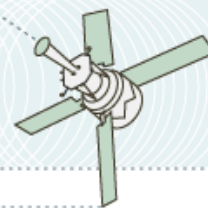# Remote Sensing Resources

Land cover classification methods

This document provides an overview of land cover classification using remotely sensed data. We will describe different options for conducting land cover classification will be described including types of imagery, methods and algorithms, and classification schemes. Using this guide, you will learn how to carry out a land cover classification project, while gaining an appreciation for the image classification process. Land cover mapping is not as difficult as it may appear, but there are many decisions, choices, and compromises regarding image selection and analysis methods that must be made. Although it is beyond the scope of this guide to provide details for all situations, you will at least be able to assess your own needs and requirements. If you lack experience with land cover mapping we suggest that you work with someone with sufficient experience.

What is the difference between land cover and land use?

Although the terms "land cover" and "land use" are sometimes used interchangeably, they are actually different. Simply put, land cover is what covers the surface of the earth and land use describes how the land is used. Examples of land cover classes include: water, snow, grassland, deciduous forest, and bare soil. Land use examples include: wildlife management area, agricultural land, urban, and recreation area. Some classified maps include a mix of land cover and land use. Although there is nothing inherently wrong with this approach, it can sometimes lead to confusion. It is important that each class on the map be clearly defined and distinct from other classes.

Some maps include both land cover and land use information. One way to do this is to represent land cover using different colors, while using different patterns or symbols overlaid on the map to represent land use.

Where do I start?

A good place to start is to determine if you really need a classified map. You might want to read the guide titled: Justification for using photo interpretation methods to interpret satellite imagery to see if your needs can be met simply using image photo products instead of a classified map. Should you decide that a land cover map is necessary, you will need to have a good understanding of why the map is being created and how it will be used. Reading this document should provide sufficient information regarding factors to consider for a needs assessment. You should write down the the results of this needs assessment. It will become obvious that there are several decisions that you need to make based upon what the map will be used for. Accordingly, these decisions will determine many of your choices throughout the classification process. The next several sections detail the steps you will take during a land cover mapping project.

What do I need to create a land cover classification map?

There are several options for creating land cover maps. The following is a list of a few data and interpretation alternatives that can be considered for mapping land cover:

- Visual satellite image interpretation
- Digital satellite image interpretation
- Aerial photo interpretation using stereo photos
- Aerial photo interpretation using single uncorrected photos
- Aerial photo interpretation using orthophotos
- Field surveys using simple angle and distance measurements
- Interpretation of videography
- Interpretation of small format photography

This guide focuses on approaches using remotely sensed data but you can apply many of the steps to other data sets.

First off, creating a land cover map using remotely sensed imagery necessitates suitable imagery. There are several choices available and the process of selecting the appropriate imagery is described in the section "How do I select the imagery I need?"

You will also need some way to visualize and process the imagery. If the imagery is in a printed form then viewing it is relatively straightforward and the classification process is limited to visual interpretation methods. However, if the imagery is in digital form you will need software to view and process the imagery. Software required for classifying imagery can range in price from free to tens of thousands of dollars. Some software options are listed in the Net Resources section of this web site and in the Viewing Data theme. There are several options available when using digital imagery and these are detailed in the section below titled "Which classification method should I use."

You will also need a certain skill level to create an accurate map. Arguably the most important skill is the ability to associate the features that can be seen in an image with what is on the ground. This ability comes from experience. A good way to begin learning this skill is to look at imagery from an area that you are familiar with and begin identifying features on the image based on your recollection of what you know exists on the ground. The importance of this type of experience cannot be over stated.

In the classroom emphasis is usually put on the mechanics of automated classification. This is certainly important information for a remote sensing practitioner, but understanding the details of how an algorithm works or the physics of remote sensing is not required to produce an accurate land cover map. There are some simple methods that work very well as long as the analyst has the ability to reliably identify features in the image.

Lastly, whether you are just learning image classification or you have been through some formal training such as workshops or university classes, it is always a good idea to get some feedback from experienced colleagues. Discussing your plan of action and periodically showing your progress to other remote sensing practitioners is very helpful. They can sometimes suggest refinements and offer other options that will help improve the land cover map.

Define the study area

Once you know how the map will be used, you need to define the region that will be mapped. This can be a fairly easy task but in some cases it can be difficult to reach consensus among project participants. For example, you might find that adding a small portion to a study area will involve the purchase of additional imagery and therefore increase the required resources to conduct the classification. After discussing this situation with project participants you may decide that the additional cost to include the small portion of the study area cannot be justified and the study area will have to be modified. When delineating the study area it may be important to include an area adjacent to the primary study area. This could be done to better understand the transition to the area outside of the study area. An easy way to do this is to create a buffer around the primary study area so adjacent areas are included (Figure 1). Buffers can be created using common GIS or image processing software.

Which classification scheme do I use?

A classification scheme effectively defines the legend that will be used for the final map. For example: will the map show forest and non-forest or will it have several or even dozens of different categories? Should the final map categories represent land cover or something else such as land use, habitat, or conservation importance? The way in which the map will be used, and some practical realities, will dictate the content of the classification scheme. One place to start is to look at some of the common classification schemes. There are a large number of classification schemes used for land use and land cover maps throughout the world. Some of the more common schemes are listed in Table 1.

Classification name	URL
Anderson	http://landcover.usgs.gov/pdf/anderson.pdf
National Land Cover Data	http://landcover.usgs.gov/classes.asp
FAO Land Cover Classification System	http://www.africover.org/LCCS.htm

Although this guide focuses on land cover there are other attributes that can be mapped such as:

- Vegetation structure
- Land cover disturbance
- Vegetation age (for example, primary and secondary)
- Distribution of taxa (Chorology)
- Land use
- Ecosystems
- Crown closure

An important point to remember is that no matter what classification scheme is selected each class must be well defined and documented. For example, if you have a class called "Forest" you need to specify what constitutes a forest. Do the trees need to be a certain height? How dense do the trees have to be? Without this sort of information it is impossible for the user of the final map to know what the different classes represent.

When choosing an appropriate classification scheme you should decide if compatibility with existing schemes is necessary or desirable. Some advantages of using an existing system is that the classes are already defined and the map you produce can be easily compared with other maps using the same system.

Defining mapping classes is often an iterative process. A balance must be struck between the classes that are desired, based on the map's purpose, and the classes that can be accurately and economically delimited. In general greater detail translates to greater cost in time and money and lower accuracy per class. A good rule of thumb is to select the minimum number of classes that are practical. When you decide upon a classification system it is important that you document the details for each class and apply these class definitions in an objective manner.

