with considerable efficiency and in some cases can have certain other advantages in respect to the kind of detail that can be photointerpreted.

In general, during a training sample site preselection process, the photointerpreter should not strive to photointerpret all details on which ground truth information is desired. For example, he may find and delineate potential training sample sites that meet the criteria for pine forest, hardwood forest, marsh, and brushland, but stop short of identifying the particular species or species association. He may also delineate potential training sample sites for cultivated areas or grassland that appear to be uniform, leaving the ultimate categorization to the field team. In essence, the photointerpreter strives to add efficiency by preselecting training sample sites that meet the general criteria so that field personnel (may also include the photointerpreter) can go directly to these preselected sites as opposed to canvassing the entire area in search of adequate training sample sites. Also, even if field personnel reject some of the preselected sites and establish substitute sites while in the field, the overall operation is usually less time-consuming than it would have been had sites not been preselected. However, depending on the type, scale, and season of acquisition of available aerial photography, the photointerpreter may deal with certain variables of land cover more precisely than they can be dealt with on the ground. For example, color infrared positive transparencies acquired during the winter season at scales of 1:30 000 or larger can be used to determine density (crown closure) categories in pine forest and/or the degree of the overstory mix between pines and leafless hardwoods as precisely and with much less effort than can be determined on the ground. Broad slope and aspect categories can also be efficiently determined through stereovision interpretation of forward overlapping photography. If the photointerpreter is not familiar with general land cover/vegetation types within the area of concern, a review of publications that give statistical information by county (such as those published by the U.S. Soil Conservation Service, the U.S. Forest Service, and the U.S. Crop Reporting Service) can be helpful.

As potential training sample sites are located through photointerpretation, the boundaries of these preselected sites are usually delineated on transparent material overlaid on the original film for office records and on prints made for use in the field. If the original photography was color infrared or color, black-and-white prints are usually made for field use; and, if the original scale is smaller than 1:63 360 (1 centimeter = 0.63 kilometer (1 inch = 1 statute mile)), a print enlarged to approximately that scale can be made to facilitate use in the field. As potential training sample sites are preselected and delineated, the photointerpreter writes a unique four- to six-digit letter/number identifier on the print and overlay adjacent to the delineation of the site. This unique letter/number identifier, to be explained later in this report, is used both for a cross reference to ground truth forms and for identification of the site during computer processing.

Site Referencing By Use of Prominent Surface Features

Another means of increasing overall efficiency by preselecting training sample sites through aerial photointerpretation is to delineate potential sites in such a manner that their locations are referenced to prominent surface features easily found in the field and detectable in the CRT display. This concept is illustrated in figure 4, in which the potential training sample site was so delineated that one side can be located by visually projecting from a road junction and another side can be located by visually projecting from a bend in a river. If such linear features are 15 meters (50 feet) or wider, they can almost always be used, subsequently, to locate the training sample site more easily in the Landsat image display. An even more effective use of this concept is to delineate potential training sample sites so that one or more sides are adjacent and parallel to straight-line interfaces between two different land cover types (e.g., forest and cropland), or, in the case of cropland and grassland, to delineate training sample sites in fields with one or more sides parallel and immediately adjacent to prominent roads. It is also often possible to project lines from the centers of two or more prominent nonlinear features such as small water bodies or built-up areas as references to the sides of potential training sample sites. Experience at ERL has shown that, if sufficient attention is given to this concept during preselection of sites and/or in field establishment of sites, very few if any training sample sites are "lost" because their locations cannot be ascertained in the image display of Landsat data.

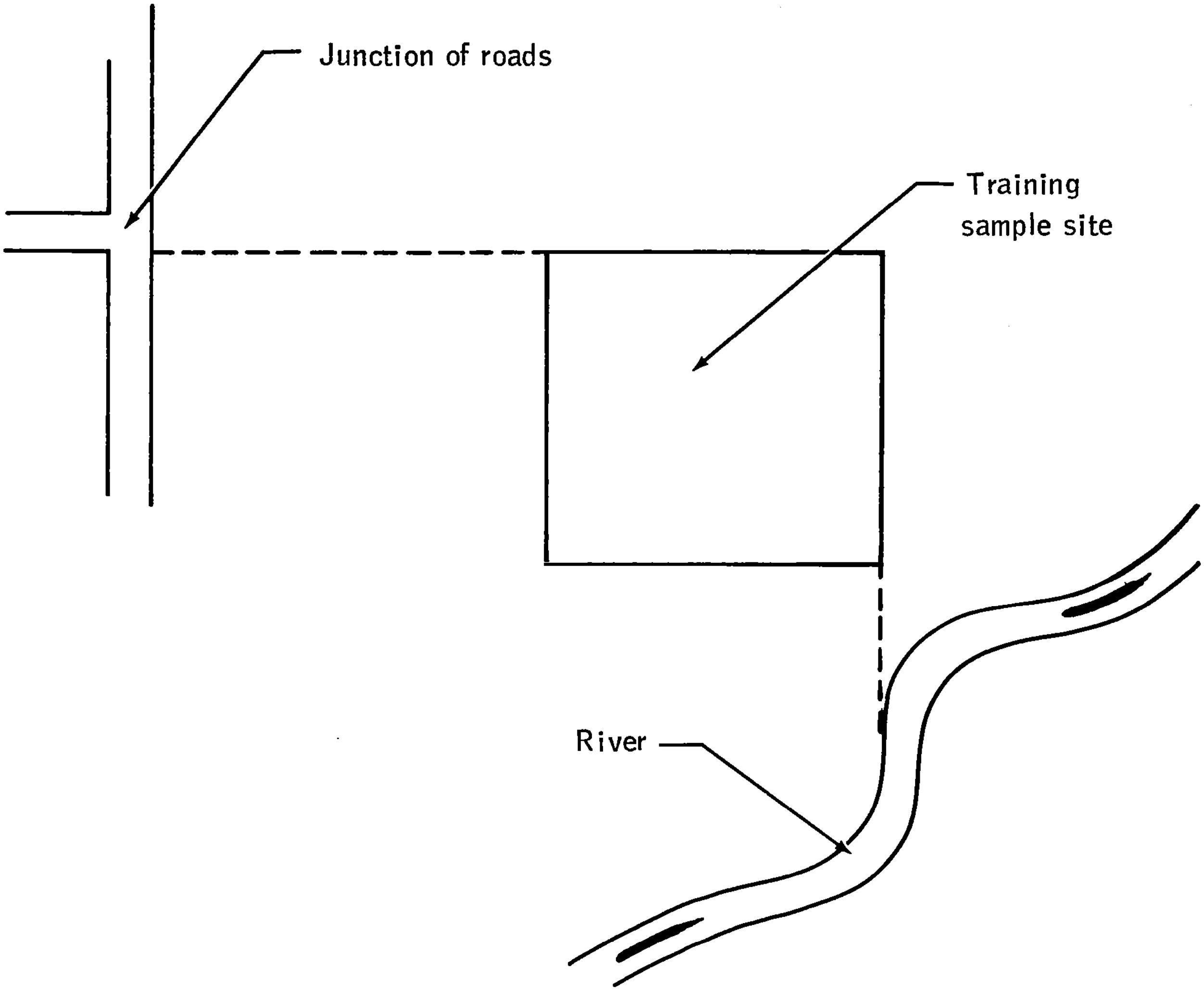


Figure 4.- Delineating training sample site by referencing location to prominent land surface features.

