Time Series/Change Analysis

This chapter gives a brief overview of the special procedures available for change and time series analysis in IDRISI.

The techniques for the analysis of change are broken down into three broad categories for this chapter. The first of these contains techniques that are designed for comparisons between pairs of images; the second is made up of techniques that are concerned with the analysis of trends and anomalies across multiple images (i.e., a time series), and the third consists of methods for predictive modeling and assessment of models.

Pairwise Comparisons

With pairwise comparisons we can further break down the techniques according to whether they are suitable for quantitative or qualitative data. Quantitative data has values that indicate an amount or measurement, such as NDVI, rainfall or reflectance. Qualitative data has values that indicate different categories, such as census tract ID’s or landuse classes.

Quantitative Data

Image Differencing

With quantitative data, the simplest form of change analysis is image differencing. In IDRISI, this can be achieved with the OVERLAY module through a simple subtraction of one image from the other. However, a second stage of analysis is often required since the difference image will typically contain a wide range of values. Both steps are included in the module IMAGEDIFF, which produces several common image difference products: a simple difference image (later - earlier), a percentage change image (later-earlier/earlier), a standardized difference image (Z-scores), or a classified standardized difference image (Z-scores divided into 6 classes). Mask images that limit the study area may also be specified.

Care must be taken in choosing a threshold to distinguish true change from natural variability in any of these difference images. There are no firm guidelines for this operation. A commonly used value for the threshold is 1 standard deviation (STD) (i.e., all areas within 1 STD are considered non-change areas and those beyond 1 STD in either the positive or negative direction are considered change areas), but this should be used with caution. Higher values may be more appropriate and in some cases natural breaks in a histogram of the simple difference or percentage change images may be more sensible as a basis for choosing the threshold values.

Image Ratioing

While image differencing looks at the absolute difference between images, image ratioing looks at the relative difference. Again, this could be achieved with OVERLAY using the ratio option. However, because the resulting scale of relative change is not symmetric about 1 (the no change value), it is recommended that a logarithmic transformation be undertaken before thresholding the image. The module IMAGERATIO offers both a simple ratio and a log ratio result.

Regression Differencing

A third form of differencing is called regression differencing. This technique should be used whenever it is suspected that the measuring instrument (e.g., a satellite sensor) has changed its output characteristics between the two dates being compared. Here the earlier image is used as the independent variable and the later image as the dependent variable in a linear relationship.

---

regression. The intercept and slope of this regression expresses the offset and gain required to adjust the earlier image to have comparable measurement characteristics to the later. In effect, we create a predicted later image in which the values are what we would expect if there were no change other than the offset and gain caused by the changes in the sensor. The equation is:

\[
\text{predicted later image} = (\text{earlier image} \times \text{gain}) + \text{offset}
\]

With the sensor differences accounted for, the predicted later image and the actual later image may then be analyzed for change. Note that this technique requires that the overall numeric characteristics of the two images be equal except for sensor changes. The technique may not be valid if the two images represent conditions that are overall very different between the two dates.

The module CALIBRATE automates the image adjustment process. The input image (the one to calibrate) is used as the independent variable and the reference image is used as the dependent variable in the regression. The output image is adjusted to the characteristics of the reference image and thus can be used in a standard comparison operation (such as IMAGEDIFF or IMAGERATIO) with any image also based on this reference, including the reference image itself.

Note that CALIBRATE also offers options to adjust an image by entering offset and gain values or by entering mean and standard deviation values.

**Change Vector Analysis**

Occasionally, one needs to undertake pairwise comparisons on multi-dimensional images. For example, one might wish to undertake a change analysis between two dates of satellite imagery where each is represented by several spectral bands. To do so, change vector analysis can be used. With change vector analysis, difference images are created for each of the corresponding bands. These difference images are then squared and added. The square root of the result represents the magnitude of the change vector. All these operations can be carried out with the Image Calculator, or a combination of TRANSFORM and OVERLAY. The resulting image values are in the same units as the input images (e.g., %).

When only two bands (for each of the two dates) are involved, it is also possible to create a direction image (indicating the direction of change in band space). The module CVA calculates both magnitude and direction images for 2-band image pairs. Figure 19.1 illustrates these calculations. The magnitude image is in the same units as the input bands and is the distance between the Date 1 and Date 2 positions. The direction image is in azimuths measured clockwise from a vertical line extending up from the Date 2 position.

![Figure 19.1](image-url)

**Qualitative Data**

**Crosstabulation / Crossclassification**

With qualitative data, CROSSTAB should be used for change analysis between image pairs and there are several types of output that can be useful. The crosstabulation table shows the frequencies with which classes have remained the same.
(frequencies along the diagonal) or have changed (off-diagonal frequencies). The Kappa Index of Agreement (KIA) indicates the degree of agreement between the two maps, both in an overall sense and on a per-category basis. Finally, the crossclassification image can readily be reclassified into either a change image or an agreement image. Note that the numeric values of data classes must be identical on both maps for the output from CROSSTAB to be meaningful.

Multiple Image Comparisons

With multiple images, a variety of techniques can be used. For analysis, the most important of the routines offered in IDRISI is Time Series Analysis (TSA).

Time Series Analysis (TSA)

TSA can produce an analysis of up to 256 input images, providing both spatial and temporal outputs. It analyzes the series as a whole based on a Standardized Principal Components Analysis. An ordered set of uncorrelated component images are produced, each expressing underlying themes (trends, shifts, periodicities, etc.) in the data of successively lower magnitude (in terms of the total variability explained in the original image set).

As a prelude to running TSA (and many of the other routines in this section), you will need to create a time series (ts) file using IDRISI Explorer. This file lists in order the names of the files of the time series.

TSA requires the name of the time series file and, among other things, the number of component images to be produced and the type of output to be used for the temporal data, the loadings. Assuming that the series contains a large number of images, you may wish to limit your output to no more than 10-12 components. For many studies, this proves to be sufficient. Also, for most purposes, integer components provide adequate precision while saving valuable disk space.

For temporal output, saving the component loadings as a DIF file provides the greatest flexibility. The output can then be brought into a spreadsheet for the preparation of loadings graphs. For quick analyses, output of IDRISI profiles is convenient.

Analysis of the results of TSA is done by examining the component images (the spatial output) in combination with the component loadings (the temporal output). The first component image describes the pattern that can account for the greatest degree of variation among the images. The loadings indicate the degree to which the original images correlate with this component image. Quite typically, most images will correlate strongly and roughly evenly (i.e., with correlations in the 0.93-0.98 range) with the first component image. Since graphing programs tend to autoscale loadings graphs, your graph of the first component may tend to look overly irregular. We therefore suggest that you combine components 1 and 2 onto the same graph. This will ease the problem for reasons that will become apparent once you try it.

Unless the region analyzed is very small, component 1 will most likely represent the typical or characteristic value over the series. Since each component is uncorrelated with the others, all successive components represent change. Indeed, the process can be thought of as one of successive residuals analysis. High positive values on the image can be thought of as areas that correlate strongly with the temporal pattern in the loadings graph while those with high negative values correlate strongly with the inverse of the graph. (To get a sense of what the inverse is, imagine that you could hold the X axis of the graph at the ends and rotate it until the Y axis is inverted. This can also be done by looking at the graph through the back of the paper, upside down.) We sometimes find it useful to invert the graph in this way and to invert the palette as well. Areas that correlate strongly with the inverted graph then have the same color as the areas that correlate

84. Inverting the signs associated with the loadings and component values constitutes a permissible rotation in Principal Component Analysis and does not affect the interpretation so long as both elements are rotated as concert.

85. To invert the palette, open the palette in Symbol Workshop. Use the Reverse function, then save the palette under a new name. Then change the palette to be used for the image display to the one just created.

Chapter 19  Time Series/Change Analysis  233
Just as strong anomalies on the image can be thought to represent areas that are strongly associated (positively and negatively) with the temporal loadings, time periods with high positive loadings can be thought of as those with spatial patterns very similar to those in the component image, while those with strong negative loadings are associated with the inverse pattern (that you would get by inverting the palette).

Interpretation of the output from TSA requires care and patience. However, the results can be enormously illuminating. Since each successive component represents a further residual, you will find yourself peeling deeper and deeper into the data in a manner that would otherwise be impossible. Here you may find latent images of trends that would be almost impossible to detect through direct examination of the data.

A Note About Masking and Component Rotation

This version of TSA offers the option of specifying a mask image to exclude portions of the images from the analysis. Generally, we do not recommend that you do this. Rather, experiments at Clark Labs have shown that it is generally preferable to mask the input images first (e.g., to 0) and include these as data in the analysis (i.e., run the analysis with a mask). If the masking is applied identically to all images (e.g., through the use of a dynagroup in Macro Modeler), it adds no source of variability to the series. Thus its effect will be completely described by the first component. All other components will be free of its effects. However, it does contribute to the intercorrelation between images. The effect is thus equivalent to rotating the axes such that the first component is focused on the commonalities between the images. We have found this generally to be highly desirable, leading to a highly consistent logic to the interpretation of the initial components, with the first component showing the characteristic pattern over time, and the second component showing the main seasonal effects.

**Time Series Correlation**

TSA is an inductive procedure that attempts to isolate the major components of variation over time. In contexts where you are looking for evidence of a particular temporal phenomenon, however, you may find it more effective to use the CORRELATE module. CORRELATE compares each pixel, over time, with the values of an index in a non-spatial time series. For example, one might compare a time series of vegetation index images with a measure of central Pacific ocean temperature on a monthly basis. The result would be an image where each pixel contains a Pearson Product-Moment correlation expressing the correlation between that vegetation index and sea surface temperature over the series.

**Time Profiling**

To examine changes in values at specific locations over time, PROFILE may be used. The module requires two inputs—a raster image defining up to 15 sites or classes to monitor, and a time series file listing the images to be analyzed. A variety of statistics can be produced (e.g., mean, minimum, maximum, range, standard deviation). Output is in the form of a graph, although you also have the option to save the profiles as values files.

**Image Deviation**

As an extension to simple differencing with pairwise comparisons, image deviation can be used with time series data. The logic of image deviation is to produce a characteristic image for the series as a whole, from which any of the individual images can be subtracted to examine how it differs from the sequence at large. The most common procedure is to create a mean image, then difference each image from that mean. This can be accomplished with Image Calculator or a combination of OVERLAY and SCALAR. If the input images are all in byte binary form, it is also possible to use the Multi-Criteria Evaluation (MCE) routine to create the mean image. To do so, simply treat each input image as a factor and apply a weight equal to \(1/n\) where \(n\) represents the number of images in the sequence. The resulting difference images must

---

then be thresholded as discussed above in the section on pairwise image differencing.

**Change Vector Analysis II**

Occasionally one needs to examine the difference between two series. For example, one might have 12 months of image data for one year and 12 months of image data for another. To compare the two years, *change vector analysis* can be used. Use CVA to produce a change magnitude image for each of the monthly pairs, then sum the results using OVERLAY or Image Calculator. Please note that change vector analysis of this type can suffer from significant problems of temporal misregistration. The assumption behind the technique is that human-defined slices of time are environmentally meaningful—however, they rarely are. For example, if the image sets represented vegetation index data for an area with a pronounced season, a small difference in the onset of rains between the years would be considered by the technique as substantial change, while there might be little difference in agricultural output and timing of the harvest.

**Time Series Correlation**

It is a common need in change and time series analysis to identify key temporal patterns occurring through the image set. For example, you may have developed one or more typical drought scenarios that tend to produce particular patterns in monthly rainfall or NDVI data. You might then wish to examine a monthly time series of images to see where that pattern has most closely occurred. The module CORRELATE is designed to calculate the degree to which each pixel location corresponds to a given pattern as recorded in an attribute values file. The measure used is the Pearson Product-Moment coefficient of correlation.