When selecting classes, a hierarchical or non-hierarchical approach can be used. In a hierarchical approach, classes are nested such that major classes are broken into sub-classes and these sub-classes can further be broken into more detail. The advantage of such a system is that it can be easily generalized and it is easy to adapt to various scales (the coarser the scale, the more detailed the hierarchy will be). Hierarchical systems are often used when there is a need to maintain conformity with a national-level classification system. A non-hierarchical approach, however; is designed for a specific purpose with a specific scale in mind. The advantage of a non-hierarchical system is that it can be modified to suit a specific application because it is not restricted by the constraints imposed by a hierarchical system. In other words, it can be easier to customize for specific project goals since it can include a mix of detailed and generalized land cover classes.

When defining classes you must decide how to classify mixed features such as transition and mosaic classes. One approach is to define explicitly these classes as mixed or transition and the other is to ignore the fact that classes are mixed and define classes by the most populous feature on the ground within a delineated unit. A third approach that is gaining popularity is called "continuous fields mapping". With this approach each class is mapped into a separate layer (image) showing the percent coverage of that cover type in each pixel. For example, if "Conifer Forest" was a cover type you would create an image with pixel values representing the percentage of conifer forest in that pixel (Figure 2). Instead of a single classified map you would have several maps (one for each class) detailing the percent coverage for each of the respective cover types.

When deciding which approach to use to represent mixed features, you must

take into account the nature of the features being mapped and how important mixed classes are relative to the intended use of the map. You should clearly document whichever method is used.

Another point that has to be considered with the classification scheme is the spatial detail that will be mapped. A minimum mapping unit defines the smallest area that is defined on a map. For instance if the minimum mapping unit is 1 hectare then any feature less than 1 hectare would not be delineated as a unique feature. Instead it would be incorporated into another feature. Minimum mapping units can vary from class to class so more important or rare classes would have a smaller minimum mapping unit to ensure that they are not lost as a result of inclusion in another class. In some cases no minimum mapping unit is used and all recognizable features are identified. No matter what approach is used, it is important that it meets the requirements of the needs assessment and should be well documented.

How do I select the imagery I'll need?

Now that you have defined the classes you want to map, it is necessary to select appropriate imagery that will allow you to accurately define these classes. Selecting appropriate satellite imagery is, more often than not, a subjective task. The selection is often limited to the resources that are available. If the project budget is limited you may be restricted to using free or inexpensive imagery, which in turn affects the level of information that can be extracted for the map. Ultimately, this may require that the needs assessment be adjusted to reflect the practical limitations of the project.

The spatial detail of the final maps will dictate the required resolution of the basemap and this will significantly limit the list of possible image types that can be used. For example, if you were interested in identifying individual trees you would need a few fine resolution satellite images or aerial photographs. If, on the other hand, you wanted to create a global land cover map you would look for a much coarser resolution. The selection of an appropriate image resolution is somewhat of an art and experience will indicate the limits of available image types. One rule of thumb that can be used as a rough guide to select an appropriate resolution is that the area of the minimum mapping unit should be ten times the surface area represented by a single pixel. In other words, if you had a minimum mapping unit of 1 hectare a Landsat Thematic Mapper image with a 30 meter resolution would be appropriate since one hectare equals approximately 11 pixels (10,000m/30m x 30m). Perhaps a better way to get a feel for the level of detail that can be extracted from a remotely sensed image is to view it on a computer display and zoom and roam around the image to see how well individual features can be identified. If you can see it on the screen there is a good chance it can be properly classified.

The specific classes you are looking for will also have some effect on the type of image data you are looking for. For example, some sensors are better designed for collecting information about aquatic environments and others are more suitable for specific terrestrial environments. This is an area where advice from someone with experience classifying similar features would be useful. Some additional information is available in the guide titled What you need to know when searching for satellite imagery.

What about preprocessing?

Preprocessing can be divided into two categories; radiometric and geometric processing. Radiometric processing is used to reduce the effects that prevent us from measuring surface reflectance using remotely sensed imagery. This can involve reducing the effects the atmosphere has on modifying solar energy (sunlight) before and after it hits the Earth's surface and the effects due to terrain and sensor configurations. Minimizing this "noise" reduces variation in the value of a pixel due to factors independent of surface cover. When conducting automated classification, reducing these effects tends to improve classification accuracy since you are essentially reducing the noise in the image. With visual interpretation, however, this is not always the case. For example, in mountain environments there is often one side of a mountain that is brightly lit while the other side is in shadow. If not corrected, this effect will introduce problems for automated classification since the same land cover will look very different. However, with visual interpretation this illumination effect can actually help since it tends to accentuate 3-dimensional features (Figure 3). Moreover, a trained analyst familiar with the study area is able to accurately classify and identify the

vegetation even though the shadows and topography make the vegetation features appear like two distinct classes.

In the past, using radiometric correction methods required a significant level of expertise and data that was not widely available. In recent times, however, radiometric correction algorithms have become more accessible to the non-specialist and as this trend progresses it is likely that these tools will be able to improve significantly the ability of non-specialists to benefit from these corrections.

Geometric correction typically involves warping an image to match a specific map projection. Whenever an image is warped a resampling of the image pixels takes place which degrades the original data to some degree. This has more of an influence when applying automated methods since the individual pixel values are modified during resampling. This effect may or may not be significant, depending on the application, but in general it is best to perform the geometric corrections after the final classified map is produced. This is not an issue when doing visual interpretation because the changes in pixel value due to resampling are not as visible to the human eye.

What other types of data do I need?

In addition to satellite or aerial imagery, other data can be used to increase the accuracy of the classification. Some possible ancillary data are digital elevation models (DEMs) and their derived datasets (slope and aspect), climate data such as rainfall and temperature, and vector overlays such as roads, rivers, and populated places. If there is an attribute that effects the distribution of land cover that exists in a mapped form it should be incorporated into the classification process. For example, DEMs often supplement satellite data when mapping land cover since vegetation classes are often limited to specific elevations or aspects.

Sometimes, however, we might be aware of an attribute that we want to incorporate into the classification but there is not an available dataset that is appropriate for the study. For example, vegetation distribution is affected by rainfall but many rainfall datasets are too coarse to be useful in classifying vegetation. As time goes on more of these environmental layers are being improved so that they can be used to map vegetation more accurately.