During preselection of potential training samples, the photointerpreter should also observe the road network and, whenever possible, locate potential training sample sites so as to facilitate access and take best advantage of road networks. Site selection of this sort will lessen field work by reducing time spent in walking and in backtracking vehicle routes.

Concentration of Sites

A final means of gaining efficiency consists of establishing potential training sample sites in concentrated groups distributed throughout the area encompassed by a particular Landsat scene of interest. This can be accomplished by having the photointerpreter begin by selecting 8 to 12 aerial photographs, depending on scale, from all the photography available for the area encompassed by a particular Landsat scene. This selection can be made so that each photograph encompasses a variety of land cover types or in such a manner that each photograph focuses on a particular land cover type, depending on whether field teams are organized along multidisciplinary or disciplinary lines. In either case, it is desirable that each photograph covering an area of concentrated training sample sites fall completely within the area covered by one of the four CCT's in order to preclude the establishment of training sample sites at the abutment of tapes. Also, if field personnel are organized relative to political or management units (e.g., a county forester), it is also desirable that all or most of the photographs fall within that particular unit.

If the photointerpreter does not encounter a sufficient number of potential training sample sites that meet the predetermined criteria for each land cover condition with the original selection of photographs, he can select additional photographs in areas that optimize the chances of finding sites relating to those land cover conditions lacking after the first iteration. The net effect of delineating potential training sample sites in concentrated groups distributed throughout the Landsat scene is the reduction of travel time between sites during field operations and, subsequently, the reduction of time required to locate training sample sites in the image display of Landsat data.

Acquisition of Aerial Photographs

At the present time, aerial photography that has been taken within the last 5 years and that is of a type and scale suitable for training sample site preselection is available for most of the United States. If not already in the possession of agencies planning Landsat ground truth information gathering operations, the existence and coverage of aerial photography acquired through various Federal programs can be ascertained through the Earth Resources Observation Systems (EROS) Data Center at Sioux Falls, S. Dak., operated by the U.S. Geological Survey. Landsat coverage for a particular area defined by latitude and longitude can be verified through and purchased from the EROS Data Center. If recent aerial photography is not available for a particular area of interest, it may be cost-effective to acquire a limited amount of new aerial photography, especially in forest or marsh areas with poor accessibility. If the acquisition of new aerial photography is carefully planned, it should be possible to acquire sufficient aerial photography to gain cost-efficiency in ground truth information gathering operations by covering no more than 2 percent of the area with aerial photography (e.g., 36 frames or 12 sets of 23-centimeter (9 inch) format stereo triplets at 1:24 000 scale per Landsat scene). If the aerial photographs are to be used for ground truth information gathering activities for specialized (e.g., pine as compared to hardwood stratifications for forest inventory) rather than composite classifications, limited color infrared photography during the winter season at a time close to a cloud-free Landsat pass could completely eliminate the need and cost of field activities.

Use of Prints or Maps

Even if aerial photography is not used for preselection of potential training sample sites, it is desirable to provide aerial photoprints to field personnel to be used as a map base on which to delineate the field-established training sample sites. In case of the ASVT project to which this report relates, preselection of potential training sample sites through photointerpretation was accomplished within seven counties for which specific applications were demonstrated, but ground truth information gathering operations were conducted without preselection in the remaining 75 counties in the State of Mississippi. However, field personnel were provided with either aerial photoprints or photobase maps with broad land-cover-type delineations. The original photography was 1:120 000-scale color infrared, but the prints were reproduced in black and white to reduce cost and then were enlarged to 1:60 000 to facilitate use in the field. A print of one photograph (16- by 16-kilometer (10 by 10 mile) effective area) was provided for each of 65 counties, and a township-size (9.7- by 9.7-kilometer (6 by 6 mile)) photobase map with land cover types delineated was provided for 10 counties. The net effect was to concentrate established training sample sites in either 16- by 16-kilometer (10 by 10 mile) or 9.7- by 9.7-kilometer (6 by 6 mile) areas within each county.

Inasmuch as the training sample sites were established by field personnel assigned to each respective county (e.g., County Extension agents and county foresters) or a management unit within a county (e.g., State park or game management unit), all field personnel were very familiar with their respective area of responsibility. In some cases, the field personnel were so familiar with their areas that, after orientation by use of the photoprint, they could delineate satisfactory training sample sites, each representing some specific vegetation/land cover condition in their area, and fill out corresponding ground truth forms without leaving their offices. It is mainly in this situation, in which field personnel are very familiar with a localized area, that training sample sites can be established efficiently without preselection through photointerpretation.

Ground Truth Packages

Prior to field implementation of a ground truth information gathering operation, ground truth packages should be assembled for field personnel. In the case of this ASVT project, the packages prepared for disciplinary personnel located in each county consisted of the following.

1. An aerial photograph or a photograph-based land cover map

2. A county map that shows the outline of the area encompassed by the aerial photograph

3. Applicable blank ground truth forms (appendix C)

4. An instruction sheet

5. A sheet defining letter symbols for each vegetation/land cover condition that could be characterized by a four-digit combination of letters

To show an example of letter symbols, a pine forest that was old (more than 50 years) and sparse was characterized as PFOS. Field personnel were asked to write the appropriate letter symbols adjacent to each training sample site delineated on the photoprint. Instructions also called for a unique two-digit number to be added to the four letter characters as each training sample site was established and delineated on the photoprint. For example, 09 added to PFOS would mean that PFOS09 was the ninth training sample established. This six-digit identifier was also used on the ground truth form (which contained additional information) corresponding to a particular training sample site as a cross reference. In the case of the seven counties for which preselection of sites through photointerpretation was conducted, the photographs taken to the field had from one to four letter symbols, depending on the degree to which photointerpretation of land cover conditions was possible, together with a unique two-digit number recorded adjacent to the delineation of the potential training sample site.

When preselection of sites is made through photointerpretation and when the field team is not very familiar with the local area, it is recommended that the appropriate location of each preselected site be plotted with an "X" on a small-scale map such as a 1:250.000 topographic quad map. This map can then be used to assign field teams to specific areas outlined on the map with regard to site locations, the road network, and lodging facilities.

IMPLEMENTING FIELD PROCEDURES

Organization of Field Personnel

Although there are various ways to organize a field team for gathering ground truth information, this section will focus only on what are considered to be two basic options. The first option is to organize the effort around field personnel employed by those agencies that are the anticipated users of the land cover classification. This option involves an organized effort in which each individual is responsible for establishing training sample sites within his local area for his area of specialty. For example, a county forester would only establish training sample sites to represent the various forest vegetation conditions within the county to which he is assigned. With this form of organization, each individual involved would use only a fraction of his time for the establishment and visitation of training sample sites and most work could be conducted in the course of carrying out routine activities as opposed to initiating a separate effort.

The other option consists of employing a team representing several disciplines (e.g., a forester, an agronomist, and a botanist) whose primary responsibility would be to gather ground truth information. The possibility of this option is envisioned in a situation in which it would be feasible to make such a team part of the staff of a center that processes remotely sensed data. As will be explained later in this report, it is believed that three disciplinary personnel could furnish ground truth information, for a State as large as Mississippi, using 80 percent of their time, leaving 20 percent for performing certain steps during data processing and for interfacing with disciplinary personnel from user agencies.