**Predictive Change Modeling**

In some cases, knowing the changes that have occurred in the past may help predict future changes. A suite of modules in IDRISI has been developed to provide the basic tools for predictive landcover change modeling. These tools are primarily based on Markov Chain Analysis and Cellular Automata.

**Markov Chain Analysis**

A Markovian process is one in which the state of a system at time 2 can be predicted by the state of the system at time 1 given a matrix of transition probabilities from each cover class to every other cover class. The MARKOV module can be used to create such a transition probability matrix. As input, it takes two landcover maps. It then produces the following outputs:

- A transition probability matrix. This is automatically displayed, as well as saved. Transition probabilities express the likelihood that a pixel of a given class will change to any other class (or stay the same) in the next time period.

- A transition areas matrix. This expresses the total area (in cells) expected to change in the next time period.

- A set of conditional probability images—one for each landcover class. These maps express the probability that each pixel will belong to the designated class in the next time period. They are called conditional probability maps since this probability is conditional on their current state.

ST_CHOICE is a stochastic choice decision module. Given the set of conditional probability images produced by MARKOV, ST_CHOICE can be used to produce any number of potential realizations of the projected changes embodied in the conditional probability maps. If you try this, however, you will find the results to be disappointing. The output from MARKOV has only very limited spatial knowledge. To improve the spatial sense of these conditional probability images (or in fact, any statistic), use DISAGGREGATE. Given an image of the likely internal spatial pattern of an areal statistic, DISAGGREGATE redistributes the statistic such that it follows the suggested pattern, but maintains the overall area total. NORMALIZE can then be used to ensure that probabilities add to 1.0 at each pixel (this may need to be

Chapter 19  Time Series/Change Analysis
Cellular Automata

One of the basic spatial elements that underlies the dynamics of many change events is proximity: areas will have a higher tendency to change to a class when they are near existing areas of the same class (i.e., an expansion phenomenon). These can be very effectively modeled using cellular automata. A cellular automaton is a cellular entity that independently varies its state based on its previous state and that of its immediate neighbors according to a specific rule. Clearly there is a similarity here to a Markovian process. The only difference is application of a transition rule that depends not only upon the previous state, but also upon the state of the local neighborhood.

Many cellular automata transition rules can be implemented through a combination of FILTER and RECLASS. Take, for example, the case of Conway’s *Game of Life*. In this hypothetical illustration, the automata live or die according to the following criteria:

- An empty cell becomes alive if there are three living automata in the 3x3 neighborhood (known as the Moore neighborhood) surrounding the cell.
- The cell will stay alive so long as there are 2 or 3 living neighbors. Fewer than that, it dies from loneliness; more than that, it does from competition for resources.

This can be implemented using the following kernel with the FILTER module:

```
1 1 1
1 10 1
1 1 1
```

followed by the following RECLASS rule:

```
0 - 2  =  0
3 - 4  =  1
4 - 11 =  0
12 - 13 =  1
14 - 15 =  0
```

The critical element of this rule is the use of the 10 multiplier in the central cell. As a result of the filter step, you know that the central cell is occupied if the result is 10 or greater. The CELLATOM module can be used to implement this kind of Cellular Automaton rule. However, a cellular automaton procedure very specific to the context of predictive landcover change modeling is implemented with the CA_MARKOV module.

CA_MARKOV takes as input the name of the landcover map from which changes should be projected, the transition areas file produced by MARKOV from analysis of that image and an earlier one, and a collection (zip) of suitability images that express the suitability of a pixel for each of the landcover types under consideration. It then begins an iterative process of reallocating landcover until it meets the area totals predicted by the MARKOV module. The logic it uses is this:

- The total number of iterations is based on the number of time steps set by the user. For example, if the projection is for 10 years into the future, the user might choose to complete the model in 10 steps.
Within each iteration, every landcover class will typically lose some of its land to one or more of the other classes (and it may also gain land from others). Thus within the consideration of each host within each iteration, claimant classes select land from the host based on the suitability map for the claimant class. Since there will commonly be competition for specific land parcels, this process of land allocation is undertaken using a multi-objective allocation procedure (the MOLA module).

The Cellular Automaton component arises in part from the iterative process of land allocation, and in part from a filtering stage with each iteration that reduces the suitability of land away from existing areas of that type. By default, the module uses a 5x5 mean filter to achieve this contiguity constraint. By filtering a Boolean mask of the class being considered, the mean filter yields a value of 1 when it is entirely within the existing class and 0 when it is entirely outside it. However, when it crosses the boundary, it will yield values that quickly transition from 1 to 0. This result is then multiplied by the suitability image for that class, thereby progressively downweighting the suitability as one moves away from existing instances of that class. Note that it is possible to apply a different filter by specifying an alternative filter file (fil). Also note that class masks are defined at each step to incorporate new areas of growth.

The net result of this iterative process is that landcover changes develop as a growth process in areas of high suitability proximate to existing areas. CA_MARKOV is computationally intensive—a typical run might involve several thousand GIS operations. Thus you should start the run when you can leave your computer for 15-30 minutes.

Model Validation

An important stage in the development of any predictive change model is validation. Typically, one gauges one's understanding of the process, and the power of the model, by using it to predict some period of time when the landcover conditions are known. This is then used as a test for validation. IDRISI supplies a pair of modules to assist in the validation process.

The first is called VALIDATE, and provides a comparative analysis on the basis of the Kappa Index of Agreement. Kappa is essentially a statement of proportional accuracy, adjusted for chance agreement. However, unlike the traditional Kappa statistic, VALIDATE breaks the Kappa down into several components, each with a special form of Kappa or associated statistic (based on the work of Pontius (2000)):

- Kappa for no information = Kno
- Kappa for location = Klcoation
- Kappa for quantity = Kquantity
- Kappa standard = Kstandard
- Value of Perfect Information of Location = VPIL
- Value of Perfect Information of Quantity = VPIQ

With such a breakdown, for example, it is possible to assess the success with which one is able to specify the location of change versus the quantity of change.

The other validation procedure is the ROC (Relative Operating Characteristic). It is used to compare any statement about the probability of an occurrence against a Boolean map which shows the actual occurrence. It can be useful, for example, in validating modifications to the conditional probability maps output from MARKOV. Note that LOGISTICREG incorporates ROC directly in its output.
Exercise 5-7
Vegetation Analysis in Arid Environments

In this exercise, we will explore the use of different vegetation index calculation models available in the VEGINDEX, TASSCAT and PCA modules to analyze vegetation cover. Before continuing, you may find it useful to read or review the Vegetation Indices chapter in the IDRISI Guide to GIS and Image Processing. That chapter provides an extensive overview of many vegetation indices, only some of which will be used in this exercise.

Introduction to Vegetation Indices

Vegetation cover was an early focus of research in natural resources management using space-born satellite images, especially with the release of the Earth Resources Technology Satellites known as Landsat in 1972. Landsat, SPOT and NOAA data offer time series images that are widely used to monitor and assess the status of vegetation at the global, regional, national and local levels. Vegetation indices use various combinations of multi-spectral satellite data to produce a single image representing the amount of vegetation present, or vegetative vigor. Low index values usually indicate little healthy vegetation while high values indicate much healthy vegetation. Different indices have been developed to better model the actual amount of vegetation on the ground. The index that is most appropriate for use in a particular environment can best be determined through calibration with sample measurements of biomass. In the absence of biomass measurements, these index values can be used indicators of the relative amount of vegetation present.

Vegetation has a characteristic spectral response pattern in which visible blue and red energy is absorbed strongly, visible green light is reflected weakly (hence giving vegetation its green color) and near infrared energy is very strongly reflected. Because of this characteristic spectral response pattern, many of the vegetation index models use only the red and near-infrared imagery bands.

Introduction to the Data and the Study Area

In this exercise, we will assess vegetation cover and its changes in an area of southern Mauritania.

a) Display the image MAUR90-BAND3 with the Greyscale palette and choose to autoscale the image using Equal Intervals.

The area covered by the images in this exercise is near the Senegal/Mauritania border and contains part of the Senegal River flood plain as well as the lower section of the Gargol River flood plain (partially visible at the upper-left corner of the image). This is a tributary of the Senegal River. These sections of the two rivers are covered by permanent vegetation dominated by the Acacia nilotica species, the preferred species for fuelwood and charcoal. Other woody species such as Boswellia forskalii and Ziziphus boehnia are used as building material. Rainfed and flood recession agriculture and grazing are

1. Of the 19 vegetation indices produced in the VEGINDEX module of IDRISI, only the RVI and NRI indices produce images with high values indicating little vegetation and low values indicating more vegetation. If you are using a vegetation index model not provided in VEGINDEX, you must determine whether the index values are proportional or inversely proportional to the amount of vegetation present before you can properly interpret the image.

also practiced in this region.

Once a relatively humid area, persistent rainfall deficits since the late 1960s have left the study area, as well as more and more of the Sahel, semi-arid. Much vegetation has shifted from savanna to steppe. Relics of the savanna vegetation are only found along river valleys on clay, clay sand and sandy clay soils, since these retain moisture better than other soils in the area. Increasing pressure from populations trying to adapt to the continuous drought conditions has been the main cause of vegetation cover degradation in this environment.

Quantifying the low density vegetation cover that characterizes arid and semi-arid lands is especially challenging because vegetation cover is not complete - most pixels contain an average reflectance of vegetation and bare soil. Some of the vegetation index models will be used as developed specifically to help account for the effects of background soil reflectance.

The data we will use are Landsat Multi-spectral Scanner (MSS) images. These images were taken on October 10, 1980 and October 12, 1990 by Landsat 4. There are eight images provided in the dataset, four from each year: MAUR80-BAND1, MAUR80-BAND2, MAUR80-BAND3 and MAUR80-BAND4 for 1980; MAUR90-BAND1, MAUR90-BAND2, MAUR90-BAND3 and MAUR90-BAND4 for 1990. These correspond to MSS bands visible green, visible red, near-infrared and a slightly longer-wavelength near-infrared, respectively. Since the two scenes were taken at two different dates, they must be registered to one another if we are to do analysis between them. This task has already been performed using a methodology similar to that described in Exercise 4.5. We will begin the exercise by producing and comparing several vegetation indices for the 1990 scene, then we will analyze changes between the two scenes.

Creating Vegetation Index Images

There are three major families of vegetation indices that we will explore: Slope-Based, Distance-Based and Orthogonal Transformation vegetation indices.

The Slope-Based VI's

The slope-based VI's use the ratio of the reflectance of one band to that of another, usually the red and the near-infrared. The term slope-based is used because in comparing resulting VI values, we are essentially comparing the slopes of lines passing through the origin and the pixels as plotted on a graph with the reflectance of one band as the X-axis and the reflectance of the other as the Y-axis.

b) Before beginning our exploration of vegetation indices, select User Preferences from the File menu and set the "Automatically display the output of analytical modules" feature on. We will always display the VI images with a user-defined palette named NVDI. Go to the Display tab of the User Preferences dialog box and enter NVDI as the Quantitative Palette. Also, choose to show titles, but do not show legends (this will maximize display space). Click OK to save the settings and exit User Preferences.

c) Use the module VEGINDEX (Image Processing/Transformation menu) twice to produce images for two of the slope-based models: Ratio and NDVI. Use MAUR90-BAND2 as the red band and MAUR90-BAND3 as the near infrared band. Call the resulting images 90RATIO and 90NDVI. Examine each of the output images. Consult the on-line Help System for details about the equation used for each index.

1. What similarities and differences do you notice between the two output images? (To answering this question, it may be useful to look at the pair of images with other quantitative palettes as well, such as Grey-scale or Quant.) What is the purpose of normalizing the Ratio to create NDVI? (You may wish to consult the Vegetation Indices chapter for help in answering this question.)

The slope-based VI's are simple linear combinations that use only the reflectance information from the red and infrared
bands. In contrast, the second family of Vegetation Indices that we will explore, the distance-based VIs, uses information about the reflectance characteristics of the background soil in addition to the red and infrared bands.