Incorporating these ancillary data into the classification process is not always easy. Some of the classification methods discussed below allow the incorporation of assorted datasets but others are more restrictive. For the more restrictive methods there are still some ways to utilize ancillary data. One is to use these data to stratify the study area into regions based on one or more environmental variables. We might use slope and aspect and create unique strata for different slope and aspect ranges. For example we could create two elevation categories-- below 500 meters and 500 meters or above -- and four aspect categories-- north, east, south, and west-- and from these create the following 8 strata: below 500m north aspect, below 500m east aspect, below 500m south aspect, below 500m west aspect, 500m or above north aspect, 500m or above east aspect, 500m or above south aspect, and 500m or above west aspect. The classification can then be carried out within each of the individual strata. Defining strata will generally have a positive effect on the accuracy of a classification.

Which classification method should I use?

The classification process involves translating the pixel values in a satellite image into meaningful categories. In the case of land cover classification these categories comprise different types of land cover defined by the classification scheme that is being implemented. There are dozens, if not hundreds, of classification methods that can be used to group image pixels into meaningful categories. Unfortunately there is not a single "best" approach to image classification. The choice you make depends a lot on the algorithms that are available to you with the image processing software you use and your familiarity and experience with the different methods. To facilitate the discussion of the different methods we will group them into automated, manual, and hybrid approaches.

Automated

The majority of classification methods fall in this category. With this approach an algorithm is used to assign individual pixels or groups of pixels to one of the valid categories. One advantage of automated approaches is that the algorithm is

applied systematically throughout the entire image relatively quickly. The algorithm can also utilize many more layers than a human using visual methods. For example, using manual methods an interpreter will often be limited to viewing a 3-band image whereas using automated methods it is possible to incorporate data from hundreds of satellite bands. Traditionally, automated algorithms tended to be limited to using only the pixel values in the image but there are now several approaches that allow a user to easily incorporate ancillary data layers such as elevation, slope, aspect, soil type, and a host of other biophysical layers to improve the classification accuracy.

In order for an automated classification algorithm to associate pixel values with the correct land cover category, some input from an analyst is necessary. When this information is provided before the algorithm is run the procedure is referred to as supervised classification. With this approach the user identifies sample pixels in the image that can be used as representative examples for a particular land cover category and then the sample pixels are used to train the algorithm to locate similar pixels in the image. When a supervised classification is run the result is a land cover map with all of the pixels labeled as a particular cover type (Figure 4).

The other way to classify an image is to start by letting the computer group similar pixels together into unlabeled classes (clusters) and then have the analyst label the clusters with the appropriate land cover category. This approach is called unsupervised classification since the algorithm works without a-priori input of information about existing land cover samples.

Both supervised and unsupervised classification methods can produce reliable results; however, there is greater variety in the available algorithms geared toward supervised classification. The primary difference between classification algorithms is the way in which they determine how an individual pixel is assigned to a land cover category.

We present a short overview of some of the more popular classification approaches to provide you with insight into the range of possibilities. Many of these algorithms come from the field of machine learning and they can be quite complex. Providing details of these algorithms is beyond the scope of this guide. More detailed information is available in remote sensing textbooks, training courses, or on the Internet. The best way to get a feel for how these different algorithms work is to practice applying them using remote sensing software packages. Although it is not required, an understanding of how an algorithm works will help the analyst use the selected method more effectively. Again, talking with someone with experience with the different classification approaches can be helpful in deciding which one to use.

ISODATA unsupervised classification

The ISODATA algorithm is the most common unsupervised classification tool. It is used to create a user-defined number of clusters or classes in an image that must later be labeled to create a land cover map. Before you run an ISODATA program you need to define several parameters that control how many clusters will be generated, how many iterations will be run, and other parameters that control how clusters are combined and split as the program executes.

When the program is finished you will have an image with several classes which will not be labeled. For example, if you specified that the ISODATA program should create 25 classes the output map will have 25 classes with each comprised of pixels with similar pixel values. The next step is to label these classes with the correct land cover type. If you were trying to create a map with two classes, forest and non-forest, you would look at each of the 25 classes in the ISODATA output image and label them as either forest or non-forest (Figure 5). In some cases it will be difficult to decide how to label a particular class because it will contain more than one land cover type. When this happens you can choose to run the ISODATA program again to output more classes or you can apply the ISODATA program to run only on those classes that contain multiple land cover types, a technique known as "cluster-busting".

Supervised statistical classification

There are several types of statistics-based supervised classification algorithms. Some of the more popular ones are (in increasing complexity); parallelepiped, minimum distance, maximum likelihood, and mahalanobis distance. With

supervised statistical classification algorithms the analyst must first locate and define samples in the image of each class that are required for the final map. For example if you were interested in creating a map with forest and non-forest classes you would select sample areas in the image that represent the different types of forest and non-forest. These samples are called training areas. Remote sensing software that supports supervised classification provides tools to allow users to draw lines around training areas and label them. Once a sufficient number of training areas are selected you can run the supervised classification. The algorithm then compares each pixel in the image with the different training areas to determine which training area is most "similar" to the pixel in the image. Once the most "similar" training area is found the image pixel is labeled with the according land cover class (Figure 4).

The difference between the different types of supervised statistical classification algorithms is how they determine similarity between pixels. In general, the more complex algorithms take longer to process but they also tend to do a better job at assigning the right land cover label to image pixels.

Artificial Neural net classification

Artificial neural network algorithms attempt to mimic the human learning process to associate the correct land cover class label to image pixels. Neural networks have their roots in the field of artificial intelligence and although the algorithm still pales in comparison with the human brain its use in image classification has been quite successful. One advantage to using neural networks is that it is easy to incorporate ancillary data in the classification process. This in itself can greatly improve classification accuracy.

The training process for a neural network classification can be time consuming and is not as simple as the supervised statistical approach. There are a number of parameters that are not very intuitive and experience plays an important part in learning how to use these methods effectively.

Binary decision tree classification

Decision trees are a common machine learning tool that have taken hold in the remote sensing arena. Decision trees are a set of binary rules that define how specific land cover classes should be assigned to individual pixels. (Figure 6) illustrates a simple decision tree that was used to classify forest and non-forest classes using two Landsat TM bands. At each node a true/false decision is made thereby creating a branch in the tree. The terminal node (the bottom tips of the tree) along each path defines an individual land cover class. This approach makes it easy to integrate ancillary data into the classification process.

Creating the rules for a decision tree requires decision tree generator software available in most popular statistics packages. This software takes sample data from training areas in the satellite imagery and ancillary data and works to define the most effective set of rules to define the nodes in the decision tree. Experience is helpful when working with decision trees as they require a good deal of editing (pruning) to produce suitable land cover maps.

Image segmentation

Image segmentation is not a classification tool; rather, it is a method to group contiguous pixels into areas (segments) that are relatively homogeneous (Figure 7). It can be thought of as a preprocessing step before classifying an image. The segmentation is done using sophisticated algorithms that compare a pixel's value with the values of the neighboring pixels. If neighboring pixels are relatively similar then they are added to the contiguous group and if they are not then another segment is started. These programs provide variables that allow an analyst to specify the relative size and sometimes even the shape of the segments.