There are advantages and disadvantages associated with each option. For the first option, one advantage involves the use of local field personnel's detailed knowledge of local vegetation/land cover conditions and road systems. In addition, it is believed that local field personnel, who eventually become users of the classification products, could make better use of the products by having been personally involved in their production. Finally, the use of local field personnel substantially reduces the funds needed for travel and per diem expenses. The principal disadvantage in using a large number of local field personnel is that it involves substantial coordination. It is believed that the magnitude of the coordination effort for a State as large as Mississippi and with a comparable number of operating agencies would be such that a designated coordinator would expend at least 25 percent of his time in coordinating activities related to ground truth gathering. In addition, the effort required in time and cost to orient a large number of local field personnel in ground truthing procedures is substantial.

Conversely, the second basic option discussed has the principal advantages that the little coordination required could be accomplished by the team itself, and the team could be formed by personnel already trained in the use of remotely sensed data. In addition, the team could give continuity to the total operation and could perform analysis during data processing more effectively than local field personnel. The main disadvantage of a small, centralized team is that the cost per training sample site established may be higher because of additional travel/per diem costs and because of some lost field time caused by the lack of familiarity with local road systems and conditions.

Distribution of Responsibility

When ground truth data are to be gathered by a large number of local field personnel, it is most important to have a well-conceived plan to distribute responsibility. However, because such a plan must consider the exact organization of field personnel to be involved, it is impossible to provide much more than general guidelines.

Ideally, training sample sites that are established to relate to a particular Landsat scene should be distributed so that some are located on each of the four CCT's. Because there is some shifting in the Landsat coverage from pass to pass and because areas covered by individual CCT's may not encompass all of a land unit (e.g., a county) to which field personnel relate, it is impractical to assign responsibility for areas corresponding to each CCT. If field personnel are assigned to counties or management units within a county, such as a game management unit or a State park, a practical and simple manner of assigning responsibility is to request that each field person establish a given number of training sample sites in each land cover condition within his respective county. In the

ASVT project, county field personnel were requested to establish one training sample site for each vegetation/land cover condition occurring within the area encompassed by an aerial photograph or map selected for each county. Since field personnel were organized along disciplinary lines, the result was that county foresters established one training sample site for each forest vegetation condition, County Extension agents established one for each cropland and pasture condition, etc.

The effect of supplying field personnel with one aerial photograph or township map within the respective county was to have some concentration of ground truth within each county, thereby saving time during the location of sites in the Landsat data at the same time that a distribution of ground truth throughout the Landsat scene was attained. The size of counties in Mississippi is such that, on the average, there are 12 counties within each Landsat scene; therefore, with a rule of one training sample site per land cover condition per county, 12 training sample sites for each land cover condition are theoretically possible. However, because all land cover conditions did not occur within the area covered by the aerial photograph or township map selected for each county, the actual outcome varied from three to eight land cover conditions per Landsat scene.

In situations where the average size of counties is substantially larger or smaller than the average Mississippi county, the guideline used for the ASVT project could be adjusted accordingly. For example, in the case of smaller counties, if there were from 20 to 24 counties per Landsat scene, one may select one-half the counties for ground truth activities and still use the simple instruction of one training sample site for each land cover condition. The set of four CCT's for a given Landsat scene is usually processed as one data set; however, because of cloud problems, it may be necessary to use one CCT from one scene acquired on a given date and three CCT's from another scene acquired on a different date. Consequently, it is desirable to select counties for ground truthing so that established training sample sites are likely to occur in the areas encompassed by each CCT in the nominal Landsat scene. In addition, if there are different physiognomic areas within a particular Landsat scene, counties should be selected to be somewhat proportional to the area encompassed by each physiognomic unit. For example, in Mississippi, a Landsat scene may encompass both an alluvial plains agricultural area and an uplands area with mixed land use; in this case, counties would be selected to represent each of these two physiognomic units roughly in proportion to the extent of each.

Timing of Ground Truth Gathering

Although the gathering of ground truth can take place during any time of the year, it is desirable to restrict field activities to be within a prime time defined for each season, thereby avoiding transitions in respect to seasonal change and/or agricultural land use. For example, in the transition between winter and spring, forest vegetation may reflect a condition in which deciduous trees are neither leafless nor fully foliated; and agricultural fields may be in a state in which some fields show stubble from the previous crop, some are being plowed, and some are being planted. There are two reasons for avoiding transition periods: there is a greater number of land cover conditions with which to deal, and the ground truth information is valid only for a short period. For example, a field with stubble in early spring may be plowed within days after ground truth has been gathered.

Although the commencement or termination of the phenomena associated with a given season may fluctuate somewhat from year to year because of weather patterns, the prime time for ground truth gathering for each season is generally considered to be a 4- to 6-week period that best typifies the following conditions in the season.

Winter - Deciduous vegetation is leafless; green, growing winter crops or pastures cover at least 40 percent of the soil underneath; or, if in northern latitudes, snow covers the surface of cropland and pasture areas.

Spring - New leaves on deciduous vegetation have fully foliated, but still have the color and texture of new leaves; land that will support crops during the summer is essentially exposed soil (plowed, disked, harrowed, planted, or germinated) but the vegetative portion of emergent plants covers less than 5 percent of the surface.

Summer - All vegetation is in a green, growing state with crops and/or pastures planted during the previous spring covering at least 40 percent of the soil underneath; in coastal marshlands, summer vegetation is predominant in those areas where there is a change in species predominance between spring and summer.

Fall - Deciduous vegetation is in a state of growth decline with most forest and brushland foliage in some stage of color change but not yet leafless; summer crops have been harvested, and most cropland reflects a stubble or plowed condition; pasture, range, and wetland grasses are in a yellowing or browning condition.

Since the accuracy with which various land cover types can be classified varies between seasons, ground truth activities for specialized classifications should be performed during the prime time for the selected season. For example, ground truth for a classification of coastal marsh vegetation should be gathered during prime time for the summer; the best time to separate evergreen forest from other vegetation is during winter; the best time for an agricultural crop classification is during summer; etc. If a good composite classification were desired with one set of data, a ground truth gathering operation during the spring season would be most appropriate. However, the best possible vegetation/land cover data base could be built up with a classification during each season, followed by subsequent updating as needed.

Another consideration as to the timing of ground truth information gathering concerns the means by which field activity is initiated. There are several possible alternatives. One alternative is to make a GO/NO GO decision based on observations of cloud cover at the time of each satellite overpass during the prime time. In this situation a GO decision would be made for the first cloud-free or relatively cloud-free (95 to 100 percent) pass, after which field personnel would be immediately notified to gather ground truth information within 10 to 15 days. This means of initiating field activity assures that ground truth will be close to the date for which Landsat data are acquired for processing, but requires a high degree of coordination between the weather observers, the decision-maker, and the field personnel. In addition, this method limits the amount of time available for field personnel to perform their work.