**The Distance-Based VIs**

The reflectance values recorded by the sensor for each pixel constitute an average reflectance of all the cover types in the instantaneous field of view (i.e., the pixel). When vegetation cover is not complete, which is particularly the case in arid and semi-arid regions, the average reflectance values are greatly influenced by the background soil type. The distance-based VIs address this problem of separating information about vegetation from information about soils in remotely sensed data.

The distance-based indices are based on the concept of a soil line and distances from that soil line. A soil line is a linear equation that describes the relationship between reflectance values in the red and infrared bands for bare soil pixels. This line is produced by running a simple linear regression between the red and infrared bands on a sample of bare soil pixels. Once the relationship is known, all unknown pixels in an image that have that same relationship in red and infrared reflectance values are assumed to be bare soils. Unknown pixels that fall far from the soil line because they have higher reflectance values in the infrared band are assumed to be vegetation (based on the characteristic spectral response pattern for vegetation where the infrared band reflectance values are relatively higher than those of the red band). Those that fall far from the soil line because their red reflectances are high are often assumed to be water (based on the characteristic spectral response pattern for water where the red band reflectance values are relatively higher than those of the infrared band).

Inputs to the calculation of the distance-based VIs are the red band, the infrared band, the slope of the soil line and intercept of the soil line. (In addition, some of these VIs also require a scaling factor.)

The first step in calculating the soil line is to identify a sample of bare soil pixels in the image. We will use the 90NDVI image created earlier to develop a mask image for bare soil. (If better knowledge of the area were available, we could illustrate digitize known bare soil areas.)

2. If you assume that any pixel having a higher infrared than red reflectance is vegetation and everything else is bare soil, what threshold value could you use with the 90NDVI image to separate vegetation from bare soils? (Hint: Use the NDVI equation with some example values to help you answer this question.)

3. Run RECLASS with 90NDVI to create the image SOILMASK. Assign the new value 1 to bare soil areas and the new value 0 to vegetated areas.

Once the bare soil areas have been identified, the values for those areas in the infrared and red bands are submitted to linear regression to calculate the soil line. The soil line calculation is not the same, however, for all the distance-based VIs. Some are based on a regression where the red band is evaluated as the independent variable, and some are based on a regression where the infrared is evaluated as the independent variable. Since we will be creating both types of distance-based VIs, you will need to run the regression twice to determine two soil lines.

4. Run REGRESS (from the GIS Analysis/Statistics menu) twice, between the MAUR90-BAND2 and MAUR90-BAND3 images, using SOILMASK as the mask image. Write down the slope (b) and intercept (a) values for the case in which the red band is treated as the independent variable and for the case in which the infrared band is the independent variable.

5. What are the slope and intercept when the red band is the independent variable? When the infrared is the independent variable? What is the coefficient of determination (r²)?

---

3. The equation written at the top of the REGRESS display is in the form y = mx + b, where y = dependent variable, b = intercept, m = slope, and x = independent variable.

---

Exercise 5-7 Vegetation Analysis in Arid Environments 242
The coefficient of determination is quite high, indicating that the relationship between red and infrared reflectance for these bare soil pixels is described well by a linear equation.

f) Run VEGINDEX three times to produce the distance-based VI's PVI, PVI3, and WQVI. For each VI, refer to the Help System section Determining Slope and Intercept Values under VEGINDEX to determine which soil line parameters to use for each particular VI. Also refer to the Vegetation Indices chapter in the IDRISI Guide to GIS and Image Processing for details about the equation used for each index.

4 What are the major differences you see in the displays of the three distance-based vegetation index images produced?

5 Is there a noticeable difference between these three images (on average) and the two slope-based images (on average) produced earlier? In other words, would you be able to separate the five output images into two families based solely on the resulting images?

The Orthogonal Transformation VI's

The final group of vegetation indices we will explore are the Orthogonal Transformation VI's. With these VI's, four or more bands of imagery are transformed into a set of new images, one of which describes vegetation. We will explore the use of the Tasseled Cap and Principal Components transformations for producing vegetation images.

The Tasseled Cap transformation uses a set of four MSS multi-spectral images to produce four new images. The Green Sniff or Green Vegetation Index (GVI) image represents vegetation. Other images produced represent Soil Brightness Index (SBI), Yellow Vegetation Index (YVI) and Non-Soil Index (NSI). The name of the transformation describes the shape of a plot of pixels in GVI-SBI space for an image having vegetation in many stages of development. The Tasseled Cap was developed to represent the most important information from a multi-band agricultural scene in only two images - GVI and SBI.

g) Run TASSCAP from the Image Processing/Transformation menu. Indicate that you will be using MSS data and enter the four bands for the 1990 scene. Give GVI as the prefix for the output files. This will produce four images called 90GREEN, 90BRIGHT, 90YELLOW and 90NSOSUCH. Display the four images. (Auto-display is disabled for modules that produce more than one output image.)

6 Why do you think the areas indicated as having high amounts of vegetation in the green vegetation image show low values in the soil brightness image?

The Tasseled Cap transformation uses global constants (i.e., the values don't change from scene to scene) to weight the bands being transformed. Because of this, it may not be appropriate to use in all environments. Principal components analysis, on the other hand, is a scene-specific transformation of a set of multi-spectral images into a new set of component images. The component images are uncorrelated and are ordered according to the amount of variation they explain from the original band set. The first of these component images typically describes albedo, or brightness, (which includes the background soil) and the second typically describes variation in vegetative cover.

h) Run PCA from the Image Processing/Transformation menu. Choose to calculate covariances directly. Enter 4 as the number of input bands and enter the four 1990 MSS images as input bands. Enter 4 as the number of components to be extracted. Give 90 as the output file prefix. Accept the default to use unnormalized variables. When the processing is finished, display the resulting four images, 90CMP1 through 90CMP4.

The tabular information produced by PCA indicates that the first component describes nearly 93% of the variance in the original set of four bands. All the input bands have high and positive loadings for component one. We might then interpret this component as describing the overall image "brightness." The second component has positive loadings for both

Exercise 5-7 Vegetation Analysis in Arid Environments
infrared bands and negative loadings for the visible green and red bands. It can be interpreted as an image describing vegetation, independent of the overall scene brightness. Components three and four describe little of the original variance and appear to represent atmospheric and other noise in the images.

The equation used for the GVI image of the Tasseled Cap transformation\(^2\) also weights the infrared bands positively and the visible bands negatively, though the weighting values are somewhat different. It is therefore not surprising to see great similarity between the second component image and the GVI image produced earlier.

**Comparing Vegetation Indices**

It is possible to visually compare all of the vegetation index images we have produced. Some obviously have better contrast than others. Some seem to show more variation within the low-value areas. However, without ground-truth information about the status of vegetation in the areas in 1990, we cannot determine which indices are most useful. What we will do is analyze the set of images as a whole to see what different characteristics are illustrated by the various indices.

To do this, we will submit all of the VI images we have created in this exercise to a principal components analysis. The IDRISI PCA module requires that input data be byte binary. Our vegetation index images are in real format. To convert the vegetation index images to byte binary format, we will use the module STRETCH. We will perform a linear stretch on each image to transform the original range of values to the new range of 0-255.

1. Run STRETCH with each of the seven VI images produced earlier. (Do not use 90OSUCH or 90YELLOW) Choose a linear stretch with 255 as the number of levels in the output. Give output filenames plus a "-BYTE" suffix (e.g., to create 90RATIO-BYTE, 90NDVI-BYTE). To save time, you may wish to use Macro Modeler.

2. Run the PCA module. Choose to calculate covariances directly and indicate 7 input bands. Enter the names of the seven stretched VI images. Choose to extract 4 components. Give VI as the output image prefix and choose to use standardized variables creating integer images (the last option). The output images will be called VIGMP1, VIGMP2, VIGMP3 and VIGMP4. Display these images.

The component images describe the most important "patterns" present in the 7 input vegetation index images. The first component image shows that pattern which is most common to all the input images. The second component image shows the next most important pattern remaining after the first has been removed, and so forth. The statistics produced by PCA include information about the present variance explained by each component and the weighting (loadings) of each input image on each component.

- **Compare VIGMP1 with the input stretched VI images. Which resemble it most? Are the loadings of those input images high compared to the others for that component?**

Recent research\(^6\) indicated that in a similar study comparing 25 VI images, the first component described a general vegetation index, including elements of greenness and soil background. The second component represented those VI's that correlated for soil background, and the third described soil moisture.

---

5. GVI = \(|0.2564215| + 0.5625557 + 0.0000506 + 0.0000337\) in the naming of the image file for this exercise, MAURUS0-RAND1 corresponds to IRS4 in the equation, MAURUS0-RAND2 to IRS5 and so forth.


---

Exercise 5-7 Vegetation Analysis in Arid Environments
Change Analysis using Vegetation Index Images

We will now undertake an analysis between the two dates of imagery. We will be concerned with identifying areas that have undergone significant change between 1980 and 1990.

b) Display MAUR30-BAND3, the near infrared band of the 1980 image, using the Grey scale palette and autoscaling with Equal Intervals.

Unfortunately, the data we have for 1980 has significant horizontal "striping" effects due to sensor miscalibration. It is, however, the best available data for that time and study area, so we will use it.

c) Choose any one of the vegetation indices you used with the 1990 scene and produce a corresponding image for the 1980 data. If you choose a distance-based VI, you will need to find new soil line parameters for the 1980 data since soil moisture conditions may be quite different between the two dates and areas of bare soil may have changed.

The most elementary of change analysis techniques is visual comparison.

d) Look at the VI image pair for the two dates and try to determine areas where changes in vegetation are evident. The striping that is apparent in the 1980 scene is an artifact of the sensor system. Use HISTO with the two vegetation images and note the average value for the entire image.

Are there differences in vegetation between 1990 and 1980?

The closest rain-gauge station to this area is the town of Mbout, located outside the image to the east. The station recorded approximately 200 mm of rain in 1980 and 240 mm of rain in 1990. Since rainfall and vegetation cover are highly correlated, we can expect to see generally higher vegetation index values in the area for 1990 than for 1980.

There are many quantitative methods we can use to analyze change between images. Here we will explore only one, simple differencing. For a more complete treatment of change analysis techniques, see the Time Series/Change Analysis chapter in the IDRISI Guide to GIS and Image Processing. You may use the data from this exercise to explore on your own many of the techniques presented in that chapter.

With simple differencing, we merely subtract one image from the other, then analyze the result. The critical move then becomes one of selecting an appropriate threshold for the difference image beyond which we consider real change, as opposed to ephemeral variation, to have occurred. Ground truth information would normally be used to identify these thresholds.

e) Use OVERLAY to subtract your 1990 image from your 1980 image. Call the resulting image 1980-1990. Use HISTO with 1980-1990 and change the class width to be small in relation to the range of values in 1980-1990. (The class width will differ depending on the particular VI you chose to use. Make sure there are at least 100 "bins" or divisions in the histogram.) Note the distribution of values, as well as the mean and standard deviation.

In the absence of ground truth information to guide our selection of a suitable change/no-change threshold, we will use the standard deviation. We will consider that only those pixels lying beyond two standard deviations from the mean in either the positive or negative direction constitute real change and those lying within two standard deviations represent normal variation. In a normal distribution, 90% of the values fall within two standard deviations. By setting this as our threshold, therefore, we are identifying the outer 10% of pixels as our significant change areas.

f) Use RECLASS with 1980-1990 and the mean and standard deviation values you found above to create a new

Exercise 5-7 Vegetation Analysis in Arid Environments 245
What is the distribution of positive and negative change areas in the study area? (Try to disregard change that is due to the sensor misregistration in the 1980 imagery.)

Optional:
Re-examine steps 1 through 6 for several other vegetation indices and compare the results. How much does the choice of vegetation index influence the final assessment of change?

