CSIRO Publishing

International *Journal* of Wildland Fire

Scientific Journal of IAWF

VOLUME 10, 2001

Address manuscripts and editorial enquiries to:
International Journal of Wildland Fire
Editor in Chief
Dr Gwynfor Richards
Department of Mathematics and Computer Science
Brandon University
Brandon, Manitoba, Canada R7A 6A9
Telephone: +1 204 727 7362
Fax: +1 204 728 7346
Email: richards@brandonu.ca

International Association of Wildland Fire

Address subscription enquiries to:
CSIRO PUBLISHING
PO Box 1139 (150 Oxford St)
Collingwood, Vic. 3066
Australia
Telephone: +61 3 9662 7644
Fax: +61 3 9662 7611
Email: ijwf@publish.csiro.au

CSIRO PUBLISHING

Published by CSIRO Publishing
for the International Association of Wildland Fire

www.publish.csiro.au/journals/ijwf

International Journal of Wildland Fire, 2001, **10**, 277–287

A primer on mapping vegetation using remote sensing

Tom Bobbe[AC], *Henry Lachowski*[A], *Paul Maus*[A], *Jerry Greer*[A] *and Chuck Dull*[B]

[A]USDA Forest Service, Remote Sensing Applications Center, 2222 W. 2300 S., Salt Lake City, UT 84119, USA.
[B]USDA Forest Service, Engineering, 14th & Independence Avenue, Washington, DC 20250, USA.
[C]Corresponding author; Telephone: +1 801 975 3751; email: tbobbe@fs.fed.us

This paper was presented at the conference 'Integrating spatial technologies and ecological principles for a new age in fire management', Boise, Idaho, USA, June 1999

Abstract. The use of information based upon remotely sensed data is a central factor in our 21st Century society. Scientists in land management agencies especially require accurate and current geospatial information to effectively implement ecosystem management. The increasing need to collect data across diverse landscapes, scales, and ownerships has resulted in a wider application of remote sensing, Geographic Information Systems (GIS) and associated geospatial technologies for natural resource applications. This paper summarizes the use of digital remotely sensed data for vegetation mapping. Key steps in preparing vegetation maps are described. These steps include defining project requirements and classification schemes, use of reference data, classification procedures, and assessing accuracy. The role of field personnel and inventory data is described. Case studies and applications of vegetation mapping on national forest land are also included.

Keywords: remote sensing, GIS, mapping, geospatial, project planning.

Introduction

People in general do not realize how pervasive the use of remotely sensed information is in their daily lives and in the commerce of the USA. Information gathered by remote sensing is used to predict crop production, identify drought effects, map insect outbreaks, map oceanographic conditions, evaluate global weather conditions, and to analyse and plan land use. The use of remote sensing and geospatial data in vegetation and land cover mapping is clearly established. As demands for goods and services by the human population increase, the value of renewable resources increases. The need for more and more specific data about the extent and condition of vegetation cover becomes more pronounced. If we are to succeed in managing ecosystems to ensure that we are meeting management goals, we must depend on data from a variety of remote sensing sources. Used in a geographic information system, the data provide the decision-maker with critical information that can help meet the needs of land and water managers.

Data acquisition and analysis requires the investment of limited time and money. Any effort to use digital and ancillary data must be well planned if the user is to gain the most benefit. There are a number of basic considerations which decision-makers and technicians must consider if an analysis is to be useful in the assessment of natural resources.

Project design and analysis

Introduction

In the following eight subsections, entitled *Defining project requirements*; *Project management*; *Scale and resolution*; *Selecting the correct imagery and data*; *Use of reference data*; *The classification process*; *Assessing accuracy*; *and Deriving polygon layers and other GIS data*, we review the steps we use to complete an assessment project.

Defining project requirements

A successful resource assessment project based on the use of digital data is the result of well-planned and effective project management. The need for a clear plan and understanding of the project objectives cannot be overstated. Project management is the planning, organizing and management of personnel, equipment, time, money, and data resources to accomplish a specific objective. There are five important

10.1071/WF01029 1049-8001/01/030277

steps in this phase of a resource mapping or assessment project.

First: Identify the user's needs. An interdisciplinary team working with the decision-maker should determine very specifically what questions are to be answered and what information is to be displayed on vegetation or resource maps. All subsequent steps depend upon a clear statement of the need for information. It is critical that the team understands the decisions that will be based on their assessment.

Second: Identify data requirements. Refer constantly to the objectives of the project and find out what data are available that will help answer the questions. Identify gaps in the data and develop a plan to acquire data that are needed but not available.

Third: Identify analytical needs. Will the data be interpreted manually or with a computer-aided process? Are software programs needed for the project available or must they be developed?

Fourth: Identify the processing system requirements. Determine if existing hardware and software are adequate for the analysis. Are the data in a useful format? Is there adequate data storage space in the computer system? Can raw data and interpreted data be archived safely?

Fifth: Identify staffing needed for the project. Does the necessary staff have time to do the project? Is it a priority job for all of them? Is outside expertise needed to fill holes in staffing? Determine what is needed, calculate the cost, and cover this in a budget. At the same time, determine if staff training is needed for the project and schedule it before the project starts.

Project management

A full discourse on planning is beyond the scope of this paper but a brief review is necessary. It is important to recognize that there are three phases to project implementation: plan the project, monitor the project, and finally wrap-up the project.

Writing the project plan is the first step after the project requirements are identified. The plan should include all the material identified as requirements. It should also include a budget that sets out all the costs in time and materials. A time line should be established and the work should be divided into smaller, manageable tasks. All participants must be able to take the plan and understand what their part is in the project. Accuracy standards must be established. The plan must set out the classification scheme with complete definitions and descriptions of the various classes.