Another alternative is to preselect a scheduled Landsat pass date during prime time, then instruct field personnel to gather ground truth within a given number of days (e.g., 10 to 15 days) from that date. This alternative is easier to implement and gives the field personnel more flexibility in planning and conducting their activities to fit their own schedules, but has the disadvantage that the cloud condition may not be acceptable on the preselected overpass date. Of course, this does not preclude using ground truth acquired in this manner to process data from another pass closest to the preselected date.

A third alternative is to initiate field activities to occur within a defined 6-week prime time period without regard to the satellite overpass dates. This instruction offers the greatest flexibility for field personnel to schedule and conduct their work and is the easiest to implement, but increases the chance that some training samples will have to be discarded during data processing because land cover conditions changed between ground truthing and Landsat data acquisition. The latter approach was followed in the ASVT project and resulted in a discard rate of only 3 percent of all training samples because of an apparent change in conditions.

Ground Truth Information Forms

Ground truth information forms should be developed individually for each major land cover category or associated categories rather than as a single form for all land cover categories. For example, one form may be prepared for forest and brush vegetation, another for pasture and crops, another for urban areas, etc. Separate forms of this nature allow disciplinary field personnel to deal only with those forms pertinent to their responsibility, can be developed in a simpler format, and serve to reduce the total bulk of paperwork to be handled. Examples of forms used for this ASVT project are shown in appendix C. Another option in the development of forms concerns a "checkoff" versus a "fill in the blank" approach. A checkoff approach is preferred because it not only saves time but is much easier to use under field conditions. For most information, the checkoff approach is easy to develop and use; however, for some information, the fill-in-the-blank approach may be necessary. In the case of natural vegetation, unless the person developing the form is aware of all possible species associations in the area of concern, it is better to use a fill-in approach so as not to preclude obtaining ground truth on some species associations. However, even when the checkoff approach is used, field personnel should be instructed to establish training sample sites for any land cover condition encountered that meets the established criteria, even though the condition is not indicated on the form.

Orientation of Field Personnel

The ground truth package mentioned earlier contains an instruction sheet (see appendix D for an example) for field personnel that explains field procedures. However, it is desirable to hold orientation meetings with designated field personnel to deliver the package, review all details of its contents, outline areas of responsibility, discuss timing of ground truth gathering, etc. In the case of the ASVT project, orientation meetings were held at various locations throughout the State, usually in the district offices of each agency involved. Each meeting averaged approximately 3 hours; the first hour was used to explain the basics of satellite data acquisition and processing, and the last 2 hours were used to review the ground truth package contents, explain procedures and areas of responsibility, etc. However, if time and travel funds permit, it is preferable to assemble field personnel for orientation meetings at the location of the data processing equipment and provide them with a full day of orientation. This would allow a system demonstration, including the mechanics of locating training sample sites in the CRT display of the Landsat data image. Experience at ERL shows that such a demonstration gives field personnel a better feel for the process in which sites can be located through reference to other features in the image and visually emphasizes the need to establish sites that are uniform and homogeneous in respect to the particular land cover condition.

Fieldwork

The essence of ground truth fieldwork is to verify or establish the location of each training sample site that is uniform and homogeneous with respect to the particular land cover condition and to fill out a ground truth information form for each site. If potential training sample sites were preselected through photointerpretation, the field agent simply locates the delineated area on the ground, verifies that the area delineated is uniform and homogeneous, and makes the necessary observations to fill in the ground truth information form. If the delineated area is found to fail to meet the uniformity criteria, it may be discarded and another area selected. In other cases, it may be preferable to erase part or all of the delineated boundary and indicate a new boundary, shifted slightly from the original delineation. In some cases, the photointerpreter may have delineated a uniform and homogeneous vegetation condition but may have interpreted it incorrectly. For example, the photointerpreter, using photography acquired during the winter season when hardwood trees were leafless, may have been misguided by an evergreen understory component (e.g., holly or wax myrtle) visually apparent through the leafless overstory and identified it as pine forest. In this case, the field person could simply change the letter/number identifying symbol recorded by the interpreter to the correct symbol and fill out the form accordingly.

In filling out the ground truth information form, the field person may take various approaches. If a training sample site delineates an agricultural crop in a 16-hectare (40 acre) field bounded by roads on two or more sides, the field person may make most observations from a vehicle, stopping only to make two or three spot checks by walking into the field. In the case of natural forest vegetation, he may use pacing and a hand compass to keep his bearings as he follows some pattern to assure adequate coverage (as suggested in appendix E) along which he stops occasionally to make visual observations. As the location and delineation of training sample sites and corresponding completion of the ground truth information form proceeds, it is extremely important that the letter/number identifier symbol (as described previously in this report) be recorded (both on the ground truth form and on the aerial photoprint or map) adjacent to the delineation of the corresponding training sample site. In addition, it is very helpful if field personnel staple all ground truth forms to the photoprint or map on which training sample sites corresponding to those forms are delineated. It is desirable to delineate training sample sites on recent aerial photoprints or photograph-based maps; however, if such are not available, training sample sites can be delineated on 7.5'-series (1:24 000) topography maps, or, in the absence of those, 15'-series (1:62 500) topography maps, provided that such maps are not grossly outdated. However, it is recommended that maps at scales smaller than 1:63 360 (1 centimeter = 0.63 kilometer (1 inch = 1 statute mile)) should not be used for training sample site delineation.

If potential training sample sites are not preselected through photointerpretation, local field personnel may delineate some sites on aerial photographs or maps in the office based on their knowledge or on office records of the area. Visitation of these sites as well as delineation and visitation of additional sites can usually be performed in the course of the local field personnel's routine work. However, if the time period indicated for the ground truth gathering activity is short, a special effort may be required. After sites are located, the work proceeds in the same manner as described for sites preselected through photointerpretation.

If personnel are employed exclusively for ground truth gathering, it is most efficient if they function as a team by meeting at the end of each day to keep a master list of training sample sites established and to plan the next day's activity. In this manner, preselected sites that may have been rejected or lost because of access problems may be substituted for by another teammember.

The involvement of field personnel in producing a land cover classification with Landsat data may end with delivery of aerial photoprints with delineated training sample sites and corresponding ground truth information forms. However, it is desirable that field personnel also assist in the location of training sample sites in the Landsat data. Experience at ERL has shown that assistance from field personnel can save time, both through more rapid location of sites in the CRT display and in catching possible recording errors.

Once training sample sites have been established for the first land cover classification, ground truth information for additional classifications can be obtained with substantially less effort. Except in the case of agricultural land on which use changes from year to year, it is only necessary to ascertain that no drastic change has occurred since the first ground truth effort; consequently, ground truth forms can be greatly simplified. An example of a ground truth form prepared for revisits to training sample sites established for forest vegetation is shown in appendix F.

Time Required and Cost

A tally of time required to make observations within a training sample site, to delineate the site on an aerial photograph or map, and to fill out the ground truth information form showed the following distribution for the ASVT project.