Answers

1. Answers from this visual analysis will vary. In both images, the drainage patterns show up quite clearly as having higher vegetation index values, as one would expect. The differences between the two images are more subtle. If you reduce the upper display saturation points for both images (in Layer Properties), it appears that the NDVI image has relatively more positive-value pixels in the low vegetation areas and the ratio image has relatively more low-value pixels in the low vegetation areas. The normalization of the NDVI serves to minimize topographic effects and division-by-zero errors.

2. The equation for NDVI is: (TR - Red)/(TR + Red).

If we assume that vegetation has higher reflectance in the infrared than in the red, then the NDVI for vegetation would usually be positive. Therefore, our identification of bare soil pixels might be NDVI values less than or equal to 0. Because of the "just less than" wording of the RECLASS module, you might assign the new value 1 to values from -1 to those just less than 0.000001 and the new value 0 from 0.000001 to just less than 1. This would include the value 0 in the bare soil category.

3. When red is independent and infrared is dependent: Y=1.01+1.00X, r²=98.41. When infrared is independent and red is dependent: Y=2.36+0.98X, r²=98.41. The first number in the equation is the intercept and the second is the slope (numbers shown are rounded to two decimal places). Your equations will be slightly different if you used different reclassification criteria for the mask image than those given in question 2.

4. These three images are very different. The PVI3 image, in particular, identifies large contiguous areas of relatively higher vegetation that are not identified in the other two images (nor in the slope-based images). More of the low vegetation areas have relatively lower values with the PVI image than with the WDI images.

5. PVI and WDI could be distinguished from Ratio and NDVI because the former have lower values in the low vegetation areas. However, the PVI image would not easily be grouped with either "family.

6. The areas with much vegetation are darker than areas with less vegetation because vegetation is darker than bare soil.

7. The loadings for the VICMPI image are highest on the ratio, NDVI, PVI, and WDI images. The lowest loading is for PVI3, which verifies the PVI3 as being rather dissimilar from all the others. The highest loading for the second component is for PVI3.

8. The answers here will depend on the index chosen, but it should appear that the area, overall, has more vegetation in 1990 than in 1980. However, the grading of the 1980 image makes this difficult to assess.

9. The areas that show the greatest increase in vegetation from 1980 to 1990 seem to be those along the drainage system. Much of what is classified as negative change appears to be attributable to the grading problems in the 1980 image.
Tutorial Part 6: Land Change Modeler (LCM) Exercises

Land Change Modeler Exercises

Projects and Change Analysis
Transition Potential Modeling
Change Prediction
Validation
Dynamic Road Development
Habitat Assessment, Change and Gap Analysis
Species Range Polygon Refinement and Habitat Suitability
Biodiversity Analysis
Reserve Selection with Maxxan

Data for the exercises in this section are installed (by default—this may be customized during program installation) to a folder called \IDRISI Tutorial\LCM on the same drive that the IDRISI program directory was installed.
Exercise 6-1

LCM: Projects and Change Analysis

This next set of tutorial exercises will explore the basic functionality of the Land Change Modeler. By no means do these exercises cover the depth of all that is available. Several case study areas are used to illustrate best the section under consideration and the breadth of what LCM has to offer.

In this exercise, we will explore the Change Analysis tab within LCM. Here you will find a set of tools for the rapid assessment of change, allowing one to generate one-click evaluations of gains and losses, net change, persistence and specific transitions both in map and graphical form. Specifically in this exercise, we will look at the process of establishing a LCM project and performing a basic change analysis. For this we will use the first of several study areas – Central Massachusetts, USA – the home of IDRISI.

a) First we need to set our default Working Folder to CMA under the IDRISI Tutorial folder. Assuming the IDRISI Tutorial folder was installed with the default settings to the C drive, open IDRISI Explorer, click on the Projects tab, move the cursor to an empty area of the Explorer view and right-click the mouse button. Select the New Project option. Then, browse for the folder named IDRISI Tutorial/LCM/CMA. This will create an IDRISI project named CMA.

b) Now display the file named LANDCOVER85CMA using a qualitative palette of the same name. This is the region between the outskirts of Boston (Route 128/195 is at the eastern edge) and the core of Massachusetts. The resolution of the data is 60 meters. Then display the file named LANDCOVER89CMA. Although it may not be very evident at this stage, there was enormous change during this period, as we will see with LCM.

c) On the IDRISI toolbar, locate the Shortcut Input box. Click within it and press the letters “T”. The Land Change Modeler will immediately show up as the first entry. Click on the green arrow beside it and LCM will launch. Alternatively you can access LCM from Modeling/Environmental/Simulations Models menu.

d) If IDRISI Explorer is open, minimize it against the left-hand edge to make as much room as possible for LCM.

e) In the LCM Project Parameters panel, click on the create new project button and enter the text “CMA” (for “Central Massachusetts”). For the earlier landcover image, enter LANDCOVER85CMA for the later landcover image, enter LANDCOVER89CMA. For the basis roads layer, enter ROADS85CMA, and for the elevation layer, enter ELEVATION85CMA. You will have noticed that the default palette has filled in automatically. This is an optional element and any palette file can be used. Finally, click on the Continue button.

f) You are now presented with a graph of gains and losses by category. Notice that the biggest gain is in the residential (>2 acres) category. Notice that the default unit is cells. Change that to be hectares. The minimum size of the residential category (>2 acres) is approximately 1 hectare (actually 0.81 hectares).

g) Now click on the Contributors to Net Change button and select the residential (>2 acres) category. As you can see, it is mostly gaining from forest, and to a lesser extent agriculture (cropland and pasture).

1 Notice that some land is lost to smaller residential. What is this process called?

---

1. LCM works best on wide-screen or dual monitor displays. If your screen is capable of a higher resolution, you may wish to use the Display Properties feature in Windows to change it.
h) Now return to the gains and losses graph. Notice that most classes are primarily either gaining or losing land but that the open land category is doing both. Select open land in the Contributors to Net Change drop-down list.

What information in this graph would allow you to conclude that the major character of open land is secondary forest regrowth?

i) Now click on the Gains and Losses button again and change the units to ”% change.” Notice that this confirms that the open land category is very dynamic (as is the Barren Land category).

j) Change the units back to hectares and select deciduous forest from the Contributors to Net Change drop-down list. As you can see, open land is the chief contributor.

k) To complement these graphs, go to the Change Maps panel and click on the Create Map button. Notice that you didn’t need to specify an output name – it created a temporary filename for you. There are a number of cases in LCM where you may want to produce outputs in quick succession without necessarily keeping any of them (because you’re exploring). These will all indicate that the output name is optional. However, if you want to keep an output, give it a name!

The map you just created shows a bewildering pattern of change! Since we know that the biggest contributor to change is residential (>2 acres), we will now use the tools in LCM to see if we can begin to understand it better.

l) In the Change Maps panel, click on the Map the Transition option. In the first drop-down list (from), choose the All item. Then in the corresponding “to” box, choose the residential (>2 acres) category. Click Create Map. This shows all the areas that changed to the residential (>2 acres) category by the origin category.

m) Although we can begin to see a pattern here, we will use the spatial trend tool to see more detail. Expand the Spatial Trend of Change panel by clicking on its arrow button. Then select All in the Map Spatial Trend from drop-down list and then residential (>2 acres) in the “to” drop-down list. Leave the order of polynomial at the default of 3 and click the Map Trend button.

As you can see, this analysis takes considerably longer than the simple change analysis. However, it provides a very effective means of generalizing the trend. From this it is evident that the change to large residential properties is primarily concentrated to the north-east and south-east of the image.

n) Each in the Change Analysis panel, create a graph of the Contributors to Net Change experienced by cropland. Notice that in addition to losing land to development categories, it also loses to open land (i.e., secondary forest). Create a third-order trend of cropland to open land.

Comparing the trend map of change to large residential to the trend map for Cropland, what can you conclude about the main driving forces of change in this area of Massachusetts?
Exercise 6-2

LCM: Transition Potential Modeling

In this exercise, we will explore the Transition Potentials tab. This tab allows one to group transitions into a set of submodels and to explore the potential power of explanatory variables. Variables can be added to the model either as static or dynamic components. Static variables express aspects of basic suitability for the transition under consideration, and are unchanging over time. Dynamic variables are time-dependent drivers such as proximity to existing development or infrastructure and are recalculated over time during the course of a prediction.

Once model variables have been selected, each transition is modeled using either logistic regression or an extensively enhanced multi-layer perceptron (MLP) neural network. The result in either case is a transition potential map for each transition—an expression of time-specific potential for change.

For this exercise, we will use a data set from a rapidly changing area in the Bolivian lowlands known as Chiquitania. The data for this analysis were developed by and are used here with the permission of Conservation International’s Center for Applied Biodiversity Science at the Museo Noel Kempff Mercado in Bolivia.

a) Before we begin, we need to change our default Working Folder to the CT folder. Using IDRISI Explorer, click on the Projects tab, move the cursor to an empty area of the Explorer view and click the right mouse button. Select the New Project option. Then, browse for the folder named IDRISI Tutorial LCM/CT. This will create an IDRISI project named CT (short for “Chiquitania”).

b) Display the file named LANDCOVER8CT.

Chiquitania is about 200 km to the north/northwest of Santa Cruz de la Sierra – Bolivia’s boom town of petrochemicals and agribusiness in the Amazon basin. This is a region of rolling hills at the ecotone between the Amazonian forest and deciduous dryland tropical forest. It is not well suited to mechanized agriculture, but has economic value for both cattle and timber production. In addition, there is some subsistence agriculture. Note that the classification does not distinguish between settlements and agriculture. This map was intended for ecosystem monitoring and so both are designated as anthropogenic disturbance. This also includes secondary forest – once disturbed, land remains in that class. The vast majority of disturbed areas are used for pasture – either for dairy (primarily in the south east) and beef production.

c) For a sense of how the area is changing, now display the file named LANDCOVER94CT. In this tutorial exercise, we are going to model this change and predict what the landscape might look like in the future if the nature of development stays the same (this is important wording).

d) Go to the Window List menu entry and close all map windows. Then minimize IDRISI Explorer. Launch LCM and create a new LCM project called Chiquitania. Enter LANDCOVER89CT as the Earlier land cover map, LANDCOVER94CT as the Later land cover map, ROADS94CT as the basis roads layer and ELEVATION94CT as the elevation model. Notice that it automatically fills in the palette. This is because the landcover maps each have palettes of the same name as the image files. Now click on the Continue button.

In contrast to the change in the first LCM tutorial, this one is very straightforward. This is largely because of the definition of the disturbed class. It simply consumes the natural landscape!

---

1. This is not necessary, particularly if you have a wide-screen or dual monitor display. However, if you don’t, you will want the entire space.

2. The term “basis” here refers to the fact that it will be used as the basis for building new roads. In this case, the basis landcover map will become the basis layer for building new landcover changes.
e) Click on the Create Map button on the Change Maps panel. As you can see, the amount of change that has taken place between 1986 and 1994 is extensive and involved seven separate types of transition. However, some of these are quite small. For example, change the units to hectares and then move the cursor over the minimum bar in the gain and losses graph for evergreen (a palm important for its oil and thatch). Notice how the graph tells you the exact quantity. This amount of loss is as much likely to be map error as anything else — at a total of 27 panels out of almost a million in the entire image, it is not worth modeling. Therefore, click on the Ignore Transitions Less Than checkbox in the Change Maps panel and enter a value of 500 hectares in the Edit box beside it. Then click on the Create Map button again.

Notice how this has reduced the transitions to just 4 — the main transitions that are taking place in the area. In order to predict change, we will need (at any moment in time) to be able to create a map of the potential of land to go through each of these transitions. These maps will be called **transition potential maps**.

f) To model transitions, click on the second tab in LCM — the Transition Potentials tab, then expand the Transition Sub-Models: Status panel by clicking on its button.