The plan will include a section on monitoring and this critical step must not be overlooked. As the project is carried out, the manager must compare what is actually accomplished against the planned time line and expectations. Monitoring gives the manager the information needed to reallocate resources to keep the project on track.

Finally, an effective project wrap-up will ensure that the data get to those who need the information. Project results must be presented in a format that can be easily understood by the decision-maker or other end users. The primary goal is to prepare the data for future use. This is an important part of the whole project and the time needed should be planned and funded. The data must be archived and print materials must be available to the users. Metadata must be completed and verified. This is a good time to review the entire project and ask what was learned, how were problems solved, and to look at the methods used and determine which were most efficient and which were the most effective.

Scale and resolution

Scale and resolution are two specifications that are central to every decision concerning the collection and analysis of data to be used to map vegetation or other resources. These specifications will normally be considered during the second step previously discussed. They are discussed here in more detail because a very basic and clear understanding of these concepts is important to project management and success.

As the need for larger scales and greater resolution increases, the cost of the data increases. The required level of detail needed (which is a user-defined need) must be honestly evaluated so that the data collected are just sufficient to provide the information needed by the decision-maker. The goal should be to identify the minimum scale and resolution needed that will permit resource specialists to interpret the data for the decision-maker. In other words, the questions to be answered must be linked to scale and resolution.

Scale and resolution are not the same. While they are loosely correlated, scale refers to the relation between a measure on the image and a comparable measure on the ground. It is usually expressed as a phrase such as '1 inch equals 1 mile' or by a representative fraction, 1/15 840. Resolution is the character of data or an image that limits the ability of a user to detect and identify an object or feature of interest on an image or within data. Resolution falls into four classes: spatial, spectral, radiometric, and temporal.

Spatial resolution is a measure of sharpness or fineness of spatial detail and it determines the smallest object that can be identified in the data. For digital imagery, spatial resolution is controlled by the pixel size of the sensor. It is roughly analogous to the term 'grain' in photographs. In vegetation mapping, the minimum mapping unit determines the minimum spatial resolution that a user needs. Spatial positioning of the data and the relative ability to detect objects or features can be enhanced through image preprocessing and enhancement operations.

Spectral resolution is a measure of the specific wave length intervals that a sensor can record. Both photographs and digital sensors are sensitive to a certain wavelengths of light, both visible and invisible to the human eye. If there is

a need to discriminate between a wide variety of plants and ground cover, greater spectral resolution may be needed.

Radiometric resolution is a measure of a sensor's ability to distinguish between two objects of similar reflectance. It is the ability of a sensor to finely discriminate between reflectance values. If mapping vegetation requires data that will show the subtle differences between plants of similar color, greater radiometric resolution will be helpful.

Temporal resolution refers to the time period between 'visits' by the sensor over the same area. Temporal resolution does not describe a single image but rather a series of images as they are captured by the same sensor over time. The temporal resolution of a satellite depends upon orbital characteristics. The temporal resolution of aerial photography depends upon when the flight planners for aerial photography missions schedule the repeat coverage.

Selecting the correct imagery and data

This step will follow logically from a well developed plan. All the issues raised in the previous steps form the foundation for choosing imagery and data. Before any data can be selected, however, the classification scheme must be defined. This is possibly the most important activity in the entire process and requires special attention by all concerned. We recommend that classifications be designed using published standards such as those currently outlined by The Federal Geographic Data Committee (FGDC) (FGDC 1997).

The development and maintenance of standard classification schemes for vegetation and other resources is a dynamic activity and users should track current literature for the latest progress on the subject. Specifically, the Ecological Society of America (ESA) is developing a national standard for vegetation classification (ESA 2000). In 1997, U.S. Interior Secretary Bruce Babbitt endorsed the FGDC National Vegetation Classification Standards. Users should note however that the FGDC standards focus on the physiognomic levels of vegetation classification and do not include an actual classification of floristically defined vegetation types. The initiative by the ESA is to develop a standard to help identify, describe, and document thousands of floristically based types nationwide. This requires the establishment of standards for plot data, nomenclature, and a review of proposed named units.

Our classification schemes were based upon the National Hierarchical Framework of Ecological Units (Avers *et al.* 1996).

There are several characteristics of a well designed classification scheme that should be considered as a classification is designed. These characteristics are reviewed in the following paragraphs.

Be exhaustive. All possible conditions must be assigned to a class. Consider the full range of possibilities in the project area.

Be mutually exclusive. Any vegetation condition must fall into to one and only one class. (We recognize however that non-mutually exclusive classes are potentially useful when using fuzzy logic.)

Be flexible. A flexible classification allows the production of many different specialty maps based on the data and it can easily respond to changing needs and definitions. For example, basing the classification on the National Hierarchical Framework of Ecological Units will build in flexibility.

Focus on basic vegetation characteristics. A flexible database contains information that is as objective as possible. Avoid interpreted and relative terms.

Define classes based on what the sensor measures. A classification should be defined with the aerial perspective of the remote sensing device in mind.

Focus on the land cover, not on land use. When using digital imagery, be wary of mixing land cover and land use in the same classification. A digital image provides a picture based on land cover reflectance values that can be classified. All land use is interpreted.

Compartmentalize for higher accuracy. Use a multi-layer approach to exploit strengths and work around the weakness of imagery. Classified layers, such as crown closure, size class, and species, should be individually assessed for accuracy.

Use a hierarchical scheme. It is possible that a classification has not accurately separated all classes. If this happens, the hierarchical scheme can be collapsed to the next layer for a more general analysis.