Number of sites	Time required
93	5 to 15 minutes
130	15 to 30 minutes
117	30 to 60 minutes
37	1 to 2 hours
8	2 to 3 hours
5	Over 3 hours

There was a noticeable difference in time required for training sample sites for different land cover conditions. On an average, crop and pasture sites required 24 minutes per site, whereas forest and brushland sites required 43 minutes. It was not possible to keep account of travel time and expense (vehicle operation and depreciation costs) because most sites were established by field personnel during their routine activities. However, ERL experience shows that a field person can be expected to establish and provide ground truth on an average of six training sample sites per day when travel time between sites within the county is included. Consequently, for programmatic purposes, it can be estimated that 25 man-days (200 man-hours) would be required to address up to 150 training samples that may be established for one Landsat scene of 34 300 square kilometers (13 300 square statute miles).

The cost of ground truth operations can be calculated by using a rate of \$10.50/hr for the estimated 200 man-hours (to cover all salary, overhead, and operating costs) and dividing by the total area of the Landsat scene (34 300 square kilometers (13 300 square statute miles)). The calculation showed that ground truth operations would cost \$0.06/km^2 (\$0.16/mi^2) for the initial classification. This calculation is compatible with past ERL cost calculations, although they were derived in a research rather than an operational environment. The ERL-derived cost calculations indicate a range of from \$0.058/km^2 (\$0.15/mi^2) for the easiest ground truth gathering (i.e., recent aerial photography, field personnel familiar with area, easy access or terrain) to \$0.12/km^2 (\$0.31/mi^2) for difficult operations (aerial photography not available, field personnel unfamiliar with area, difficult access or terrain). Actually, it is unrealistic to charge all costs for the first ground truth effort against the initial classification because revisiting the established training sites for subsequent classifications requires far less effort.

In certain situations, such as when dealing with large, inaccessible marsh or wetlands areas, ground truth gathering can be more costly. It is in these conditions that selected coverage with aerial photography (if not already available) and/or the use of helicopters should be considered. However, even after comparisons are made with costs assuming access by boat, a higher cost for use of helicopters may be considered an adequate trade-off in view of the time required for a limited number of personnel to use boat transportation.

Although ground truth for this project was acquired by local field personnel, it was mentioned earlier in this report that an option would be to use a ground truth team that would work almost exclusively for this purpose. Such an effort for a State as large as Mississippi may involve a breakout of work activity as follows.

Activity	Days
Photointerpretation and prefield preparation	150
Field work and travel within counties	180
Postfield records and location of sites in Landsat data	150
Travel from central location to counties	48
Total	528

Allocating 220 workdays per year for each of three persons (e.g., an agronomist, a forester, and a botanist) results in a total of 660 days, leaving about 20 percent of their time for interface with users or other activities such as digitizing other information. However, it may be most desirable to use local field personnel for the first complete ground truthing effort, then use a full-time team of three for revisits and updating for subsequent classifications and specialized classifications. In this manner, the first ground truthing operation could be accomplished rapidly, and the field personnel, who eventually become users of the land cover classification, become familiar with the manner in which the classification is derived from Landsat-acquired data.

CONCLUDING REMARKS

This report addressed ground truth information gathering procedures relative to performing a land cover classification using Landsat MSS digital data and the supervised approach to computer-implemented data processing. One difference between supervised and cluster analysis approaches is that, in the former, ground truthing takes place first so that the computer can be programed to recognize a land cover condition elsewhere; whereas, in the latter, classification takes place first, and then a ground truth operation is launched to determine the land cover condition that corresponds to each resulting class. Modified approaches employing unaided training sample selection and supervised classification are also in use.

However, since the same basic data are utilized and the basic principles involved in the measurement of reflected and/or emitted energy are the same, this report should have relevance to ground truthing activities irrespective of the approach to data processing.

It is unlikely that new techniques will result in drastic changes to the procedures outlined in this report in the near future, However, techniques that are currently being developed and/or tested may cause some slight changes in ground truth gathering procedures. When raw data are registered to a given map projection so as to permit the development of techniques to allow automated location of training sample sites in the Landsat data, it will be necessary to determine the map coordinates (ref. 6) that define the location of each training sample site. If techniques currently under development and testing at ERL to define categories of mixed vegetation through distribution relationship analysis of classified data are successful, there may be no need to establish training sample sites for some mixed vegetation categories (e.g., an oak-pine mix). Also, when techniques to merge land cover information from seasonal classifications into a master composite classification are perfected, it may be desirable to conduct ground truth activities during each season of the year in such a manner that each seasonal activity encompasses only those land cover categories that can be most accurately classified with Landsat data acquired during that particular season.

This report was written primarily to help field personnel understand the principles and procedures involved in ground truth information gathering. However, it is believed that it would be beneficial for field personnel to familiarize themselves with aspects of data processing and analysis. Such familiarization would not only enhance their understanding of ground truth information gathering, but would also give them a better understanding of both the advantages and limitations of using land cover classifications derived from Landsat data.

Lyndon B. Johnson Space Center
National Aeronautics and Space Administration
Houston, Texas, September 19, 1977
177-52-89-00-72

REFERENCES

1. Butera, M. K.: A Technique Development for the Determination of Louisiana Marsh Salinity Zones from Vegetation Mapped by Multispectral Scanner Data: A Comparison of Results from Satellite and Aircraft Data. NASA TM X-58203, 1977.

2. Joyce, A. T. and Griffin, R. H.: The Use of Landsat Digital Data and Computer-Implemented Techniques for an Agricultural Application. NASA RP-1016.

3. Whitley, Sidney L.: A Procedure for Automated Land Use Mapping Using Remotely Sensed Multispectral Scanner Data. NASA TR R-434, 1975.

4. Jones, Clay: Implementation of an Advanced Table Look-Up Classifier for Large Area Land-Use Classification. Proceedings of the Ninth International Symposium on Remote Sensing of Environment, vol. II, Environmental Research Institute of Michigan, Ann Arbor, Mich., Apr. 1974, pp. 813-836.

5. Whitley, S. L.: Low-Cost Data Analysis Systems for Processing Multispectral Scanner Data. NASA TR R-467, 1976.

6. Graham, Marcellus H.: Digital Overlaying of the Universal Transverse Mercator Grid with Landsat Data Derived Products. NASA TM 58200, 1977.

APPENDIX A

LAND COVER CONDITIONS IN MISSISSIPPI

The major land cover categories and the various conditions of each category for the State of Mississippi are shown in table A-I. Such a listing is typical of that for a State as large and as varied as Mississippi. Some of the vegetation types listed do not occur at all times during the year; consequently, only a portion of the land cover conditions shown would be found during a ground truth operation conducted during a particular time of the year. Experience at ERL has shown that, once the vegetation types occurring at a given time of the year have been properly identified as to variation in age, density, understory, and topography, there are typically around 30 to 50 land cover conditions within a 185- by 185-kilometer (115 by 115 statute mile) area encompassed by a Landsat scene.

The term "land cover condition" is used to refer to a particular combination of surface features that are likely to influence the reflected energy as measured with the MSS. For example, a land cover condition may be a combination of surface features described as "a sparse pine forest with a grass understory on the south aspect of a 30- to 50-percent slope." Another land cover condition may be described simply as "asphalt," implying that the material itself is the only feature expected to influence the reflected energy. Of course, a categorization of training sample sites as to land cover condition does not assure that each land cover condition categorized will be spectrally separable from all others. However, without such categorization, it would be impossible to attempt a computer-implemented classification of each land cover condition with Landsat digital data.