**Important Note:**

Notice that there is a grid that lists all 7 transitions, but that only the first four are indicated as included in the model. This was caused by the area filter you applied on the Change Analysis tab to ignore minor transitions. It has given each transition a default name (which you may change at any time). **In order to predict change, you will need to explicitly model each of these four transitions. You have two tools to do this: logistic regression and a multi-layer perceptron (MLP) neural network.** If you use the former, then each of these transitions must be modeled separately. However, if you use MLP, you have the opportunity of modeling several or even all transitions at once. This is only reasonable if you think the driving forces for these transitions are the same and that a common group of explanatory variables can adequately model all of the transitions that are collected together into a sub-model. **If you wish to group several transitions into a sub-model, all that is required is that you give them a common name, as you will see in the sequence that follows.** Your final model can range from one that consists of a single sub-model describing all transitions to a separate sub-model for each transition.

For our purposes, it is reasonable to conclude that all four of the transitions have the same origin. Thus we will collect all four into a single sub-model.

g) We will use the Transition Sub-Models: Status tab to group all four transitions. Notice the left-most column in the grid signifying the transitions to be included (denoted by a yes in the column). We will group the four ‘yes’ transitions into a new group named disturbance. Click into the Sub-Model Name entry of the grid for each of the four transitions we are grouping together and enter the sub-model name “disturbance.” Notice that the drop-down list labeled Sub-Model to be Evaluated is automatically changed to “disturbance.” This determines what is being modeled in the panels on other parts of this tab.

h) Now comes the issue of which variables can explain the change that occurred from 1986 to 1994. Display the image named DIST_FROM_DISTRIBUTION86CT.

It is logical to assume that between 1986 and 1994, new disturbance tended to be near to areas of existing disturbance (for reasons of access). This map was created by extracting the disturbed areas from the earlier landcover image, filtering it with a 3x3 mode filter to remove extraneous pixels and then running the DISTANCE module on the result.

i) To see the nature of its relationship to change, go back to the Change Analysis tab and create a map of the transition from All to Anthropogenic Disturbance from 1986 to 1994. Call the output map CHANGEALL. Then, using CHANGEALL, use the module RECLASS to create a Boolean map of change called CHANGE8694.

---

3. We tested 12 techniques, including all of the procedures found in other landcover change models at the time of publication. Of these, only these two procedures surfaced as viable techniques, and our experience has been that the MLP is the most robust — hence it is the default.
Assign a 1 to all the old values from 1 to 999. Then use the module HISTO (click on the HISTO icon next to the GPS icon) and enter DIST_FROM_DISTURBANCE86CT as the input file and CHANGE8694 as a mask. Change the maximum value to display to be 10000 (meters). Then click OK.

As you can see, there is a very sharp decline in the frequency of change as we move away from existing areas, to the point where it drops to virtually nothing after four kilometers. This is a non-linear relationship. If we were to model using logistic regression, we would need to linearize it by applying a log transformation (using the Variable Transformation Utility panel on the Transition Potentials tab). However, we will be using MLP which is quite capable of modeling non-linear relationships. Therefore we will leave the variable as it is.

i) Go back to the Transition Potentials tab and click on the Test and Selection of Site and Driver Variables panel button. Click on the Pick List button for the Evaluate input box and select DIST_FROM_DISTURBANCE86CT. Then click the Test Explanatory Power button. This is a quick exploratory tool for seeing whether these might be some value in including that variable as part of your model. It indicates the degree to which the variable is associated with the distribution of landcover categories. Although it gives you an overall Gamma’s V (a measure of association that ranges from 0-1), it is the individual class values that are more important here. The values you see here are generally those one would expect from a variable with strong predictive power. Therefore click on the Add to Model button. Notice how the Transition Sub-Model Structure panel opens up with that variable as the first entry. As a contrast, test the explanatory value of the image named DIST_FROM_STREAM8CT. Overall it is not a strong variable. However, it does have some relationship to the location of areas of disturbance, so we will use it. Click on the Add to Model button for this as well.

k) Notice the model grid in the Transition Sub-Model Structure panel also allows you to enter variables directly. Click the Number of Files up-down button and increase this number to 6. Then enter directly into the grid the following variables: DIST_FROM_ROADS94CT, DIST_FROM_URBANC8, ELEVATION8CT and SLOPE8CT. Click on the individual Pick List button to add the files.

Notice that all of these variables are continuous quantitative variables. Both logistic regression and the MLP require this. What if we wanted to include a qualitative variable such as landcover? There are two ways we can do this. One would be to create a separate Boolean layer of each landcover class and add them to the model. In regression analysis, these are known as “dummy” variables. However, the downside is that this potentially increases the number of variables in the model substantially, which can impact model performance (a phenomenon known as the **Hueger** phenomenon). We will therefore use a different approach.

j) Open the Variable Transformation Utility panel and select the Evidence Likelihood option. Enter CHANGE8694 in the Transition or Land Cover Layer input box and the earliest landcover map, LANDCOVER86CT as the input variable name. Call the output EVLikelIHOOD LC. Be sure the checkbox correctly indicates this is a categorical variable. Then click OK. Notice that you now have a quantitative variable that you created from one that was categorical. It was created by determining the relative frequency with which different landcover categories occurred within the areas that transitioned from 1986 to 1994. The numbers thus express the likelihood of finding the landcover at the pixel in question if it were an area that would transition. Now test its potential explanatory power by entering EVLikelIHOOD LC into the Evaluate input box of the Test and Selection of Site and Driver Variables panel and click the Test Explanatory Power button. As you can see, anthropogenic disturbance has a strong association with the landcover types. This is logical when the change is for purposes of agriculture. Therefore add it also to your model. You should now have a total of 7 variables in your model as shown in the Transition Sub-Model Structure panel: DIST_FROM_DISTURBANCE86CT, DIST_FROM_ROADS94CT, DIST_FROM_STREAM8CT, DIST_FROM_URBANC8, SLOPE8CT, ELEVATION8CT and EVLikelIHOOD LC.

m) Now we come to the final transition modeling step. Close the Variable Transformation Utility panel and open up the Run Transition Sub-Model panel. The default procedure is the MLP neural network, which is what we
will use. Notice that it lists the number of cells that transitioned during the training period (1988-1994) for the smallest transition in the group of four that you are modeling in this run, as well as the number of cells that could have transitioned, but did not (i.e., persistence). This allows you to gauge the sample size you will use. Although you can select a smaller sample, there is no need to do so – just let it use the suggested size.

Important Note:

Before running the model, it is useful to explain briefly the MLP procedure since it is a dynamic process. The first thing that will happen when you click the Run Sub-Model button is that it will create a random sample of cells that experienced each of the four transitions we are modeling and an additional set of random samples for each of the cases of pixels that could have, but did not go through the transition. Thus the neural network will be fed with examples of eight classes, four transition classes and four persistent classes. We are only interested in the first four of these, but the neural network will be able to train best if it has all 8. We have designed a special automatic training mode that allows you to simply watch the training process and wait for it to finish. Although you can stop the training process at any point, make adjustments to the parameters, and then start it again, do not do so here – just watch what it does. The on-line Help System can give you more details about how the MLP works, but the key thing to understand at this point is that it is using the examples you gave it to train on and is developing a multivariate function that can predict the potential for transition based on the values at any location for the 7 explanatory variables. It does this by taking half the samples it was given to train on and it reserves the other half to test how well it is doing. The MLP constructs a network of neurons between the seven input values from the explanatory variables and the eight output classes (the transition and persistence classes), and a web of connections between the neurons that are applied as a set of (initially random) weights. These weights structure the multivariate function. With each pixel it looks at from the training data, it gauges its error and adjusts the weights. As it gets better at doing this, you will notice that the accuracy (determined from the validation data) increases and the precision improves (i.e., the RMS error declines). When the MLP completes its training, it is up to you to decide whether it has done well enough and whether it should re-train either with the same parameters, but a different random sample, or with new parameters. When you achieve an acceptable training, you will then need to click the Create Transition Potential button.

n) Now click the Run Sub-Model button and watch what happens. It may indicate that it needs to adjust the sample size. This is normal and just fine – it relates to the random selection process. Remember, just wait until it finishes its default 10,000 iterations. You should achieve an accuracy rate somewhat in the vicinity of 80%. If it finishes and it achieves less than 75%, click on the Run Sub-Model button again. Otherwise, click on the Create transition potential button. It will then create and display the four transition potential maps. These express, for each location, the potential it has for each of the modeled transitions.

This completes the first stage in developing a prediction. The transition potential modeling phase is extremely important as it has the largest bearing on the success of any prediction. In the next exercise, we will actually create a prediction. If you are taking a break at this stage and closing down LCM, be sure to save your LCM project (it will prompt you to do so). You will then be able to start up exactly as you left off.

Exercise 6-2 LCM: Transition Potential Modeling
Exercise 6-3

LCM: Change Prediction

In this exercise, we will use the transition potentials we modeled in the previous exercise to create several types of predictions. The Change Prediction tab in LCM provides the controls for a dynamic land cover change prediction process. After specifying an end date, the quantity of change in each transition can be modeled. We will use the Markov Chain analysis to model these transitions.

Two basic models of change are provided: a hard prediction model and a soft prediction model. The hard prediction model is based on a competitive land allocation model similar to a multi-objective decision process. The soft prediction yields a map of vulnerability to change for the selected set of transitions. In general, the results of the soft prediction are preferred for habitat and biodiversity assessment. The hard prediction yields only a single realization while the soft prediction is a comprehensive assessment of change potential.

In the next exercise, we will validate the results of our prediction.

a) If you closed LCM after the last exercise, launch it again and reload your LCM project (e.g., Chuquistania). You will notice that everything is filled in exactly as you left it. Now move to the Change Prediction tab and open the Change Demand Modeling panel. This is where you specify the end year of your prediction and consequently determine the amount of change that is going to happen. The default procedure for doing this is a Markov Chain. If you wish, you can choose to edit the transition probabilities or you can enter the transition probabilities as a data file from some external program. We will use the default option here and let LCM work out the quantities automatically. Therefore, simply enter 2000 as the prediction date (i.e., a 6-year prediction). We will do this because we have an actual image for 2000 which we can use to validate how well the prediction process works.

b) Next, open the Change Allocation panel. By default, it is set to create the prediction in one step. Notice also that by default, an option is checked for a soft prediction. Click this off for the moment and click the Run Model button. Notice that the prediction process takes 4 passes – one for each of the 4 transitions. The result is what is called a hard prediction – a prediction of a specific scenario for the future date (in this case, 2000).

c) When the hard prediction run has finished, click on the soft prediction option. You will notice that it now enables a grid that shows each of the included transitions. In this case, we will elect to include all of the transitions (the default option). Then run the prediction model again.

The result will be both a hard prediction and an additional map of vulnerability to the set of transitions selected. Since we are modeling four transitions to disturbance, the result is a map of vulnerability to anthropogenic disturbance.

The distinction between hard and soft prediction is very important. At any point in time, there are typically more areas that have the potential to change than will actually change. Thus, a commitment to a single prediction is a commitment to or “best guess” at just one of many highly plausible scenarios. If you compare the results to what actually occurred, the chances of getting it right are quite slim. A soft prediction, however, maps out all the areas that are thought to be plausible candidates for change. If the concern is with the links to habitat and biodiversity, this may be the better output format.