Once an image has been segmented it can be classified at the segment level instead of the pixel level. There are several advantages to this approach:

- it runs much faster since the number of segments is much less than the number of pixels in an image,
- the relative scale of the segmentation output can be specified so different segmentation runs can be used to capture features of different sizes,
- the classification algorithm can utilize the spectral characteristics (the pixel

values) of an image as well as a host of other segment characteristics that describe the segment, such as mean pixel value, standard deviation, shape of the segments, and dimensions of the segment, and

- the resulting image does not suffer from the "salt and pepper" effect common to the pixel-by-pixel classifiers mentioned above.

This approach holds a lot of promise for imagery with high spatial resolution although it is becoming quite popular for classifying moderate resolution imagery as well. With moderate resolution imagery the accuracy obtained using this approach is often similar to other methods but the advantages mentioned above make it an appealing alternative. As with the other newer methods, this classification approach requires experience to become proficient.

Manual

Manual, or visual, classification of remotely sensed data is an effective method of classifying land cover especially when the analyst is familiar with the area being classified. This method uses skills that were originally developed for interpreting aerial photographs. It relies on the interpreter to employ visual cues such as tone, texture, shape, pattern, and relationship to other objects to identify the different land cover classes. The primary advantage of manual interpretation is its utilization of the brain to identify features in the image and relate them to features on the ground. The brain can still beat the computer in accurately identifying image features. Another advantage is that manual classification can be done without a computer, instead using a hardcopy version of a satellite image.

The downside of manual interpretation is that it tends to be tedious and slow when compared with automated classification and because it relies solely on a human interpreter it is more subjective. Another drawback is that it is only able to incorporate 3 bands of data from a satellite image since the interpretation is usually done using a color image comprised of red, green, and blue bands. More information about using manual interpretation methods is available in the Justification for using photo interpretation methods to interpret satellite imagery guide.

The technique used in manual interpretation is fairly simple. The analyst views the image on either a computer screen or a hardcopy printout and then draws a polygon around areas that are identified as a particular land cover type (Figure 8). If the land cover delineations are done on a computer screen the land cover map is created during the delineation process. If the interpretation is done on a hardcopy image the resulting map will have to be digitized to convert it into a machine readable format.

Hybrid

A hybrid approach combines the advantages of the automated and manual methods to produce a land cover map that is better than if just a single method was used. One hybrid approach is to use one of the automated classification methods to do an initial classification and then use manual methods to refine the classification and correct obvious errors. With this approach you can get a reasonably good classification quickly with the automated approach and then use manual methods to refine the classes that did not get labeled correctly.

The editing process requires that the analyst be able to compare the classified map with either the original satellite image or some other imagery that can be used to identify land cover features. To compare a classified map with imagery it is helpful to have access to software that allows the analyst to flicker between two images (the land cover image and the original satellite image) or slide one image over the other on the computer display using a technique often called "swiping". By doing these comparisons the analyst gets a sense of the quality of the classification. When errors are spotted they can be edited using tools common in many image processing software packages.

It is surprising that many remote sensing practitioners will not edit a classified map even if it is obvious that a certain area is misclassified. If the purpose of the land cover classification study is to produce the best map possible then the analyst should use all means possible to meet that goal. In most cases, visually editing a classified map will improve the accuracy of the final product.

Do I need to visit the area I'm mapping?

Although land cover maps are often made without visiting the field, there are good reasons why field visits should be made. The two primary reasons for visiting the area that is being mapped are to collect data that can be used to train the algorithm or the interpreter and to collect data that can be used to evaluate the land cover map and estimate the accuracy of the individual classes (a process called validation). At a minimum, these data can be collected in one trip but often two or more trips are preferred so that validation information can be systematically collected using a sampling design based on the classification results.

Data collected in the field must be georeferenced so that the point where the data were collected can be located on the imagery. GPS receivers are commonly used to record this location information. The type of information collected can range from detailed notes describing a site to a photograph of the site. Some of the detailed information that can be recorded includes: type of vegetation, crown closure slope, aspect, soil type, and other bio-physical characteristics that are important to identify the land cover type. If photographs are taken it is a good idea to record the direction the camera was pointed and to make notes about the area to supplement the content in the photograph. For example, you could add information about species composition, tree height, and possibly land use.

When land cover maps are created without using field data from the region of interest it is difficult to predict the accuracy of the final land cover map. An analyst with significant experience may be able to produce a land cover map of high quality but without validation information the true accuracy of the image classification quality is not known.

How do I know if my map is any good?

In the needs assessment some thought should be given to the required accuracy of the final map. The accuracy can refer to either spatial accuracy or thematic accuracy. Spatial accuracy is directly related to the base information that is used. If you are using satellite imagery the spatial accuracy of the final map will be dictated by the spatial accuracy of the satellite image that was used to create the map. Thematic accuracy specifies how well individual mapped classes match what is on the ground. For example, if the map indicates that there is forest at a particular location, what is the actual land cover type? Error is inherent to the mapping process and determining an acceptable level of error requires careful thought. Factors affecting accuracy include: quality and type of the base data, experience of the analyst, the analyst's familiarity with the area being classified, and the level of detail of the classification scheme. The accuracy of a land cover map can be assessed using statistical sampling procedures that are outlined in many remote sensing textbooks. The basic idea is to select a number of sample sites from each of the cover types in the final image and then go into the field to see what type of cover is actually on the ground. This information is then compiled in a contingency table (Table 1) so that the accuracy of each class can be determined.

Table 1:

Contingency Table: A pixel by pixel comparison of ground reference class to satellite-based map class.

Cover Type	Satellite Map Class (pixel counts)				Total	Producers Accuracy, Percent Correct	
	Conifer	Hardwood	Grass	Barren			
Conifer	911	20	1	0	932	97.7	
Hardwood	40	343	72	2	457	75.1	(1-omission error)
Grass	0	62	176	14	252	69.8	
Barren	0	0	19	27	46	58.7	
Total	951	425	268	43	1687		
Users Accuracy, Percent Correct	95.8	80.7	65.7	62.8		86.4	

(ground reference labels the rows)

(1-commission error)

The sampling statistics for this process can be fairly straightforward but the practical issues such as limited access to the study area, insufficient funds to visit all of the sites, and a lack of time tend to impose limitations that must be accommodated in the sampling design. Dealing with this less than ideal situation

may necessitate some creative solutions that require an understanding of statistics beyond the expertise of the image analyst. When this is the case it is good to have the sampling design reviewed by someone with a solid understanding of statistics. More information about accuracy assessment can be found in the Overview of accuracy assessment of land cover products guide.