TABLE A-I. - MISSISSIPPI TRAINING SAMPLE MASTER FILE CATEGORY LIST

Category	Code	Condition
Brushland	BD	Brush debris (e.g., recent clear-cut)
	BE	Brush evergreen
	BH	Brush deciduous
	BM	Brush mixed
Cropland	CA	Potatoes, sweet
	CB	Grain sorghum
	CC	Cotton
	CE	Exposed soil, cleared
	CF	Crop fallow
	CG	Forage
	CH	Peppers
	CI	Plowed
	CJ	Disked
	CK	Harrowed
	CM	Wheat
	CN	Corn
	CO	Stubble
	CP	Peanuts
	CQ	Cucumbers
	CR	Rice
	CS	Soybeans
	CT	Oats
	CW	Watermelons
	CZ	Field peas
Extractive	EC	Clay extraction
	EG	Gravel pit
	EM	Strip mine, coal
	EQ	Quarry/limestone
	ES	Sand pit
	EZ	Sand/gravel pit
Forest land	FP	Oak-pine mix
	HA	Elm-ash-cottonwood
	HB	Maple-beech-birch
	HC	Cypress-tupelo
	HD	Leafless hardwood with deciduous understory
	HE	Leafless hardwood with evergreen understory
	HH	Oak-hickory
	HM	Hardwood mixed
	HO	Oak-gum-cypress
	HP	Hardwood plantation
	HW	Willow
	PL	Loblolly-short leaf
	PP	Pine plantation
	PS	Longleaf-slash

Category	Code	Condition
Inert materials	IB	Barren or rock outcrops
	IG	Building
	IH	Hard surface (asphalt, concrete)
	IM	Mud flat
	IS	Sand beach or bar
Marshland	BA	*Baccharis halimifolia*
	BB	*Baccharis halimifolia*/goldenrod
	DS	*Distichlis spicata*/*Scirpus*
	JA	*Juncus roemerianus*/*Spartina cyanosuroides*
	JB	*Juncus roemerianus*/*Baccharis halimifolia*
	JC	*Juncus roemerianus*/*Distichlis spicata*
	JD	*Juncus roemerianus*/*Spartina alterniflora*
	JS	*Juncus roemerianus*/*Spartina patens*
	ME	*Cyperus*/Eleocharis cellolsa
	SC	*Spartina cyanosuroides*/*Scirpus*
	SJ	*Spartina alterniflora*/*Juncus roemerianus*
	TY	*Typha*
Natural grassland	NF	Native field grass
	NW	Native woodland grass
Orchards	OC	Citrus
	ON	Pecans
	OP	Peaches
Pasture and hayland	IA	Alfalfa
	IB	Bermuda
	IC	Bahia
	ID	Dallas
	IE	Combination
	IF	Fescue
	IO	Other
	IT	Temporary (e.g., ryegrass)
Urban buildup	UH	High density
	UL	Low density
Water	WC	Catfish pond
	WD	Deep lake, reservoir
	WO	Other
	WR	River
	WS	Shallow lake, reservoir

APPENDIX B

DEFINITION OF MAJOR LAND COVER/VEGETATION TYPES

The definitions and criteria of various land cover categories and conditions are included in the following list.

Cropland - A specified unit area that is usually planted to an agronomic crop or grass on an annual basis after soil preparation

Pasture/grassland - Specified unit area of which 90 percent or more of the surface covered with foliage is covered with foliage of grasses, generally used for grazing or hayland on other than an annual basis

Forestland - Specified unit area of which 10 percent or more of the surface area is covered with foliage of trees

Pine forest - Forest in which 66-2/3 percent or more of the area covered with foliage of trees is covered by foliage of evergreen trees as seen from above

Hardwood forest - Forest in which 66-2/3 percent or more of the area covered with foilage of trees is covered by foliage of deciduous trees as seen from above

Mixed pine/hardwood - Forest that does not meet the preceding criteria for evergreen or deciduous forest

Brushland - Specified unit area of which 90 percent or more of the surface area covered with foliage is covered with foliage of multistemmed, perennial shrub species

Forested wetlands - Forested areas that are seasonally flooded for prolonged periods (usually 3 months or more) and/or flooded because of diurnal tidal action directly or indirectly through water backup

Marshland - Specified unit area that is frequently inundated for prolonged periods and contains plant species typical of nonforested wetlands covering 90 percent or more of its surface

Species association - A vegetation type in which two or more plant species grow intermingled, with the foliage of each species covering at least 25 percent of the surface area as seen from above

Sparse crown closure - Forested area in which 10 to 65 percent of the surface is covered by crowns (foliage and branches) of overstory trees when in leafed condition

Dense crown closure - Forested area in which 65 to 100 percent of the surface is covered by crowns (foliage and branches) of overstory trees when in leafed condition

APPENDIX C

GROUND TRUTH DATA FORMS

Examples of the ground truth data forms used by field personnel for forest, brush, and orchards; for crops and pasture; for extractive land uses; for urban areas; and for marsh vegetation are shown in figures C-1 to C-5, respectively.

GROUND TRUTH DATA FOR FOREST, BRUSH, AND ORCHARDS

TAKEN BY: ______________________________ DATE: __________

TRAINING SAMPLE IDENTIFIER ______________ MAP OR AIR PHOTO INDEX # __________

ESTIMATED FIELD SIZE: __________ ft X ______ ft. or ______________ ACRES

LOCATION ____

County ¼ ¼ Section Township Range

KIND OF VEGETATION (Check One)
( ) Natural Forest
( ) Forest Plantation
( ) Brush Vegetation

IF NATURAL FOREST, INDICATE:

(1) Major forest type (check one)

( ) Maple-Beech-Birch	( ) Elm-Ash-Cottonwood	( ) Aspen-Birch
( ) Oak-Hickory	( ) Loblolly-Shortleaf	( ) Oak-Pine
( ) Oak-Gum-Cypress	( ) Longleaf-Slash	( ) Mixed Hardwood

(2) Overstory Crown Closure

( ) Dense (65% to 100%) ( ) Sparse (10% to 65%)

(3) Overstory species composition (to nearest 25%)

Species	%

(4) Understory species compostion (to nearest 25%)

Species	%

(5) Average age class of upper canopy trees (check one)

( ) Less than 20 years
( ) 20 to 50 years
( ) 50 to 100 years
( ) over 100 years

(6) Average height class of upper canopy trees (check one)

( ) Less than 20 feet
( ) 20 to 50 feet
( ) 50 to 100 feet
( ) over 100 feet

Figure C-1.- Ground truth form for forest, brush, and orchards.

(FOREST, BRUSH, ORCHARDS CONTINUED)

(7) Slope (Check One)

( ) 0% to 10%
( ) 10% to 30%
( ) 30% to 50%
( ) 50% or more

(8) Predominant Aspect (Check One)

( ) North ( ) South ( ) East ( ) West

If Forested Wetlands are flooded at time of observation, indicate depth of water:

( ) less than 1' ( ) 2' to 4' ( ) greater than 4'

or if not flooded at time of observation indicate:

( ) appears subject to flooding by water backup due to tidal action.

( ) appears to have been flooded for a prolonged period prior to observation.