In both of the above cases, we modeled the change in one step. This is fine if all the variables in the model are static (i.e.,

---

1. Soft prediction is based on aggregating the transition potentials for each of the solicitated transitions. By default, a logical OR is used for aggregation. Very simply, this recognizes that the vulnerability to transition is highest if several transitions have interest in the same area.
they do not change over time). This is clearly true of the elevation, slope, distance from streams and likelihood of land cover variables. However, there is one variable in our model that is clearly dynamic rather than static – distance from disturbance. As new areas of disturbance emerge, the distance from disturbance changes. LCM incorporates the concept of dynamic variables in several ways.

d) Go back to the Transition Potentials tab. Find the entry for the variable named DIST_FROM_DISTURBANCE36CT in the Transition Sub-Model Structure panel. Notice that it is listed as being static by default. Click into the Role cell for this variable and change it to dynamic using the drop-down list box. Then click into the basis layer type column for this variable and select land cover. You will then be presented with a list of the landcover classes. Select anthropogenic disturbance in this case and click the Insert button to make it the dynamic landcover class. Then click OK.

e) Now go back to the Change Prediction tab. Since we have identified a variable as being dynamic, now set the number of dynamic variable recalculation stages to be 3. Check the Display Intermediate stage Images checkbox option out and the Create AVI Video option. Finally, change the output name to be LANDCOV_PREDICT_2000_D3. Be sure that soft prediction is turned on. Then click the Run Model button again. Notice that now there is a lot more work being performed. There are several differences with this analysis:

- At each stage, distance from disturbance is being recalculated.

- At each stage, the explanatory variables (including this revised one) are re-estimated to MLP Multi-Layer Perceptron and then applied to the previously calculated connection weights to the revised explanatory variables to calculate new transition potentials.

- The prediction at each stage calculates change in proportion to the number of stages.

- A video (in AVI format) is created of the images at each stage. This video can be viewed with Media Viewer in IDRISI or with any player that supports the AVI format.

f) You saw the intermediate results as they were created. Now open Media Viewer (from the Display menu), maximize it and select the AVI video named LANDCOV_PREDICT_2000_D3. Notice that you are seeing 4 frames in this video. This is because it starts from the landcover map which is the basis for the prediction – the 1994 map. Then open the AVI video entitled LANDCOV_PREDICT_2000_D3_SOFT and review the results.

g) When you are finished reviewing the prediction results, close Media Viewer. Display the final predicted image LANDCOV_PREDICT_2000_D3, then use Composer to add LANDCOV_PREDICT_2000 as an additional layer on top of it (and use one of the landcover palettes). In Composer, click the checkmark beside this top-most layer on and off. This will highlight the difference. In general, it is best to grow long predictions in stages in order for dynamic variables to be adjusted. You can have any number of variables that are designated as dynamic.

h) Now we will add new infrastructures and a constraint. Go to the Planning tab and open the Planned Infrastructure Changes panel. Click the spin button for the number of changes and indicate that there will be three new infrastructural development stages. In the first row of the grid, enter the file named NEW_ROADS_96CT and set the effective date to be 1996. For the second row, enter NEW_ROADS_98CT and 1998 respectively, while in the last row, enter NEW_ROADS_00CT and 2000.

---

2. Use this option with care. It uses a substantial amount of Windows resources to display images. With a prediction that has many stages, it is possible to completely exhaust available memory.

3. In actual use, these roads would be planned infrastructural developments. However in this case, the file indicates actual road developments so that we can validate how well the prediction process worked (next exercise).
Before we enter our constraint, display the image RESERVESCT.

This is a constraints map that delineates indigenous forest reserves (the black areas) in which transition potentials need to be lowered to reduce the possibility of development. A constraint and incentives image acts as a multiplier. A multiplier of 1.0 has no effect. Multipliers greater than 1.0 act as incentives (they increase the transition potential) while multipliers less than 1.0 act as disincentives. A multiplier of 0.0 acts as an absolute constraint. RESERVESCT is an image where indigenous forest reserves have been set a very low multiplier value (0.01). These are areas that were designated for indigenous forest use in the 1990’s by the Bolivian National Institute for Agrarian Reform (INRA). Traditional subsistence agriculture does lead to some forest conversion, but the rate is very low—hence the low multiplier. Thus it is not a hard constraint but rather a very strong disincentive. All other areas have been assigned 1.0.

To apply this multiplier, open the Constraints and Incentives panel. In the Incentives / Constraints map column of the grid, enter the image RESERVESCT for each of your four transitions.

Next we will need to set our roads layer as dynamic also. Go back to the Transition Potentials tab and change the DIST_FROM ROADS94CT layers to dynamic. Then click on the basis layer type entry and choose roads. You will then be presented with a dialog that shows the three road categories. Select primary, secondary and tertiary, click the Insert button and then click OK. This information will have more meaning when we run dynamic road building. However, we are activating this layer as dynamic now because the addition of new infrastructure needs to which explanatory variable needs to be updated with the new roads when they reach their implementation date and which road classes should be included in the calculation of distance from roads.

Now return to the Change Prediction tab and the Change Allocation panel. Under optional components, click on the apply infrastructure changes and zoning — constraints/incentives options. Then set the number of dynamic variable recalculation stages to six (i.e., each year will be modeled separately). If you have plenty of RAM and your screen is fairly clean of displayed images, turn on the Display intermediate stage images checkbox — you will find it interesting to see the effects of the new infrastructure as it is added over time. Otherwise leave it off, because you can view it in the AVI movie afterwards. Finally, set the output to be LANDCOV_PREDICT_2000_DCI6 (disturbance/constraints/infrastructure in 6 iterations) and then run the model. This output is required for the next exercise. Depending upon the speed of your computer, this will take between 3 and 10 minutes to complete. Notice the effect of the new roads and the forest reserves disincentive in your hard and soft results. After it finishes, also view the AVI movies for both the hard and soft outputs.

Try a long prediction (e.g., 30 or more years) and look at the impact of the number of dynamic stages on the end result (e.g., try it in 1 stage, then 2 stages, then 4 stages, etc.). Is there a point where it doesn’t make any difference?

---


5. Depending upon the context, you may find the need to designate different constraint/incentive images for different transitions. For any transitions for which no constraint or incentives apply, simply specify “none” (without the quotes).

6. Notice that when modeling landcover as a dynamic dataset, we started with a basis layer that was for the earlier year (1994) whereas when we are modeling roads, we use a basis layer for the later year (1994). This logic contained in the prediction process. Thus, the new roads for 2000 are used when the prediction for 2000 is formed.

Exercise 6-3 LCM: Change Prediction
Exercise 6-4

LCM: Validation

In the previous exercise, we created a prediction in both a hard (scenario) and soft (vulnerability) sense for the year 2000 based on information about the landcover in 1986 and 1994, and information about road developments and development constraints. How good was it? Given that the prediction was to 2000, we know the result! In this exercise, we will find out and explore the answer to determine its implications for predictive landcover change analysis. To continue with this exercise, you should have completed the previous exercise and have your default Working Folder set to the LCM/CT folder.

a) Open LCM and load the project used in the previous exercises, e.g., CT. The earliest land cover image should be LANDCOVER86CT and the later land cover image should be LANDCOVER94CT. The final land result from the previous exercise was named LANDCOVER_PREDICT_2000_DCI6. Display it and then display the image named LANDCOVER00CT. This is the actual landcover in 2000.

Clearly there is quite a difference. What is immediately apparent is that the quantity of change was far larger than what the history from 1986 to 1994 would have suggested. In fact, there was a major change as a result of the land reform process enacted in the mid-1990’s. In order to keep title to land, it was necessary to show that it was being used, which in turn led to a spike in deforestation in the late 1990’s.  This provides a first hand lesson about landcover change prediction – past history is not always a good indicator of the future.

b) Now go to the GIS Analysis menu and click on the Change/Time Series submenus to locate and launch the VALIDATE module. For the comparison image, enter LANDCOVER_PREDICT_2000_DCI6 and for the reference image, enter LANDCOVER00CT. Click off the Mask or Stata image option and click OK. If you look at the value in the % Correct row for your, you can see that it indicates that the prediction was actually quite good! This doesn’t seem to accord well with what we see. To learn a bit more, click on the More button.

Since we did not analyze by strata (regions), only two types of disagreement exist – disagreement in quantity and disagreement in location (grid cell). You can see that in absolute terms, these components are small. Notice also that the disagreement due to quantity was bigger than the disagreement in location and that the agreement in location is the largest component. We only modeled the transitions to anthropogenic disturbance so most of the area stayed exactly the same from 1994 to 2000 – hence the high agreement. VALIDATE tells us about how well we did with the entire map and not in a specific group of transitions.

To examine more carefully how we did with the specific task of predicting change to anthropogenic disturbance, we will use the Validation panel in LCM. Validation uses a three way crosstabulation between the later land cover map, the prediction map, and the map of reality.

c) Go to the Validation panel in LCM under the Change Prediction tab. The later land cover map is that stated as the later land cover image in the LCM Project Parameters panel, LANDCOVER94CT. Specify second image, the current prediction map, as LANDCOVER_PREDICT_2000_DCI6 and the third image as LANDCOVER00CT, the map of reality. Call the output VALIDATE_DCI6 and hit the Validate button.

The cases where we predicted correctly are called hits and are green. For example, looking at the legend locate the class 1|8|8 - Hits. This is the case where in 1994 it was woodland savanna, actually transitioned to anthropogenic disturbance in 2000, and we predicted the same transition. The cases where the we predicted change but in reality it did not are called

---

false alarms. Misses are the ones where we predicted no change but in reality it transitioned. Correct rejections are those cases we did not predict (background) that dominate the map (now it is easy to see why the accuracy rate was reported to be high by VALIDATE).

Notice that the misses are predominantly the large deforested areas away from the roads. These are the big changes by private owners rushing to establish claims to forested areas. The earlier history we had could not have predicted this. If we ignore these, then we notice that hits, false alarms and misses tend to happen in generally the same locations. This would suggest that we are able to get the general locations of change fairly well, but we have room for improvement on the specifics. If you look at the number of false alarms relative to the number of hits, you can see that our success rate is only about 25%.

Clearly we have room for improvement, but remember that this is a scenario — a hard prediction chosen from many equally plausible scenarios. Wherever there are more eligible locations for change than the actual amount of change, it is going to make it very hard to attain an accurate hard prediction. This is where the soft prediction comes into play.

d) Display LANDCOV_PREDICT_2000_Dc16_SOFT — the soft prediction that was created from your last run. Then add ACTUAL_CHANGE9400CT as a layer on top of it and choose the third uniform blue palette. Make the background of the layer transparent (using the rightmost button on Composers in the middle group). Notice that most of the areas that truly changed (with the exception of some of the large fields that resulted from the land tenure policy change) were considered to be vulnerable.

e) To qualify this, go to the GIS Analysis menu, Change/Time Series submodule and select the module named ROC. This module calculates the ROC statistic (also known as the Area Under the Curve ROC Statistic). It is used to determine how well a continuous surface predicts the locations given the distribution of a Boolean variable. In this case, you should specify LANDCOV_PREDICT_2000_Dc16_SOFT as the input image and ACTUAL_CHANGE9400CT as the reference image. Set the number of thresholds to be 100 and leave all other parameters at their default values. Then click OK. Your answer may be a little different because of the stochastic component of the MLP used in the model. However, you should have a value near to 0.80 — quite a strong value!

1 Given that there was a major policy change that had a huge impact on land cover change, what can you conclude about the relative benefits of soft prediction? What are the potential drawbacks?

2. We are ignoring misses because we know we had the quantity wrong. By computing hits to false alarms, we can evaluate the quality of the area that our model indicated would change.
Exercise 6-5

LCM: Dynamic Road Development

In the third exercise for LCM, we created a prediction for 2000 in which we were able to add new infrastructure as we went along. If we do not have any information on future roads, and if we plan on projecting long into the future, we run into a problem. Proximity to roads is typically a very strong factor in landcover change. If we project into the future without the roads growing along with development, our model is increasingly forced to make decisions without a critical component. In this release of LCM, we have included a tool for dynamic road development that attempts to predict where they will grow. This is the focus of this exercise.