How can I use a classified map?

Classified maps are used for a host of quantitative analysis applications such as species or landscape modeling, fragmentation analysis, and conservation priority. They can also be used as visual aids in a presentation or as a layer in a Geographic Information System (GIS). No matter how a classified map is used, it is useful to have some information on the accuracy of the map. This information can come from a carefully planned accuracy assessment or some less rigorous qualitative methods, but the source of the accuracy should also be known. It is important to keep in mind that a classified map is only an approximate representation of the features on the ground. The accuracy of this representation can greatly affect the results of any quantitative analysis. Another important characteristic of a classified map is spatial scale. This is especially important when using a land cover map to calculate landscape metrics. A coarse scale land cover map will produce very different results than a finer scale map.

Written and compiled by Ned Horning, Center for Biodiversity and Conservation at the American Museum of Natural History
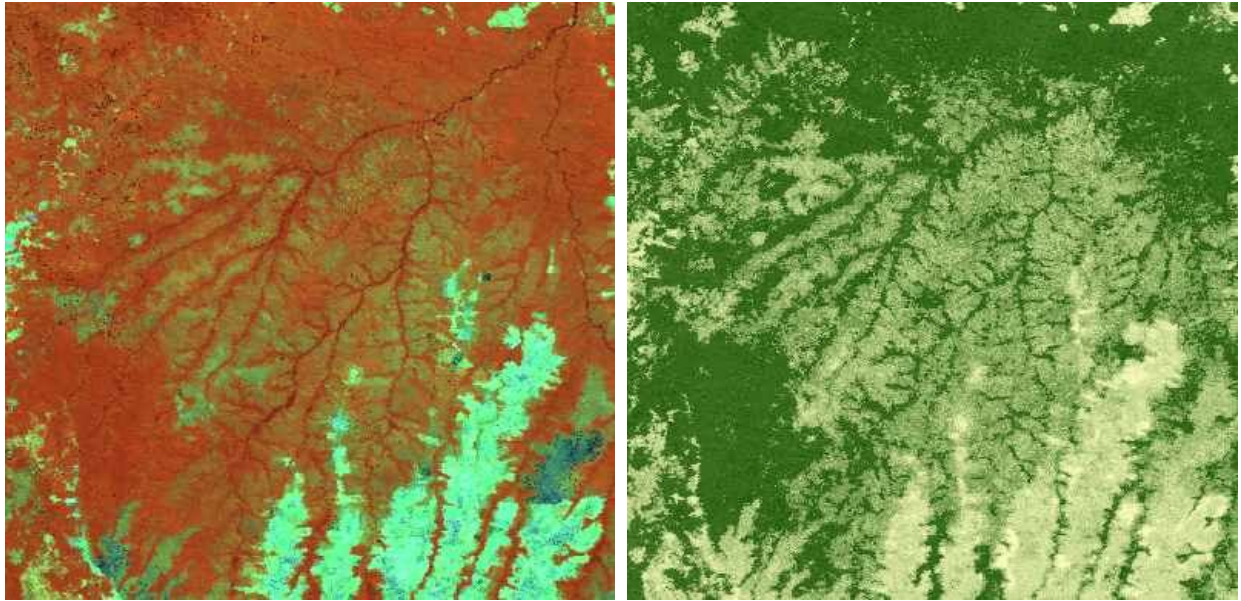
Figure 1

Adding a buffer to broaden the study area

A 1.5 kilometer buffer is shown as a transparent blue strip around Bach Ma park in central Vietnam. The image is a Landsat Enhandced Thematic Mapper Plus (ETM+) image acquired 21 April, 2003

Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*

figure 2

Comparing Images from Two Dates

These two images cover an area including northeastern Bolivia and Mato Grosso, Brazil. The image on the left is a 32-day composite MODIS image. The image on the right is a Continuous Fields for tree cover product. Each pixel in the tree cover image represents the percentage of the pixel that is covered by tree canopy. Dense cover shows up as a dark green and little or no tree cover shows up as beige. These two images were created by the Global Land Cover Facility.

Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*
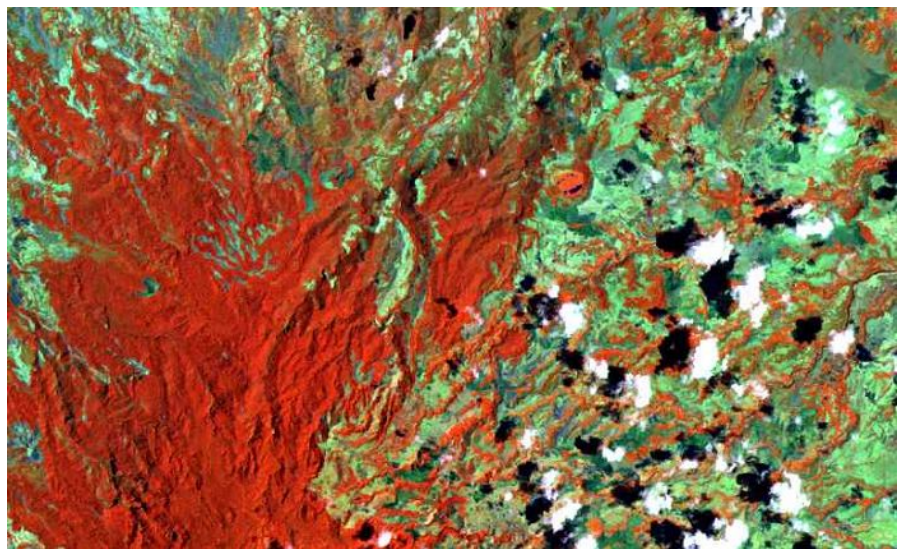
figure 3