If Forest Plantation or Orchard, indicate:

Species ____________________ Average Age ____________________

Spacing ____________________ Average Height ____________________

Row Direction ____________________

If Brushland, indicate species composition to nearest 25%:

(1) Species	%
____________________	__________
____________________	__________
____________________	__________
____________________	__________

(2) Vegetation Density:

( ) Sparse, 10% to 65% of surface covered.

( ) Dense, 65% to 100% of surface covered.

(3) If sparse density, ground level is:

( ) Grass

( ) Exposed earth

( ) Other ____________________

Figure C-1.- Concluded.

GROUND TRUTH DATA FOR CROPS AND PASTURE

TAKEN BY__________________________ DATE____________

TRAINING SAMPLE #______________ MAP OR AIR PHOTO INDEX #__________

ESTIMATED FIELD SIZE: ______ft X ______ft. or ____________ACRES

LOCATION____

County 1/4 1/4 Section Township Range

GENERAL CONDITION(1) ________________________________

DESCRIPTION (if not crop or pasture)________________________

____

CROP OR PASTURE SPECIES(2) ______________ VARIETY (if known)________

PLANTING TECHNIQUE(3) ______________ PLANT HEIGHT (to closest ft)______

ROW WIDTH____________________ PHYSIOLOGICAL STATE(4) __________

ROW DIRECTION__________________ VISUAL ASPECT(5) ____________

PERCENT GROUND COVER ( ) 0% to 20% ( ) 40% to 60% ( ) 80% to 100%
( )20% to 40% ( ) 60% to 80%

WEED INFESTATION (species & %, if greater than 20%)______________

DISEASE INFESTATION (kind & %, if greater than 20%)______________

INSECT INFESTATION (kind & %, if greater than 20%)______________

SOIL CONDITION(6) ____________________________________

SOIL MOISTURE(7) ____________________________________

SOIL TYPE(8)(if available)______________________________

____

OTHER COMMENTS (if needed)______________________________

____

(1) e.g. crop, pasture, stubble, plowed, fallow.
(2) e.g. soybean, bahia grass, etc.
(3) e.g. row, skip row, drilled, broadcast.
(4) e.g. flowering, heading, mature, etc.
(5) e.g. chlorotic, wilted, etc.
(6) e.g. freshly cultivated, rough, smooth, etc.
(7) e.g. moist, dry, waterlogged, etc.
(8) series, texture, color, slope, etc.

Figure C-2.- Ground truth form for crops and pasture.

GROUND TRUTH DATA
Extractive Land Uses

OBSERVATIONS MADE BY ______________________________ DATE ______________

IDENTIFIER NO.* ______________ Approx. Size _____ X _____ (feet) or _______ acres.

COUNTY ______________________________

LOCATION (if known) ____
Township Range Section Quarter Forty

ACTIVITY TYPE

( ) Sand pit
( ) Gravel pit
( ) Stone, dimension
( ) Stone, crushed
( ) Lime
( ) Cement

( ) Clay
( ) Chert & Tripoli
( ) Lignite
( ) Heavy mineral
( ) Other ____________________
____

Is area ( ) in-production or ( ) abandoned?

If abandoned, is area ( ) barren or ( ) revegetated?

Is the area likely to contain impounded water during all or a significant part of year ( ) yes ( ) no?

How much time did it take to make observations and fill out this form (min. and/or hours) ____________________.

*Observations should only be made on extractive areas that are at least 600 feet by 600 feet, or approximately 10 acres. Once such an area is located, its location should be delineated on an aerial photo or map sheet with colored pen or pencil, and an identifier cross-reference number should be recorded on the aerial photo or map beside the delineated area and on the ground truth data form.

Figure C-3.- Ground truth form for extractive land uses.

GROUND TRUTH DATA FORM FOR URBAN AREAS[1]

Training Sample ID No.____________________

Collected by:____________________ Date:____________

High Density Urban ( )[2]
Low Density Urban ( )[3]

If High Density Urban - Predominantly Concrete ( )
Predominantly Asphalt ( )
Predominantly Other ( ) ____________ e.g., metal roof
Inert Material Complex ( )

Comments:[4] ____________________

If Low Density Urban
Main type of inert Material - Roof tops ( )
Concrete ( )
Asphalt ( )
Other ( ) ____________

Main type of vegetation -
Grass (lawns) ( )
Pine trees ( )
Hardwood trees ( )
Mixed pine/hardwood ( )
Mixed grass/trees ( )

Comments:[4] ____________________

(1) An urban area training sample should be 1000 ft. by 1000 ft. or larger; however, if homogeneous areas of such dimensions cannot be located, areas of 500 feet by 500 feet or larger (approx. a city block) are acceptable.

(2) High Density Urban is defined as an area essentially devoid of vegetation; but with up to 10% covered with vegetation in small scattered parcels whose largest dimension is generally less than 100 feet.

(3) Low Density Urban is defined as an area within which inert materials (roof tops, concrete, asphalt) are predominant; but with up to 45% of the surface covered with vegetation, including overtopping trees, occurring in small, scattered parcels with the maximum dimension of each parcel no greater than 200 feet.

(4) Appropriate comments include identification of scenario, e.g., airport runway, industrial complex, downtown commercial area, etc.; height of buildings, e.g., one or two story, three to five story, 6 or more stories; pitch of roofs, e.g., flat, moderate angle, steep angle; or any other information pertinent to measurements made with overhead remote sensors.

Figure C-4.- Ground truth form for urban areas.

GROUND TRUTH FORM FOR MARSH VEGETATION

1. Sample number_____________.

2. Date: ____________________.

3. Time: ____________________.

4. Vegetation type:

   (1) pure stand (monotypic)________________.

      (a) species:__________________________.

   (2) intermixed (less than 6 vascular species present) ________________.

      (a) dominant species:________________________, ____________________.

   (3) intermixed (more than 6 vascular species present) ________________.

      (a) dominant species:________________________, ____________________.

      (NOTE: If a species comprises less than 5% of vegetation do not regard as major or dominant component.)

5. Homogeneity:

   (1) sub-elements (defined)

      (a) vegetation differences (clumps, patches, zones) ________________.

      (b) barren areas ____.

      (c) open water ____.

      (d) sparse vegetation/barren _____________________________________.

      (e) sparse vegetation/water ______________________________________.

      (f) other (describe) ____.

   (sub-elements (size)

      (a) less than 10 feet ____.

      (b) more than 10, but less than 20 ______________________________.

      (c) more than 20, but less than 40 ______________________________.

      (d) more than 40, but less than 60 ______________________________.

Figure C-5.- Ground truth form for marsh vegetation.

(3) distribution (of sub-elements in study area).

(a) evenly ____________________.

(b) center ____________________.

(c) peripheral ____________________.

(4) density (of vegetation as % of surface area).

(a) dense > 90 ____________________.

(b) intermediate < 70 ____________________.

(c) sparse < 50 ____________________.

6. Height of plants (stands).

(1) approximate height of major units:

(a) species ____________, height ____________.

(b) species ____________, height ____________.

(c) species ____________, height ____________.

(2) approximate height of minor units:

(a) species ____________, height ____________.

(b) species ____________, height ____________.

7. Status of vegetation:

(1) approximate (%) of dead-standing material.

(a) major units (species) ____________________.

(b) minor units (species) ____________________.

8. Stage of growth:

(1) major units

(a) dormancy (winter-no leaves) ____________________.