Define classes that are also discernible from photo interpretations and fieldwork. One test of a robust classification is that groups that can be separated on the imagery are also distinct in the ancillary data.

Use of reference data

In all vegetation mapping projects, some kind of reference data is needed to help interpret the data and classify the land cover and to assess accuracy. In remote sensing projects, reference data serve two main purposes. First, the data establish a firm link between vegetation on the ground and variations in reflectance values in the image or data. This clear link is necessary for assigning image pixels to discrete land cover classes in the image classification phase. Second, reference data are necessary to assess the accuracy of a map and establish its credibility for a decision-maker or other user.

Reference data are used by the image analyst to 'train' the computer software program so that it will assign a set of pixels to a specific classification category. Reference data come from training sites clearly identified on the ground. Training sites are representative areas of land cover that can be identified and accurately located both on the digital image and on the ground. An image-processing program can

separate some classes only if the sites are first identified and labeled by the analyst. Accurate identification is dependent upon data gathered by trained field crews.

Training sites are precisely located on the ground. Under the best of circumstances, reference data should be collected on the ground as close to the time that digital data are collected as is practicable. Data previously collected by inventory teams may be used but they should be verified by ground visits. Training sites are selected by people with knowledge of the area being studied. Sites are selected after the mapping classification scheme is developed and data are collected on-site using a data reference form based on the classification scheme. Field teams should collect all the information needed to make the data useful to the image analyst. On-site measurements must be accurate and adhere strictly to standards specified in the project plan.

With regard to collecting field data, issues of observer variability in recording data must be recognized. For example, observer variability has been shown to be the single largest factor in the reliability of bird counts (Bibby *et al.* 1992). Since the same variability may affect the collection of all resource data, planners should be aware of the potential problem and address it during the establishment of inventory procedures.

Training sites should be homogeneous with regard to vegetation and land cover attributes such as percentage canopy cover and the distribution of species or size classes. Training sites should be identified to illustrate all the variations that may be encountered within one category of the classification scheme. As far as possible, locate training sites in more accessible areas so that field time is efficiently used. Locate training sites in groups to minimize the need for travel time to reach the area. A training site should be large enough so that it can be identified and accurately located on the digital image but small enough to minimize variation of vegetation in the site.

Other valuable information can be found in related project documents, previous studies, and in the current literature. Team members should search the files for past inventories and reports and talk to other workers who may have important information about the area being studied. A search of current literature should be made to identify work that may be critical for an accurate map of the area of interest. Such data must also be evaluated for accuracy before it is used in the analysis.

The classification process

Image classification is the process of assigning pixels of an image to categories or classes usually based upon spectral reflectance characteristics. Usually, the objective of image classification is to obtain a thematic map for some land cover characteristics such as forest, water, croplands or land uses in urban areas. It is beyond the scope of this paper to present a detailed discussion of the methods of image classification. Our purpose here is to simply review the process so an interested user can move comfortably to more detailed subject matter.

Image classification results in the assignment of digital image data, the individual pixels, to categories that were defined in the project classification scheme. There are two basic ways to do this: visual interpretation and computer-aided image classification.

Visual interpretation is relatively simple. A hard-copy print or an on-screen interpretation of a digital scene can be visually interpreted in ways almost identical to the way aerial photographs are interpreted. The main advantage of doing this is that interpreters can use their personal knowledge of the area to analyse complexities in the landscape that would be almost impossible for a computer to interpret. This type of approach is particularly useful for broad area assessments where the classification of individual pixels would not be appropriate. For viewing and displaying imagery, there are specific digital band combinations that most effectively highlight certain mapping features.

Computer-aided image classification is the second basic method and there are numerous approaches using this technology. There are two broad phases in the classification process. The first phase is data preprocessing and the second phase is data classification.

Preprocessing is the first phase in any computer aided classification project. The raw digital data are prepared for use by the image processing software. This work increases the accuracy, the consistency, and the ease of interpretation of subsequent imagery. The distributor of the data may take some of these basic initial steps before the user receives it. A few of the most important data preprocessing steps are reviewed in the following five paragraphs.

Radiometric correction reduces or eliminates variations in an image that result from sensor anomalies or environmental conditions such as atmospheric haze so that image values represent as closely as possible the true reflectance of land cover features. It is an optional step and is taken if the project needs to have true reflectance values or if image quality is severely affected.

Geometric correction is the process of reorienting an image to compensate for the Earth's rotation and for variations in the position and attitude of the satellite. This process may also include positioning or warping an image into a desired map projection system so that accurate measurements can be made. This step is necessary if the resulting classification products are used in a GIS with other georeferenced information layers.

Terrain correction adjusts the image to compensate for relief distortion by using digital elevation data. This correction is highly recommended if precise location is required and the study area has relief differences greater than 150 meters (about 500 feet).

Image enhancement techniques are sometimes used prior to image classification to improve the visual interpretability of an image. The objective is to emphasize radiometric and spectral differences to make them more obvious to the interpreter. To do this, techniques such as principal component analysis or edge enhancements are performed on some or all of the image bands prior to classification. This step is particularly important if the image is interpreted visually rather than by computer.

Feature selection is a method of reducing the amount of data in a multispectral image. Feature selection involves identifying and isolating individual raw (or unprocessed) bands containing the most useful data for the particular project (Swain and Davis 1978). All subsequent processing will be performed only on the selected bands. Studies have shown that the most informative band combinations include at least one band from the visible, the near infrared, and the mid infrared spectral regions (Nelson *et al.* 1984; Sheffield 1985).