Effects of Shadow on an Image

This is a portion of a Landsat Thematic Mapper (TM) image acquired over the northern part of Madagascar in July. Notice how shadows help accentuate the terrain.
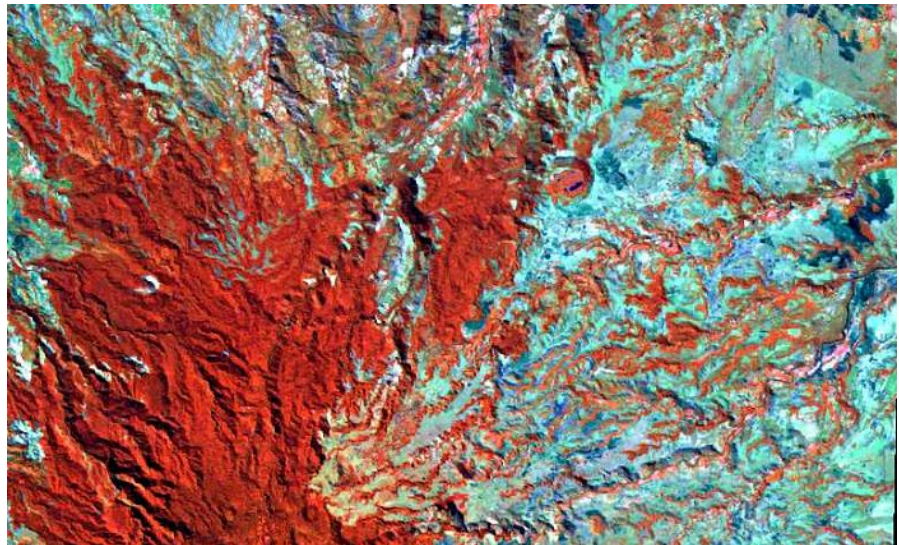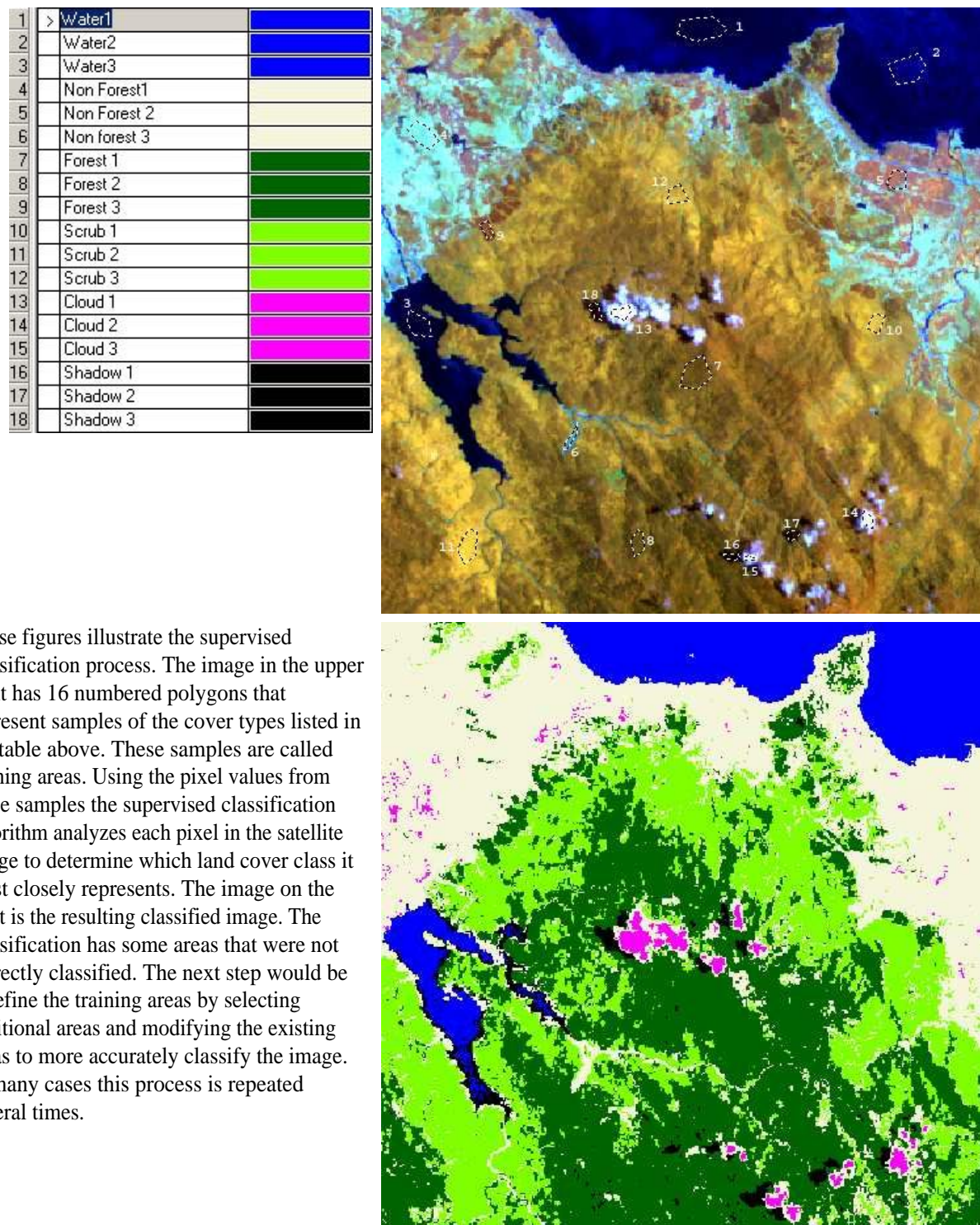