(b) dormancy (winter-leaves dead-standing) ____________________.

(c) seedlings ____________________.

(d) immature ____________________.

(e) mature ____________________.

(f) anthesis ____________________.

Figure C-5.- Continued.

(g) vigor ______________________.

(1) excellent ______________________.

(2) fair ______________________.

(3) poor ______________________.

9. Surface of substratum:

(1) covered by algae ______________________.

(2) covered by small vascular plants ______________________.

(3) covered by detritus ______________________.

(4) barren ______________________.

(5) substrate type

(a) mud ______________________.

(b) sand ______________________.

(c) sandy/mud ______________________.

10. Water level.

(1) standing on surface of marsh ______________________.

(a) covered by tidal water ______________________.

(b) covered by river overflow ______________________.

(c) combination of both (a & b) above ______________________.

(d) permanent or semi-permanent ______________________.

(2) Depth of water on marsh surface ______________________.

11. Comments:

Figure C-5.- Concluded.

APPENDIX D

TYPICAL INSTRUCTION SHEET

The following is a typical instruction sheet for field personnel documenting ground truth information for agronomic crops and pastures.

PROCEDURE FOR ESTABLISHING "TRAINING SAMPLE AREAS" AND

DOCUMENTING "GROUND TRUTH" FOR AGRONOMIC CROPS AND PASTURES

STEP #1 - Locate one typical field for each of the different crops and pastures that occur within the geographic area covered by the aerial photo (or photomap) provided. Each such field will be referred to as a "training sample area." (Note: a 10-acre field is the minimum size suitable for a training sample, but a larger field, 40 acres to 160 acres, is desirable.)

STEP #2 - Outline each training sample area located in Step #1 with pen or pencil, assign a reference number to each (starting with the number one), and record the reference number on the aerial photo (or photo map) along side of each outlined field (training sample area).*

STEP #3 - For each training area outlined and referenced on the aerial photo (or photomaps) in Step #2, fill out one "Ground Truth Data" form. Information on the form that is not readily available or not applicable can be so indicated in the appropriate blank. Record the index number of map or airphoto print on which the training sample is located in the upper right hand corner of the form.

STEP #4 - Return all materials to project coordinator as soon as Step #1 thru #3 are accomplished. This can take place between July and August; however, the earlier the better.

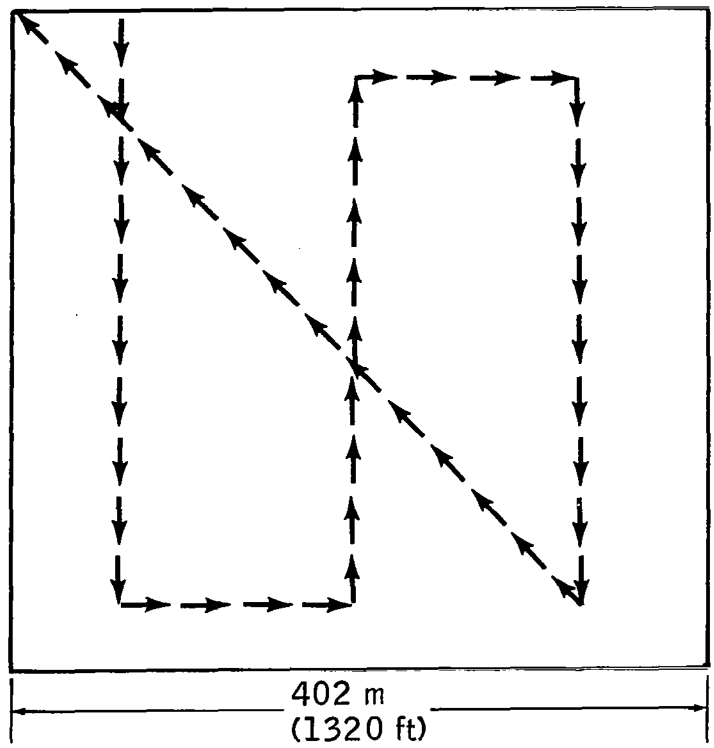
*If scale is 1:62,500 (air photo), a 40-acre field is roughly 1/4" X 1/4" on the photo; if scale is 1:24,000 (township map), a 40-acre field is roughly 2/3" X 2/3" on map.

Figure D-1.- Typical instruction sheet for ground truth field personnel.

APPENDIX E

FOREST UNIFORMITY VERIFICATION PATTERN

The following sketch illustrates the suggested coverage pattern to verify the uniformity of a forest vegetation training sample site of approximately 16 square hectares (40 acres) in size.



APPENDIX F

TYPICAL REVISIT FORM

SITE IDENTIFIER CODE (6-7 digit code recorded on air photo or land use map)________________

COUNTY:______________________

OBSERVATIONS MADE BY:__________________________________DATE:____________________

Has the vegetation within the training sample area delineated on the air photo or land use map been altered during the last year? ☐ yes ☐ no

If yes, what was the cause?

☐ logging

☐ land clearing

☐ fire

☐ heavy insect or disease mortality

☐ other (indicate) __________________________

In which month did the alteration occur (if known)?__________________________________

How much time did it take you to make observations and fill out this form (min. and/or hours)?

____

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA RP-1015		

4. Title and Subtitle PROCEDURES FOR GATHERING TRUTH INFORMATION FOR A SUPERVISED APPROACH TO A COMPUTER-IMPLEMENTED LAND COVER CLASSIFICATION OF LANDSAT-ACQUIRED MULTISPECTRAL SCANNER DATA	5. Report Date January 1978 6. Performing Organization Code JSC-12910
7. Author(s) Armond T. Joyce	8. Performing Organization Report No. S-478
9. Performing Organization Name and Address Lyndon B. Johnson Space Center Houston, Texas 77058	10. Work Unit No. 177-52-89-00-72 11. Contract or Grant No.
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D. C. 20546	13. Type of Report and Period Covered Reference Publication 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Procedures for gathering ground truth information for a supervised approach to a computer-implemented land cover classification of Landsat-acquired multispectral scanner data are provided in a step-by-step manner. Criteria for determining size, number, uniformity, and predominant land cover of training sample sites are established. Suggestions are made for the organization and orientation of field team personnel, the procedures used in the field, and the format of the forms to be used. Estimates are made of the probable expenditures in time and costs. Examples of ground truth forms and definitions and criteria of major land cover categories are provided in appendixes.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Earth resources; Pattern recognition; Crop identification; Land cover classifications; Multispectral band scanners; Ground truth; Ground cover; Landsat; Land use	STAR Subject Category: 43 (Earth Resources)

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	47	$4.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

NASA-Langley, 1978

National Aeronautics and
Space Administration

Washington, D.C.
20546

Official Business
Penalty for Private Use, $300

THIRD-CLASS BULK RATE

Postage and Fees Paid
National Aeronautics and
Space Administration
NASA-451

U.S.MAIL

20 1 1U,E,SPGEN,010978 S00903DS 740731
DEPT OF THE AIR FORCE
AF WEAPONS LABORATORY
ATTN: TECHNICAL LIBRARY (SUL)
KIRTLAND AFB NM 87117

NASA

POSTMASTER: If Undeliverable (Section 158 Postal Manual) Do Not Return