The second phase in the process is data classification. Digital data are classified according to standards that are developed to help answer questions and solve problems. Digital or numerical data are meaningful but they become more useful if they are converted to maps and images that decision-makers without technical backgrounds can understand. We use an image classification process that has yielded good results on several of our projects. The general steps we take are reviewed in the following seven paragraphs.

The first step in classification is to divide the area of interest into ecologically distinct areas. These areas may be identified and boundaries established by using digital elevation models or ecological sections or subsections (Avers *et al.*1996). Subdividing the image reduces class confusion in two ways. Smaller areas are easier to investigate and interpret, and dividing the image allows important geographic factors such as precipitation and aspect to be isolated which may help formulate generalizations about land cover.

The second step is to identify training sites across each distinct area. Training sites may be identified by using personal knowledge of ground vegetation conditions. Aerial photographs of the area are also an excellent means of identifying representative examples of land cover classes laid out in the classification scheme. In order to cover the full range of the reflectance characteristics in a particular class, most classes should be represented by more than one training site. After identifying multiple training sites for each category, the analyst establishes the perimeter of the training sites on the digital image. The spectral values and other statistics of the pixels within the perimeter represent the training sites. This process of relying on the analyst's direct involvement in the pattern recognition process is known as a supervised approach to image classification, or simply as supervised classification.

The third step is to interpret the training sites by referring to aerial photographs, other reference data, and visiting the sites in the field. Field visits to collect data are the preferred method of gathering data on ground conditions. If field visits are not possible, the interpretation of a high resolution data source such as an aerial photograph may be an acceptable alternative.

The fourth step is to perform an unsupervised classification on each area to identify variation in the image not contained in training sites. An unsupervised classification is an automated process in which the computer separates spectral classes in an image. Spectral classes are spectral groups inherent in the data. The image analyst specifies a number of parameters, including the desired number of spectral classes. The result of an unsupervised classification is a map in which each category represents a unique spectral class. The unsupervised classification supplements the supervised classification. An alternative approach to image classification is to perform only an unsupervised classification. In either case, the spectral classes resulting from the unsupervised classifications need to be identified and labeled in order to attach meaningful information to the classes. Again, aerial photographs are a good medium for identifying information about the spectral classes.

The fifth step is to analyse training sites and unsupervised spectral classes using cluster analysis to develop an optimum set of signatures. One method that maximizes the advantages of combining unsupervised and supervised techniques makes use of cluster analysis. Cluster analysis aids in labeling unsupervised classes by grouping them with supervised training sites and also helps identify potential limitations of the classification process (Chuvieco and Congalton 1988). The clustering algorithm may vary but a comparison of training sites and unsupervised spectral class can provide insights into how well the spectral data will represent the desired land cover class in the classification scheme. The goal of this process is to determine which sets of training sites and unsupervised spectral classes represent 'good', 'bad' or unknown combinations of spectral and land cover information. These relative evaluations of quality are made by comparing the classifications with ancillary data or personal knowledge of the sites to verify that the classification label is accurate.

The sixth step is to perform final image classification. The statistics representing the combined supervised and unsupervised classes from the cluster routine can be used to classify the image. At this point, the 'unknown' unsupervised classes will be labeled since they have been identified on aerial photographs and possibly in the field. The result will be a vegetation land cover map composed of supervised and unsupervised classes. Some previously described steps may need to be repeated for optimal results.

The seventh step is to model and edit problem areas. The use of ancillary or reference data in the classification process can greatly improve the accuracy of a map. Examples of ancillary data include cartographic feature files, digital elevation models, forest stand examination sample data, and ecological inventory information. Ancillary data can be used to model the classification, that is to use the computer to cross reference the classification with ancillary data, at any time before, during or after the classification process. Modeling allows the analyst to draw useful inferences about the nature of the land cover.

Assessing accuracy

Accuracy assessments are essential to all remote sensing projects. They enable the user to compare different classification methods and they provide information about the reliability and usefulness of remote sensing data for a particular project. Most importantly, the relative accuracy of geospatial data that are used in decision-making must be known and understood so that the limits of reliability can be communicated. Quantitative accuracy assessments can be time consuming and expensive but they are an integral part of any digital image classification project.

Quantitative accuracy assessment depends upon reference data. Reference data is sometimes referred to as 'ground truth'. It is well known information with a very high accuracy and precision and is theoretically 100% accurate. Accuracy assessment involves comparing the classified data for sites to the reference data for the site.

An error matrix is the standard method of presenting the results of an accuracy assessment (Story and Congalton 1986). It is a square array in which accuracy assessment sites are tallied by both their image determined category and their true category according to reference data. The matrix shows where the classifications agree and also reveal the nature of errors in the classified map in the form of errors of inclusion (commission) or errors of exclusion (omission). High numbers of inclusion or exclusion errors indicate spectral confusion. It is relatively simple to determine numerical measures of overall accuracy, of producer's accuracy, and of user's accuracy all from the matrix.

Conducting an accuracy assessment is a multi-step process. Successful completion requires a full awareness of errors in the reference data, of incorrect location of accuracy assessment sites, of changes in vegetation between the time the reference data were collected and the time the digital or satellite imagery was collected, of variations in the interpretation of reference data, and of errors in the classified map.

The eight steps for completing an accuracy assessment are briefly described in the following paragraphs.

Step 1 is to develop a statistically robust sampling scheme for the accuracy assessment. The scheme selected should have an element of randomness to help eliminate interpreter bias.

Step 2 is to choose appropriate reference data. These are the same reference data previously discussed. The highest quality reference data should be used.

Step 3 is to precisely delineate accuracy assessment sites on the reference data. Location accuracy is critical. Typically, accuracy assessment sites are delineated on aerial photographs. The standards for these sites are the same as for sites selected for image classification.