This is a portion of a Landsat Thematic Mapper (TM) image acquired over the northern part of Madagascar in November. Notice how the reduced shadows make the terrain look flatter.
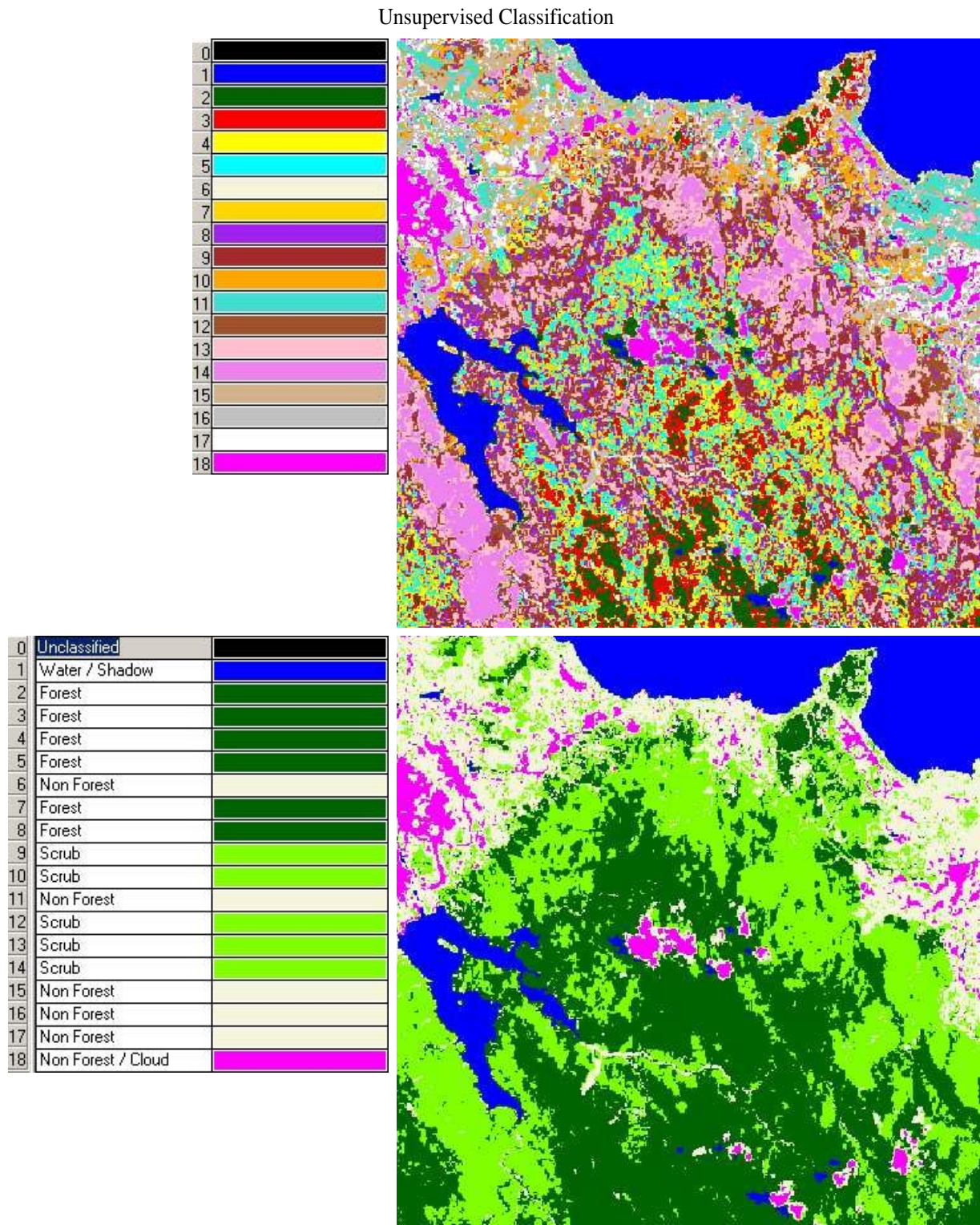
Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*

figure 4

1	>	Water1	
2		Water2	
3		Water3	
4		Non Forest1	
5		Non Forest 2	
6		Non forest 3	
7		Forest 1	
8		Forest 2	
9		Forest 3	
10		Scrub 1	
11		Scrub 2	
12		Scrub 3	
13		Cloud 1	
14		Cloud 2	
15		Cloud 3	
16		Shadow 1	
17		Shadow 2	
18		Shadow 3	

These figures illustrate the supervised classification process. The image in the upper right has 16 numbered polygons that represent samples of the cover types listed in the table above. These samples are called training areas. Using the pixel values from these samples the supervised classification algorithm analyzes each pixel in the satellite image to determine which land cover class it most closely represents. The image on the right is the resulting classified image. The classification has some areas that were not correctly classified. The next step would be to refine the training areas by selecting additional areas and modifying the existing areas to more accurately classify the image. In many cases this process is repeated several times.

Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*

figure 5

Unsupervised Classification

0	Unclassified
1	Water / Shadow
2	Forest
3	Forest
4	Forest
5	Forest
6	Non Forest
7	Forest
8	Forest
9	Scrub
10	Scrub
11	Non Forest
12	Scrub
13	Scrub
14	Scrub
15	Non Forest
16	Non Forest
17	Non Forest
18	Non Forest / Cloud

These figures illustrate the unsupervised classification process. The image in the upper right is the initial output from the ISODATA algorithm. The ISODATA algorithm is able to group similar pixels to create the number of classes specified by the image analyst. In this case 18 classes, each represented by a different color, were created. This is

illustrated in the table in the upper right. The next step involves labeling each of these classes with one of our selected land cover types. The table in the lower left shows how each of the classes were labeled to represent actual land cover. The lower right image shows the resulting classified image. In many cases it is necessary to refine this classification by using the algorithm to create more classes.

Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*

figure 6

Decision Tree Classification
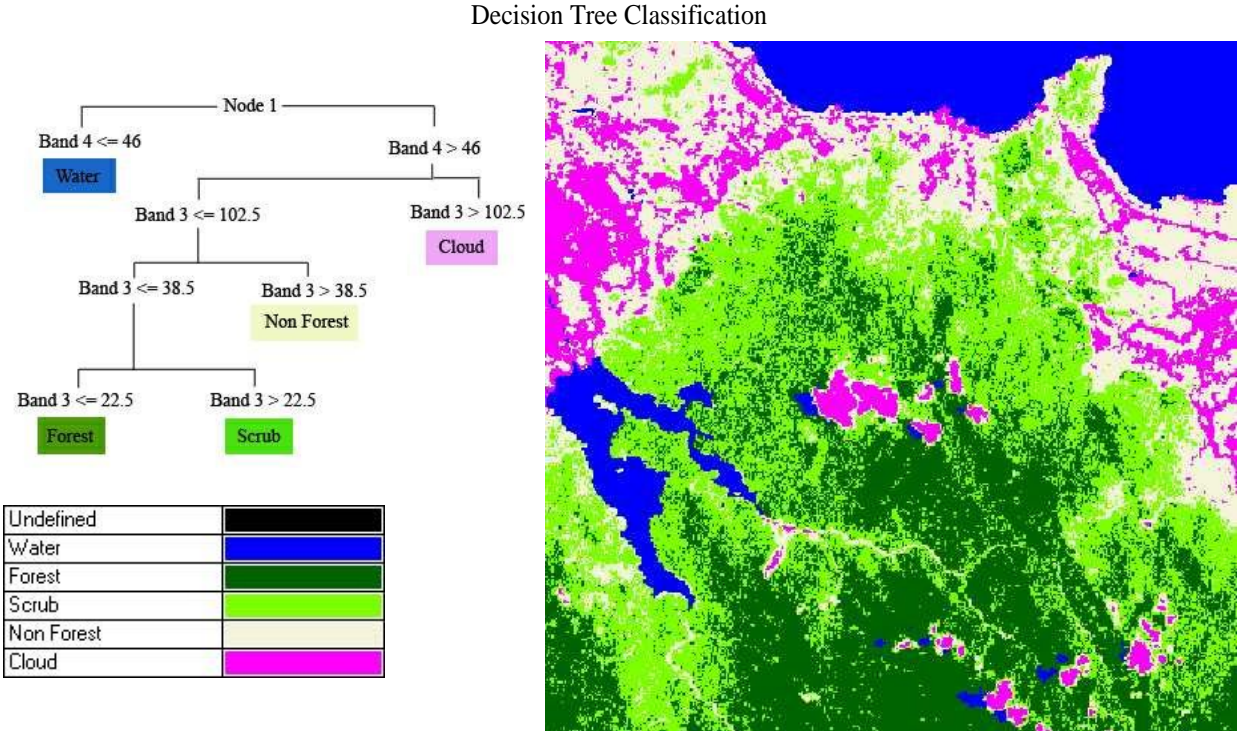
Using decision tree classification the analyst uses decision tree rule generation software to calculate a set of rules from a set of training sites (similar to supervised classification) that can be systematically applied to the satellite image to create land cover classes. In this simplistic example we use two Landsat Enhanced Thematic Mapper Plus (ETM+) bands to create a classified map with five land cover classes. The decision tree is shown in the upper left corner. This indicates the cutoffs that were determined using the rule generation software. In this example if the pixel value for band 4 was less then or equal to "46" it was assigned to the "Water" class. If band 4 is greater than "46" and band 3 is greater than "102.5" then the pixel is labeled as a "cloud". This process of successive testing continues until all of the pixels are classified.

Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*

Figure 7

Comparing Images from Two Dates
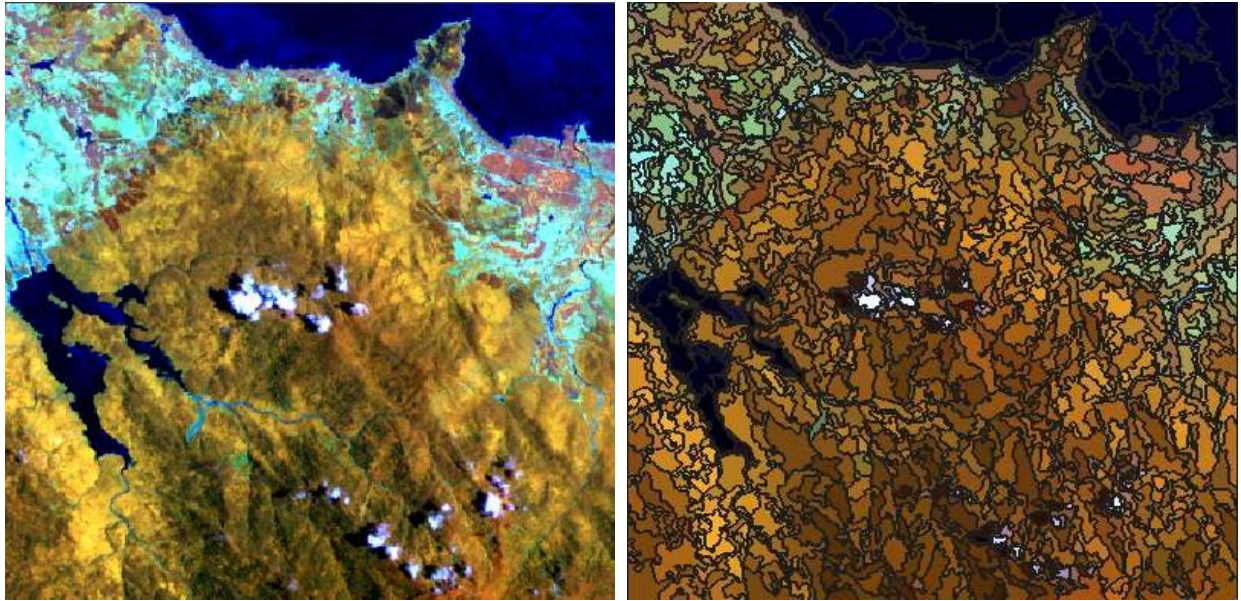
The image on the right is the result of a segmentation process applied to the image on the left. Segments are delimited with black lines. The segmentation process groups together adjacent pixels that have similar characteristics.

Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*

figure 8
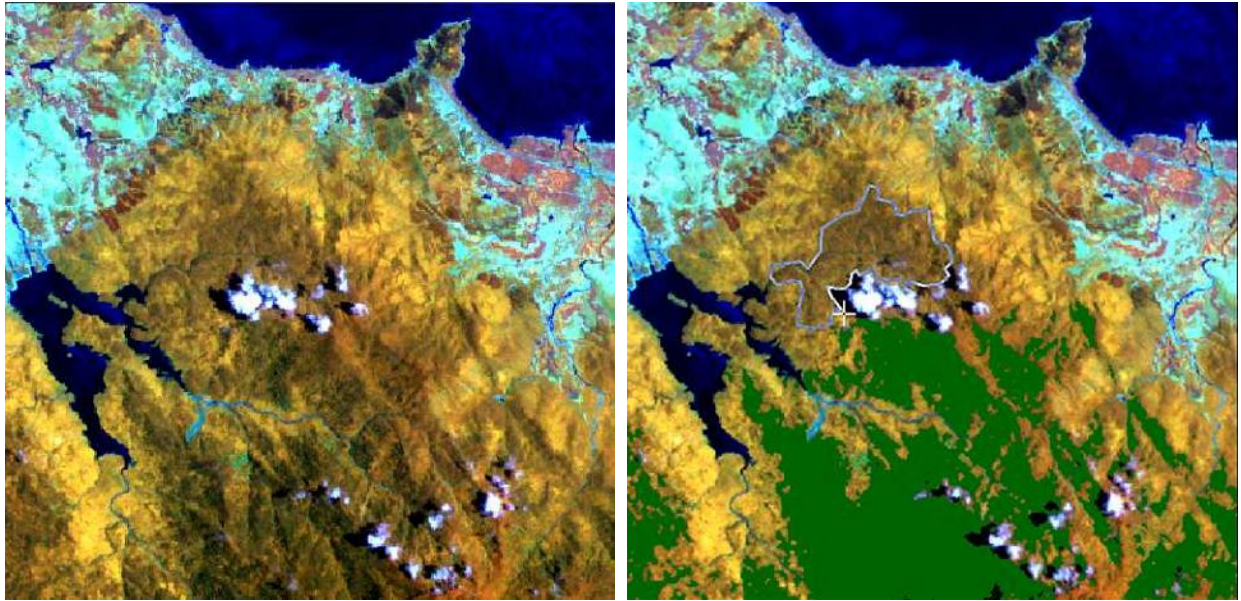
On-screen digitizing involves an image analyst drawing polygons around land cover types of interest. In this image on the right the analyst has already digitized much of the forest area (colored dark green) and an area of forest is being digitized as can be seen by the polygon next to the cursor (white crosshair).

Credit: *American Museum of Natural History, Center for Biodiversity and Conservation*

Biodiversity Informatics Facility

Biodiversity Informatics Facility

Creative Commons License