In the Change Prediction exercise, we made our roads last dynamic and we selected secondary and tertiary for development. LCM uses the following logic: primary roads can grow secondary roads and can extend themselves, secondary roads can grow tertiary roads and can also extend themselves, tertiary roads can only extend themselves. Thus we have chosen to extend existing secondary roads and grow new tertiary roads.

a) If you have completed Exercise 6-3, set your default Working Folder to the LCM/CT tutorial folder. Then open LCM and select the LCM project used to complete Exercise 6-3 (e.g., Cuiquintanci).

Then, in the Change Prediction tab, open the Dynamic Road Development panel. We will use the default choices for road endpoint and route generation and also accept the default that all transitions play a role in deciding locations of high transition potential. The critical parameters we now need to set are the spacing and length parameters. Spacing refers to the frequency with which a road type occurs along a road of a higher grade. The length refers to how much they will grow at each stage. Notice that primary roads do not appear in this grid — they can only extend themselves and do so at the same rate as the secondary roads. For secondary roads, specify a length of 5 km and a spacing of 16 km. For tertiary, 3 km for the growth length and specify 8 km for spacing. Then set the skip factor to be 2. This means that it will grow roads only at every other stage. Notice that the output name has automatically been specified as ROADS_PREDICT_2000.

b) Next, open the Change Allocation panel and check on the Dynamic Road Development option under optional components. For this run, click off apply infrastructure changes. Again choose 6 dynamic stages, create AVI and calculate soft prediction. Leave the display intermediate stage images option off to save time (dynamic road building does take time). Change the output name to LANDCOV_PREDICT_2006_DRS and then run the model.

c) When the prediction finishes, launch Media Viewer and look at each of the three AVI videos it produced — the hard and soft predictions and the dynamic road development as well.

Try different spacing and growth length parameters for the road building. What appears to look most reasonable? How sensitive is the result to these parameters?
**Exercise 6-6**  
**LCM: Habitat Assessment, Change and Gap Analysis**

In this exercise, we will explore one of the features LCM offers to gauge the implications of change: the Habitat Assessment panel. This tool would typically be used to analyze the implications of change for a single species, such as an umbrella or charismatic species. Given information on landcover, habitat suitability and parameters related to the home ranges and dispersal characteristics of the species, it designates land as belonging to five different categories:

**Primary Habitat.** This is habitat that meets all the necessary life needs in terms of home range size, access to summer and winter forage, etc. Issues other than minimum area and required buffer size are specified by a minimum suitability on a habitat suitability map.

**Secondary Habitat.** This includes areas which have the designated habitat cover types, but which are missing one or more requirements (such as area or minimum suitability level) to serve as primary habitat. Secondary habitat areas provide areas of forage and safe haven for dispersing animals as they move to new areas of primary habitat.

**Primary Potential Corridor.** Areas of primary potential corridors are non-habitat areas that are reasonably safe to traverse, such as at night.

**Secondary Potential Corridor.** These are areas that are known to be traversed by the species in question, but which constitute much lower cover types.

**Unsuitable.** There are areas that are not suited for habitat or corridors.

The spatial inputs to the Habitat Assessment tool include one of your landcover layers and optionally a habitat suitability map. In this case, we will consider habitat for the bobcat (*Lynx rufus*) in Massachusetts.

a) As we did in the first exercise, use IDRISI Explorer to set your Working Folder to the CMA (Central Massachusetts) folder under the LCM IDRISI Twomul folder. Open LCM and reload the existing LCM CMA project used in that first exercise. It should load the earlier and later landcover images LANDCOVER85CMA and LANDCOVER99CMA, respectively. Click the Continue button and go to the Implications tab and open the Habitat Assessment panel. Click on the radio button for the earlier landcover map to select it as the focus of our assessment.

b) Next, using DISPLAY Launcher display the image named HABITATSUITABILITY85CMA and examine the values. This suitability map was created using the multi-criteria evaluation option of the Habitat Suitability / Species Distribution panel. The habitat suitability map was actually created in several steps as indicated in the extended footnote below.

c) The next step is to set the cover types that compose the bobcat habitat and the gap crossing distances within their home ranges and outside their home ranges. In the landcover grid set to include as potential habitat deciduous, mixed and conifer forest areas to be Yes and leave all others as No. These enter the following values for

---

1. The parameters used in this illustration were determined from a wide variety of radio collar studies and bobcat field reports. Although we could not find data specific to the Central Massachusetts area, we adopted parameters from studies in central Pennsylvania. Although we believe the parameters are generally reasonable, differences in prey density can change the home range size substantially. In addition, some gap crossing parameters could not be definitively established. This illustration is intended only to serve as a vehicle for discussing the nature of the parameters and the character of the mapped results. No scientific conclusions should be drawn or expected from this illustration.
the gap distances

<table>
<thead>
<tr>
<th>Category</th>
<th>Gap distance within range</th>
<th>Gap distance outside of range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial / Commercial</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Residential (~2 acres &amp; multi-family)</td>
<td>0</td>
<td>120</td>
</tr>
<tr>
<td>Residential (~3 acres)</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>Transportation</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>Other Urban</td>
<td>0.2</td>
<td>2000</td>
</tr>
<tr>
<td>Forest / Waste disposal / Mining</td>
<td>0.5</td>
<td>65</td>
</tr>
<tr>
<td>Cropland</td>
<td>0.2</td>
<td>100</td>
</tr>
<tr>
<td>Pasture</td>
<td>0.2</td>
<td>2000</td>
</tr>
<tr>
<td>Open Land</td>
<td>0.2</td>
<td>2000</td>
</tr>
<tr>
<td>Deciduous Forest</td>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td>Conifer Forest</td>
<td>0.1</td>
<td>2</td>
</tr>
<tr>
<td>Wetland</td>
<td>0.2</td>
<td>100</td>
</tr>
<tr>
<td>Water</td>
<td>0.2</td>
<td>2</td>
</tr>
</tbody>
</table>

d) Now we need to specify the area and buffer requirements for each category. For primary habitat, enter a minimum core area of 42.2 km² and a buffer distance of 120 m. The default minimum habitat suitability of 0.75 is correct by design. For secondary habitat, the corresponding values should be 1.55 km², 120 m and 0.5. For primary potential corridor, set the minimum edge buffer to be 120 m and the minimum habitat suitability to be 0.25 while for secondary potential corridor, set them to be 60 m and 0.0 respectively. Check to Consider the Habitat Suitability option and specify HABITAT,SUITABILITY99CMA as the suitability map. Notice that it specifies a default output name of HABITAT_STATUS_1985. This is fine. Now click on the Create Analysis button.

e) When the analysis has finished, set the Analysis radio button to be the later landcover map and change the habitat suitability map to be HABITAT,SUITABILITY99CMA. Change the output layer name to HABITAT_STATUS_1999. Then run this new analysis. When the analysis has finished, display both habitat maps and visually compare the change that has taken place between the two dates.

f) Open the Habitat Change / Gap Analysis panel. Change the units to hectares and specify HABITAT_STATUS_1985 as the first habitat status map and HABITAT_STATUS_1999 as the second. Then click on the Run Analysis button.

1 Examine the graph of changes in habitat. What does the graph suggest about habitat for the bobcat in Central Massachusetts?

2 For primary and secondary habitat, the MCR option was used to create an initial result from 0-1. These were then rescaled into a range from 0.75 to 1 for primary habitat and 0-0.75 for secondary habitat. Additional modifications were then added as follows:

   Primary Habitat: Factors in developing the primary habitat suitability component included proximity to conifer areas (winter foraging sites), proximity to summer foraging areas (mainly the boundaries between forest and forested/shrub wetland, secondary forest, pasture and open land sites), proximity to suitable dun areas (areas with deep talus) and the presence of forest. All proximity factors were translated using nearest points of 0 and 3000 meters (the maximum distance a bobcat will typically travel in a day). Aggregation of factors was achieved using a minimum function followed by applying a forest constraint. Within habitat, conifer was assigned 1.0, mixed forest was assigned 0.85 and deciduous forest was assigned 0.75.

   Secondary Habitat: Factors in developing the secondary habitat were identical to the above except for access to suitable dun sites. Forest categories were handled in the same manner as above. In addition, other wooded areas were assigned 0.65 and large residential (> 2 acres) areas were assigned 0.55. The aggregation method was also the same.

   Primary Potential Corridor: For primary potential corridors, a very simple logic was used. Open land was assigned a suitability of 0.68 and pasture was assigned 0.32.

   Secondary Potential Corridor: For secondary potential corridors, other wooded was assigned 0.20, coniferland was assigned 0.18 and large residential (> 2 acres) was assigned 0.10.

Exercise 6-6 LCM: Habitat Assessment, Change and Gap Analysis
g) Now click on the Protection Gaps radio button and specify HABITAT_STATUS_1999 as the habitat status map and PROTECTED_CMA as the protection map. Specify GAPS as the gap map Elnename and then click on the Run Analysis button.

2. What do you conclude about the degree of protection of bobcat habitat in Central Massachusetts?
Exercise 6-7

LCM: Species Range Polygon Refinement and Habitat Suitability

This exercise will explore species range polygon refinement for increasing the accuracy of habitat suitability and species distribution modeling. The tools needed are found in the Implications tab in LCM and the Species Range Polygon Refinement and Habitat Suitability/Distribution panels.

Species distribution models require information of presence or presence-absence data that are typically collected either through expensive and time-consuming fieldwork or from museum collections or herbariums. Because of the global deficiency of this type of data, especially for rare species, it is important to take advantage of species range polygon maps—species’ ranges developed and drawn by experts on map bases—for use as input for species distribution models.

The Species Range Polygon Refinement panel allows for the refinement of such range polygon maps of species distributions. This information is exceptionally valuable, but subject to error as a result of imprecision in the base maps, projection and geodetic datum errors, and limited geographical extent of expertise (i.e., the expert delineates only in the areas where she or he has expertise).

The underlying principle of the refinement process is to uncover the common environmental logic of the areas delineated by the range polygon. It does this by creating clusters of environmental conditions according to a set of environmental variables that the user believes can characterize the niche of the species. It then compares these clusters with the range polygon to determine the proportional inclusion of clusters within the range polygon. Clusters that fall wholly or largely within the polygon are assumed to describe essential components of that niche. Those that fall mostly or wholly outside are assumed to be unlikely components. The polygon is thus refined by removing areas that fall below a designated confidence.

To explore this technique, we will use the range polygon for the Vicugna vicugna (vicuña). The vicuña belongs to the camel family and is distributed along the Andes of southern Peru, western Bolivia, northwestern Argentina, and northern Chile. In the second part of this exercise, we will model the distribution of the vicuña.

a) First we need to set our default Working Folder to Vicugna under the IDRISI Tutorial folder. Assuming the IDRISI Tutorial folder was installed with the default settings to the C drive, open IDRISI Explorer, click on the Projects tab, move the cursor to an empty area of the Explorer view and right-click the mouse button. Select the New Project option. Then browse for the folder named IDRISI Tutorial LCM. Vicugna. This will create a new IDRISI project named Vicugna.

b) Once your default Working Folder is set, open the vector file named VICUGNA. This polygon was created by NatureServe1 and modified by Conservation International - Andes CBC to include only the distribution inside countries of their interest. Now, from Composer, add the vector layer SA_COUNTRIES and specify the black outline symbol file. As you zoom out, you will see more clearly where the range polygon falls within South America.