Step 4 is to interpret the assessment sites from the reference data. The accuracy assessment interpretation must conform to the same classification scheme as that used to produce the vegetation map.

Step 5 is to compile the classified data for accuracy assessment sites. Again, accuracy assessment sites must be precisely located on the classified image. The goal in this step is to develop a label for the classified map area that is consistent with the accuracy assessment site label.

Step 6 is to perform quality control. While we list it as a separate task, in practice it is an ongoing and iterative process. Errors in accuracy assessment will appear as errors in the classification and will result in an underestimation of the classification's accuracy.

Step 7 is to construct the error matrix. It is created by tallying each accuracy site according to its accuracy assessment label and classification label. Many commercial image processing systems are now providing modules to help users create and analyze error matrices.

Step 8 is to summarize and present accuracy assessment results. In all cases, the error matrix and a discussion and analysis of the accuracy results should become a part of the classified map to prevent inappropriate uses. A synopsis of accuracy results should be included in the legend of all hard copy map products.

A recent innovation in accuracy assessments is the use of fuzzy sets and fuzzy logic. Traditional accuracy assessments as we describe here suffer from some inflexibility. Fuzzy logic helps users to recognize and evaluate data based on the fact that resource data most often are on a continuum of characteristics and may easily fall between categories established in the classification scheme. Fuzzy logic is designed to handle this ambiguity and should be considered as part of any accuracy assessment of a very complex or potentially ambiguous classification. It permits the user to rate the seriousness of errors and the absolute correct or incorrect classification of polygons. An assessment using fuzzy logic can rate a site as absolutely wrong, understandable but wrong, reasonable or acceptable match, good match, or, absolutely correct (Gopal and Woodcock 1994).

Deriving polygon layers and other GIS data

One advantage of using digital satellite data and imagery is that the derived classifications are digital and are ready to be used in a GIS. Frequently, resource managers prefer a

classification composed of polygons because they are more comfortable making decisions based on the old familiar polygon units rather than on an array of classified pixels. Some GIS applications and models require both the polygon and the pixel format. By superimposing polygons over a pixel classification, the user can determine the amount and location of variability in each polygon.

A digital image is a raster format file. A polygon map is created from a digital classification through a technique called filtering. Filtering is a term that refers to the removal of spatial features for image enhancement (Jensen 1986). The filtering process smoothes the raster data into polygon-like groups. Users of contemporary GIS software have access to may types of filters each with numerous options.

The traditional approach to polygon creation involves classifying a data set so that all vegetation characteristics are in a single file. A filter is applied to the single layer classification and the file is converted to a vector file as polygons. An alternative way is to use a hierarchical approach in which multiple layers, such as crown closure and species, are filtered individually but are later combined in an operation to produce polygons of similar vegetation. This process results in a natural classification more in line with the results of what a photo-interpreter would do.

Vegetation polygons created with filters can be used in a GIS for spatial information and can be merged with or compared to other thematic layers such as managed tracts of forests or historical data to gain more understanding of the vegetation being mapped. There are many techniques available to help create and understand vegetation maps. Polygons can be linked to the raster, or raw, data and, in a GIS relational database, can be analysed to determine the number of pixels in the polygon. Polygons can be labeled based upon the number of pixel distribution within the polygon.

Information about differences within individual polygons can be extracted for all the pixel data including crown closure, structure, size and species. Summary tables can be created from this information and a database designed to provide access. In this way, both the generated vector polygons in the GIS and the original classified pixel data can be used simultaneously, providing valuable information for vegetation analysis and management.

Another common application of using digital data in a GIS is the construction of a change detection layer. Over time, features on the ground change. With repeated satellite coverage, one image can be compared to another image from a later date to quantify and analyse changes that have occurred. Even more understanding can be gained by using a series of three or more images so that trends over time can be quantified. There are some limitations that must be considered. For example, comparing images from different seasons and those with cloud and shadow areas will affect the final product. (This is a focus of case study number 2 which follows.)

Case studies with examples of remote sensing applications

The principles outlined previously in the discussion of 'Project design and analysis' formed the framework for our analysis in each of the following case studies. While each step is not perfectly clear in these outlines, the process was used to guide each project.

Case 1. A large area analysis; vegetation and mapping ecological units

Objective: The objective of this project was to produce a land cover map with ancillary data that would guide forest managers in the implementation of ecosystem management.

The Bridger–Teton National Forest in western Wyoming is using remote sensing in support of ecosystem management. Forest managers started the Ecological Unit Inventory to map and integrate information on landform, geology, potential natural vegetation and soils. Such information is used to characterize and map ecological units that have interpretive values for management.

This project covered an area of about 900 000 ha (2 224 000 acres). Classifications were based on the National Hierarchical Framework of Ecological Units, which is a standardized mapping system for stratifying the Earth into progressively smaller areas of increasingly uniform ecological potentials (Avers *et al.* 1996). Digital satellite imagery was used in this project because it provides the broad perspective necessary for a study area of this large land area.

Ecological Unit Inventories have focused on a premapping protocol in which Land Type Associations (LTA) are first identified for the study area. This is an important first step in identifying and delineating smaller ecological units. Within the LTA, smaller premap units establish preliminary ecological unit boundaries and to a certain extent determine where field investigations will take place. Premap boundaries are refined into the final ecological unit boundaries after extensive fieldwork and stereo analysis of aerial photographs.

This type of inventory is more complex than a straightforward vegetation classification project. In the early phases of the project, two 7.5 minute USGS Topographic Quadrangles were used to test the use of slope, aspect, and spectral classes for creating ecological units. Manual and automated approaches using digital data and digital elevation models were used to identify units. This phase permitted the refinement of techniques that were subsequently applied to the entire project.

A review of the premapping products indicates that subdividing the survey area into LTA units created more meaningful results with both spectral and digital elevation model products. Fewer misclassifications of vegetation types and more useful slope classes were produced than with the automated method tested on the two preliminary topographic quads in the early phase of the project. Also, methods

developed for the premapping process allowed for the production of a field ready premap in about half the time that it would have taken to create the map from aerial photos alone.

Mapping vegetation was only one part of the process. In this project, vegetation polygons were created independently using the vegetation classification derived from Landsat TM for each LTA. Land cover and vegetation classes were grouped whenever possible. Then, filtering (scanning) techniques were used to break out heterogeneous and homogeneous vegetation. Small polygons were eliminated in the scanning process to create a final vegetation map with a 20 ha (49.42 acre) minimum mapping unit for each LTA unit.

The entire process was highly interactive. At each step, products were evaluated and decisions made by the GIS analyst and scientists familiar with the region. For this reason, it would be difficult to fully automate the premapping process. The interactive process allows for flexibility in methods and direction. Flexibility was important so that scientists could consider parameters and relationships that may be unique to the particular area under study.

Case 2. Forest plan revision; monitoring and detecting change in vegetation cover

Objective: The primary objective of this project was to apply change detection procedures, using digital satellite imagery, to monitor the forest plan. Forest land management plans typically set parameters on the extent and degree of change that would be acceptable. Managers wanted to compare actual change in vegetation resulting from disruptions to the degree of change predicted in the forest land management plan. Results were used to illustrate to the public what progress was being made through forest management to reach the desired future conditions envisioned in the plan as approved in 1986. In this project, remote sensing technology and image processing were used to determine the change in land cover over time.

Forest planning is a major activity on every national forest. It is a major expense and efforts are made to minimize the impact on budgets. The Mark Twain National Forest in southern Missouri used remote sensing data in an assessment of change in vegetation conditions over time as part of their effort to monitor progress in their forest plan. The value of remote sensing data is that it can provide snapshots of vegetation conditions at two time periods that can be compared to assess changes that have occurred.

Changes in vegetation were tracked using a new process developed during the plan. For example, clear-cuts show up as ‘major change’ and thinnings show as ‘moderate change’. Size limits of harvest openings can be determined and compared to the maximum size allowed. The total area in the 0 to 9 year age class and percentage of these areas could be determined and compared to the goal of that management unit. Distribution of these activities was also very important to managers and this was monitored using other methods developed in this study. Results of the image analysis were in the form of a digital data layer of ‘change’ in a GIS. This layer, in addition to the other data layers of ownership and management unit prescription, was used in a number of other analysis and modeling processes. Key results of the project analysis could be visually displayed and shared with the public.

In this analysis, Landsat thematic mapper (TM) images from August 1982 and August 1989 were compared. The images were goecoded and terrain corrected prior to use. A visual inspection of image registration was done by inspecting individual bands for the different dates on a computer display. In the Midwest United States, clouds routinely cause problems with change detection analyses. In this project, we created a cloud mask from an unsupervised classification. Areas classified as cloud and cloud shadow were removed from further study. This removal created gaps in approximately 10% of the study area.

Band 7 from each image was used in the analysis. Band 7 from each image was combined to produce a new two-band image. An unsupervised classification of the combined band 7 image was completed with 35 spectral classes. That classification was aggregated into five categories that represent various land cover changes. The classes defined for this study are (1) No Change; (2) Moderate Change; (3) Major Change; (4) Regrowth; and (5) Open Areas. The evaluation and labeling of the unsupervised classes on the change-detection image were based on two types of ancillary data as well as on an analysis of the unsupervised cluster means. Ancillary data consisted of forest stand information and visual identification of each category by expert personnel either on the ground or on large scale aerial photographs.

The five classes of land cover change classification were subjected to a series of filters in order to create homogeneous classes. For example, areas of moderate change were frequently composed of pixels of multiple change types. After filtering, the same area was a homogeneous area of moderate change. Vector polygons ere created from the filtered classifications and polygons less than 2 ha (5 acres) were eliminated from further analysis.

An accuracy assessment was conducted using stratified random sampling techniques over the entire project area: 164 points were generated across all five categories. The points were checked on the ground and, in instances where accessibility was a problem, assessment sites were interpreted from aerial photographs or were compared with the forest stand inventory database.

The use of thematic mapper data for classification of land cover change worked well. Methods used were easy to understand and were adaptable to larger areas of land. The

cost of implementing this process across the study area was less than 50 cents per ha (20 cents per acre). The accuracy assessment for the study area showed that 138 sites out of 164 sites were classified correctly for an overall thematic accuracy of 84%. While the accuracy assessment process was unbiased and technically sound, it created some problems in application. The randomly generated sample points sometimes were so close to the boundary of a polygon that it was difficult to determine on the ground which change condition to sample. Other problems resulted from inadequate ancillary baseline information from the early 1982 image. This made it difficult to decide whether the change category assignment for some polygons was correct.

The results of this project show that changes in vegetation can be accurately detected and classified in a rather straightforward procedure. The derived classifications are readily useable in a GIS.

Case 3. A project level analysis; range allotment mapping

Objective: The objective of this project was to gather cost-effective and timely information needed for managing the rangeland resource. On the Beaverhead–Deerlodge National Forest, Fisk *et al.* (1998) used remote sensing data to produce a land cover map for 19 range allotments in the northern Gravelly Mountains of south-western Montana. Vegetation cover maps for this area was last produced by hand, based on field sampling and aerial photographs in the late 1960s. The new land cover maps were used to evaluate and revise allotment management plans.

Since at least the early 20th Century, federal grazing policy has stirred controversy. Grazing regulations were created to settle disputes over the land's carrying capacity and to prevent overgrazing. The public land was divided into parcels called allotments. Today, the USDA Forest Service administers over 9400 range allotments. Allotment management plans are prepared to regulate and monitor grazing use.

The most basic need in range management is an understanding of the quantity and quality of vegetation on the range. Image analysts obtained information from several sources including Landsat Thematic Mapper satellite imagery, digital elevation models (DEMs), digital orthophoto quadrangles, cartographic feature files (CFFs), 1:16 000-scale natural-color aerial photographs, and from four flight line transects of large-scale color-infrared digital camera imagery.

First, range lands were separated from forest lands. This refinement reduced the project area by half. Landsat TM satellite imagery, digital orthophoto quadrangles, aerial photography, and digital camera imagery was used to further define range vegetation areas. Initially, 50 unique spectral classes of range vegetation were generated using Landsat TM imagery in an unsupervised classification.

In addition, large scale 'resource' photographs and digital camera images were used to identify over 300 sample sites of range vegetation types. These sites were transferred to the Landsat TM imagery and their derived spectral characteristics were combined with the original 50 unsupervised classes. Using spectral-analysis software to eliminate redundant and anomalous sites, 85 types were selected for a supervised classification. This automated process groups each pixel in the image with the spectral class that most appropriately matches its characteristics. This process was refined until the range specialist concluded that the spread of classes was satisfactory.

Mapping range vegetation using only the spectral characteristics of satellite imagery has some shortcomings. Physical landscape characteristics such as elevation, slope, aspect, and hydrologic features can be used to further refine the vegetation map. Layers were derived from DEMs and CFFs and these were used to model ecological conditions associated with specific land-cover types. The final map was reviewed using digital orthophoto quadrangles, aerial photography, and digital camera imagery. Resource specialists familiar with the area also reviewed the map and corrected mislabeled pixels.

Range resource specialists used 1:16000 aerial photography to assess the accuracy of the range vegetation maps. Specialists selected 449 sites from the map, overlaid them on digital orthophoto quadrangles for precise location, and transferred them to aerial photographs for interpretation. A range specialist interpreted the sites in stereo and recorded the results. These photo interpretations were used as reference data. An error-matrix table was used to compare the reference data with the range map. The accuracy assessment indicates that any site on the map had approximately a 74% chance of being correctly identified.

Resource specialists are using this information to stratify future field-sampling efforts, assess vegetation conditions, analyse allotment alternatives, and facilitate public involvement. The land cover map is used to estimate allotment forage potentials and to develop suitable range models for range allotment grazing systems. Forage productivity estimates ranging from high to low are assigned to each vegetation cover class based upon ancillary data. These estimates are used to calculate range allotment carrying capacity. Productivity estimates can then be compared to the current grazing system and the number of livestock permitted. Because all of the information is in a GIS, the range models can be adjusted to reflect new information as it is gathered.

Case 4. Burned Area Emergency Rehabilitation (BAER)

Objective: The objective of this project was to rapidly assess an area burned by a non-prescription wildfire to determine burn intensity and the effect on vegetation and soil. The after-effects of wildfire can be as devastating (or in some

cases, as beneficial) as the fire. Even before a large wildfire is controlled, a quick assessment of watershed conditions is made to determine whether life, property, and natural resources are at risk from weather events.

During a Burned Area Emergency Rehabilitation (BAER) assessment, scientists gather information on fire-altered watershed conditions. The objective is to complete inventory, analysis, and rehabilitation prescriptions before a heavy rain falls on the burned area. Because burn intensity affects post-fire flooding potential, mapping is a critical step in the survey. Typically, burn intensity is mapped by combining on-the-ground measurements of such elements as effective ground cover reduction, soil aggregate stability reduction, and hydrophobic soil development, with overlook-point or aerial sketch-mapping. BAER teams composed of resource specialists typically are a standing part of an organization and are called on as needed.

The rough maps they create are used to measure burned areas by intensity class, to estimate potential sediment yields and to locate treatment areas. Although the maps may be refined as more detailed information becomes available, the initial versions are often digitized and placed in a local geographic information system (GIS) for long-term planning. Many demands are placed on a burn intensity map. The map must be prepared quickly and at low cost. It must be available in hard copy and GIS layers and it must be available at varying scales for overlay and display.

During the summer of 1996, we demonstrated a new way to map burn intensity on the Fork Fire in northern California. A Kodak DCS 420 color infrared digital camera with global positioning system (GPS) was used to 'image' the demonstration area. Each image is a 1524×1012 array of picture elements (pixels). Mounted in a small aircraft flying 12 000 feet above ground level, each image covers an area approximately 9100×6060 feet. Each pixel covers an area of about 9×9 feet. Images are acquired every 3 seconds and are stored in digital format on a 'card' that can hold about 200 scenes.

Before the Fork Fire was controlled, flight lines were laid out on a burned 30 000 acre area of the Middle Creek watershed. Flight lines were entered into GPS navigation software to assist airplane navigation and to ensure proper image coverage. A linked GPS unit computes latitude, longitude, and elevation for each digital image. Location data are used to plot flight lines which, combined with other GIS layers, determines the relationship of images to landscape features. Ancillary data about prefire vegetation, soils, and topography came from the corporate geospatial database of the Mendocino National Forest and the Forest Service Remote Sensing Laboratory in Sacramento, California.

Acquiring digital images gave resource specialists more time to make ground observations. Their GPS data recorders were loaded with a soil dictionary to help record important burn intensity parameters. Observations included the amount and degree of litter consumed, depth and color of ash, changes in soil structure and crusting, the degree of fire caused water repellence, the weight of live fuels consumed, and site characteristics such as percentage of surface rock, prefire vegetation type, slope, aspect, soil type and, GPS location.

Processing the digital camera imagery included georeferencing 110 digital images to a common base and assembling them into a seamless mosaic of the area, which was registered to a SPOT panchromatic satellite scene. The burn intensity map for the Middle Creek Watershed was prepared by manually delineating polygons of burn intensity on the image mosaic. A digital version of the mosaic was used on a computer workstation to display specific locations in more detail. The minimum mapping unit for polygons was about 20 acres with most polygons in the range of several hundred acres.

Comparing digital images to prefire vegetation and to topographic information in a GIS was helpful. For example on the image, burned shrub communities appeared similar to severely burned forest. Ground inspection of burned shrub communities indicated that the burn intensity was moderate. Overlaying the image on the prefire cover type map in the GIS let users distinguish moderately burned areas from high intensity burn areas.

Having the digital map early in the process was an advantage. The digital map permitted the rapid identification of critical burned areas within the fire perimeter and the team could focus limited and expensive field time. The wildlife biologist immediately identified high intensity burned critical habitats. The hydrologist quickly located and calculated the percentage of watershed burned at high intensity.

Specialists used the burn intensity map to compute vegetation cover and runoff data that was used in erosion and runoff models. Soil scientists overlaid the burn intensity map on soil erosion and slope data layers and quickly determined various burn intensity soil–slope combination figures needed to predict potential soil loss.

During this project, two burn intensity maps were created, one based on traditional methodology and the second based on the collection of digital data, so that accuracy could be compared. Based on a number of ground observations, the digital image process resulted in an estimated 20% improvement in accuracy. In addition to this improvement in mapping accuracy, the Forest Service estimates it would have saved a quarter of a million dollars on the Fork Fire compared to the cost of traditional BAER assessment methods (Hardwick *et al.* 1997).

Summary

The use of remote sensing data in the classification and mapping of vegetation is becoming the primary method of

assessing the quality and quantity of natural resources. A well designed project that specifically answers questions posed by decision-makers will yield information that can improve the timeliness and quality of resource allocation decisions.

Mapping project managers must be innovative and adaptive leaders with all the skills necessary to overcome the unexpected. They must also have a strong resource background and be team players.

Vegetation mapping requires the combined efforts of a team of specialists in resource management and in computer data processing. The image analyst is a crucial member of the mapping team. The analyst's understanding of both the imagery data and the resources being mapped is not only helpful for performing image processing functions, it is critical for getting meaningful information from the project. A basic knowledge and understanding of both the virtual world and the real world allows the analyst to contribute to the design of an appropriate classification scheme and to participate in field data collection.

Acknowledgements

Much of this paper is based on work and studies previously published by the USDA Forest Service, Remote Sensing Applications Center. We acknowledge the efforts of Jan Johnson, Vaughan Landurm, Jay Powell, Vicky Johnson, Tim Wirth, Hanns Fisk, Paul Hardwick and Stan Bain. We also thank our peer reviewers, Lloyd Queen and J. Michael Scott, for their time and help with our manuscript.

References

Avers PE, Cleland DT, McNab WH, Jensen ME, Bailey RG, King TK, Goudey CB (1996) 'National hierarchical framework of terrestrial ecological units.' (USDA Forest Service: Washington, D.C.)

Bibby CJ, Burgess ND, Hill DA (1992) 'Bird census techniques.' (Academic Press: London) 257 pp.

Chuvieco E, Congalton R (1988) Using cluster analysis to improve the selection of training statistics in classifying remotely sensed data. *Photogrammetric Engineering and Remote Sensing* **54**(9), 1275–1281.

ESA (2000) This document entitled Review Draft 6.0 of 'An initiative for a standardized classification of vegetation in the United States', presents the proposed national standards for the floristic levels of vegetation classification in the United States (associations and alliances). It is a product of the Vegetation Classification Panel of the Ecological Society of America. It may be viewed on the ESA Web site at http://esa.sdsc.edu/initv60.htm.

Fisk H, Lachowski H, Suzuki K, McNamara J, Brohman R (1998) Updating Range Allotment Management Plans using remote sensing and geographic information systems. USDA Forest Service, Washington D.C. Remote Sensing Tips. October 1998. RSAC-7140-7.

FGDC (1997) 'Content standard for digital geospatial data.' (The Federal Geographic Data Committee: Washington, D.C.)

Gopal S, Woodcock.C (1994) Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing* **60**(2), 181–188.

Hardwick P, Lachowski H, Maus P, Griffith R, Parsons A, Warbington R (1997) Burned Area Emergency Rehabilitation (BAER): use of remote sensing and GIS. USDA Forest Service, Washington D.C. Remote Sensing Tips. August 1997. RSAC-7140-01.

Jensen J (1986) 'Introductory digital image processing: A remote sensing perspective.' (Prentice-Hall: Englewood Cliffs, NJ) 379 pp.

Nelson R, Latty R, Mott G (1984) Classifying northern forests using Landsat thematic mapper simulator data. *Photogrammetric Engineering and Remote Sensing* **50**(5), 607–617.

Sheffield C (1985) Selecting band combinations from multispectral data. *Photogrammetric Engineering and Remote Sensing* **52**(2), 223–227.

Story M, Congalton R (1986) Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing* **52**(3), 397–399.

Swain P, Davis S (1978) 'Remote sensing: The quantitative approach.' (McGraw Hill: New York)