1 What country’s northern border does this ‘expert’ derived range polygon seem to abruptly end at?

To refine the vicuña species range polygon, we will use the following environmental variables:

---


Exercise 6-7 LCM: Species Range Polygon Refinement and Habitat Suitability 260
NDVI – mean
NDVI – seasonal variability

Elevation

Temperature - mean
Temperature - seasonal variability

Precipitation – annual variability

c) Open LCM and go directly to the Implications tab. Then click on the Species Range Polygon Refinement panel. Select vector as the range map file type to create a new environmental cluster map. This last option will create a cluster image based on the input environmental variables. This cluster result will then be used by the program to refine the polygon.

d) Next, select confidence as the output option. This option results in a continuous surface that is proportional to the percent of the area of the cluster falling inside the range polygon with values ranging between 0.0 and 1.0. Clusters falling wholly inside the range polygon will have a confidence of 1 while those wholly outside will have a confidence of zero. It is an empirical likelihood statement of confidence that indicates how confident we are that the area belongs to the species range polygon.

e) We now need to insert the environmental variables. Click on the Insert Layer Group button and add the raster group file ENV_VARS. Notice that our six variables are now loaded in the grid.

f) Finally, specify the input range polygon map VICUGNA, name the output cluster map CLUSTER, and specify to use the background mask MASK_WATER. Name the output confidence map as CONFIDENCE_VICUGNA. When all the parameters are set, click the Run button.

g) When the process is finished, it should display the new confidence map. Add the vector country boundaries using the white outline symbol file and examine the result.

2 What are the differences, spatially and in their attributes, between the refined range map and the original range map?

Habitat Suitability / Species Distribution

b) Now that we have created a confidence map for the vampire, we are now ready to create a habitat suitability map. Open the Habitat Suitability / Species Distribution panel on the Implications tab.

Land Change Models can either theoretically or empirically model habitat suitability for a species. Theoretical models allow the users to input expert knowledge about a species in the form of a set of rules. The modeling approach option available here would be multi-criteria evaluation. When presence or presence-absence data are available, empirical modeling techniques are available that empirically determine the set of rules about a species and its distribution.

IDRISI provides two empirical models that we presence only data to model species distribution. Mahalanobis typicalities and the weighted Mahalanobis. The difference between them is that the weighted Mahalanobis uses our confidence image produced earlier to weight the environmental variables. To calculate our new species distribution map, we will use this weighted Mahalanobis approach and the confidence map created in the previous exercise.

i) Select the presence option for the type of training data to use. Then select weighted Mahalanobis as the modeling approach and vector as the training site file type. Enter VICUGNA as the input training data file and CONFIDENCE_VICUGNA as the confidence image. Enter ENV_VARS as the layer group to load the environmental variables. Name the output NEW_VICUGNA and click the Run button.
Here we are using the same variables that we used to define the polygon. However, it is possible to use different variables. For example, if the interest is to predict the distribution of the species under conditions of global warming, you could include a map of future climate derived from models of climate change. You will want to explore more with these scenarios on your own.

i) When the process is finished, display the file named NEW_VICUGNA. Add the vector layer SA_COUNTRIES with a white outline symbol file.

How does this result now compare to the original polygon?
Exercise 6-8
LCM: Biodiversity Analysis

In this exercise, we will explore the calculation of biodiversity measures that are commonly used for decision making in conservation and planning. These measures include alpha diversity, gamma diversity, beta diversity, Sørensen’s dissimilarity index, and the range restriction index.

Alpha diversity is the simplest measure of diversity, often referred to as species richness. It refers to the diversity at a single location (e.g., pixel location or ecosystem) and is usually expressed as the total number of species.

Gamma diversity measures the regional richness by calculating the overall diversity across a larger region or across ecosystems.

Beta diversity measures the change in species diversity between locations (e.g., ecosystems). Sometimes beta diversity is referred to as species turnover as you move from one region to another.

Alpha diversity (a)
Gamma diversity (γ)
Beta diversity (β)

Sørensen’s dissimilarity index is a measure of species compositional dissimilarity. It measures the turnover of species composition across regions. In contrast to Sørensen’s index, Sørensen’s dissimilarity is measured as 1 minus Sørensen’s index, where Sørensen’s index is computed as the number of species that are common between the pixel and the region to which it belongs divided by the average alpha within the region.

The range restriction index (RRI) measures how restricted a species’ range is compared to the entire region. The measure ranges from 0 to 1, where 0 indicates all species at that location (pixel location) are unrestricted from anywhere in the entire study area while a value of 1 indicates that all the species at that location (e.g., pixel location) are completely restricted. This measure would be comparable to a level of endemism.

To explore these measures of biodiversity, we will use data for the North Andean Conservation Corridor (Norandean) which is part of the Tropical Andes Biodiversity Hotspot. It is one of the most diverse regions on earth in jeopardy from urban sprawl, mining, timber extraction, cattle ranching, and agricultural expansion. The Norandean corridor has an area of approximately 84,876 km² that covers parts of Colombia and Venezuela. It is also the last refuge for many species of mammals and birds.

For our exercise, we will focus on species of one particular class—amphibians. We will use species distribution polygon data generated by NatureServe under the Global Amphibians Assessment program (http://www.globalamphibians.org) and compiled by Conservation International – Andes Center for Biodiversity Conservation.

1. Beta diversity calculates the Whitlock beta diversity measure:

\[ \beta_{bet} = \frac{\gamma}{\gamma_0} \]

\[ RR_{\gamma} = \frac{\sum \left(1 - \frac{\text{area of Region}}{\text{area of Region}}\right)}{\alpha} \]

Exercise 6-8 LCM: Biodiversity Analysis
b) Display the file named NORTHANDIAN_HILLSHADE with a Grayscale palette. This is an analytical hillshade image created with the module SURFACE from elevation data. Now, from Composer, add the vector layer SA_COUNTRIES, found in the Vucuna folder, and specify the Outline Black symbol file. As you zoom out, you will see more clearly where the range polygons fall within South America.

c) Now add another vector layer named NORTHANDIAN_CORRIDOR. This is the area of the North-Andean corridor that we will focus on.

d) Open the Implications tab of LCM and select the Biodiversity Analysis panel. Select vector composite polygon as the specialization data then leave the selected all analysis types, Leave unchecked to delete generated layers. Although this will increase the amount of disk used, it will speed the process for the second part of this exercise. For the regional definition type, select focal zone and enter a focal zone diameter of 50 km. This is the extent of the regional area for which gamma diversity and Sorensen’s Dissimilarity index will be calculated.

e) Now we are ready to enter the filenames. Input NORANDIAN_AMPH as the composite species file. This vector composite file has 556 polygons corresponding to 556 species of amphibians. You can open the MDF file of the same name with Database Workshop to see the corresponding names of the species, taxonomy and status.

Input NORTHANDIAN_HILLSHADE and the reference layer for zonation. Select to apply a land mask and input NORANDIAN_WATER_MASK. This will avoid calculation in the ocean area. Then, in order, specify the following output filenames for the remainder of the inputs: ALPHA_FOCAL50, BETA_FOCAL50, GAMMA_FOCAL50, DISSIMILARITY_FOCAL50, and RANGE_RESTRICTION_FOCAL50.

When all the parameters have been entered, run the process by clicking OK.

f) When the process is finished, display each of the diversity measures and add the vector layer NORTHANDIAN_PROTECTED to each. These polygons represent the protected areas in the region. You can find the names for each protected area by opening the MDF file of the same name in Database Workshop.

Using the results, how is the region being protected in terms of local richness, regional richness, richness change, species turnover, and protection of endemics?

We will now continue with the biodiversity analysis, but instead of using a focal zone, we will calculate biodiversity measures for regions. In doing so, we will only be creating new beta and gamma diversity outputs. Alpha, Sorensen’s dissimilarity and range restriction can be recalculated but they don’t take into account ecosystems; they only use the focal zone for their calculation. The exception is RRI, which always uses the entirety of the study area for its calculation.

g) Display the vector file WWF_ECOREGION. This is a vector file of the eco-regions for Latin America and the Caribbean created by World Wildlife Fund. We will only use a small portion of this file, the northern region of South America, for those regions that fall within our Norandean corridor.

h) Before we run the biodiversity analysis again with our eco-regions file, we need to change the values of some of the parameters. Select species group as input for the species range data. This group file was created previously when we ran the first part of this exercise. Select vector region polygons as the input for the regional definition. For analysis type, select betas and gammas.

i) Finally, we will enter in the necessary input files. Enter WWF_ECOREGION as the input region polygon file. Next enter NORANDIAN_AMPH as the species group file and NORTHANDIAN_HILLSHADE as the reference layer for zonation. Choose to apply the mask NORANDIAN_WATER_MASK. Input the file

---

**Exercise 6-8** LCM: Biodiversity Analysis

267
ALPHA_FOCAL50 created earlier for the alpha diversity file and specify the two output files for beta and gamma as BETA_ECORREG and GAMMA_ECORREG. Then click OK to run the process.

When it has finished, compare the results of beta and gamma from the previous row. How do they compare? What does gamma tell us about the biodiversity of each eco-region? Which eco-region is more diverse? Which one is the least diverse?
Exercise 6-9
Reserve Selection with Marxan

Marxan is a planning software for reserve selection originally developed by Ian Ball and Hugh Possingham (2000) at the University of Queensland. Marxan reserve selection is based on a minimum set problem, where the objective is to achieve a particular species target at the lowest cost. Marxan generates new reserve networks and permits the evaluation of current reserve networks.

IDRISI’s Land Change Modeler application includes a front-end utility that calls the Marxan program. Note that only a subset of Marxan functionality is implemented in this version of LCM. In order to run this exercise, you will first need to install the Marxan program. Marxan is freely available from the University of Queensland and can be downloaded at: http://www.uq.edu.au/marxan/index.html?page=77554&ep=1.1.4.1. More information about the IDRISI Marxan front-end utility can be found in the Help System. The Marxan manual (available at the same site as above) can also be freely downloaded for additional details.

In this exercise, we will explore the use of Marxan to evaluate Bolivia’s current protected area network and select a new protected area network to fulfill a specific species area target.

a) Before starting the exercise, we need to set up default Working Folders to MARXAN under the IDRISI Tutorial data folder. Assuming the IDRISI Tutorial folder was installed with the default settings to the C drive, open IDRISI Explorer, click on the Projects tab, move the cursor to an empty area of the Explorer view and right-click the mouse button. Select the New Project option. Then, browse for the folder named IDRISI Tutorial Data\LCM\MARXAN. This will create an IDRISI project named MARXAN.

In order to run Marxan, the following input images are necessary.

Planning units map: This is the base map that will be used for the land allocation to define the protected areas. The planning units map should contain unique identifiers for each location that corresponds to a different planning unit. During a MARXAN run, each planning unit will be evaluated on whether it should be included in the reserve network. This map can be considered the minimum mapping unit for the protected area allocation.

Planning units can be specified in different ways. For example, it is possible to consider each pixel in the image a different planning unit. However, in reality, management of protected areas does not occur at a square pixel level. In this exercise, we will use the administrative units of river basins, ecoregions, and land use to identify the different planning units.

b) Display the map BOLIVIA_LU. This map shows land use in Bolivia in 2004. Disturbed areas are either urban or agriculture and will have a particular ID in the planning units map.

c) Display the map BOLIVIA_ROADS. This map shows the roads in Bolivia. Since the resolution of the image is 5 km, it represents a buffer of 5 km along all roads.

d) Display the map BOLIVIA_PA. This map shows the location of national parks in Bolivia and was extracted from the IUCN database of protected areas (http://www.iucn.org/rdp/). Each protected area will be given a unique ID in the planning unit since they are managed differently.

e) Display the map BOLIVIA_PROV. This is an administrative units map of the provinces of Bolivia.

f) Display the maps BOLIVIA_BASINS and BOLIVIA_ECO. These are maps of the country river basins and the ecoregions for Bolivia, respectively.
Given that reserve management can be constrained by administrative boundaries, the planning units map of the provinces was used. We then subdivided the provinces based on basins and ecoregions. Finally we added the information on land use, roads and protected areas.

Display the planning units map BOLIVIA_FU to view the unique planning units.

Species distribution maps: The species distribution maps are Boolean images with values of 1 in locations where the species is present and 0 where the species is absent. For this study, we utilized customized range polygon maps from the NatureServe database (www.natureserve.org).

Planning unit tenure (or planning unit status): The planning unit tenure map specifies which locations are available for selection in a final reserve system. Values of zero or one are given for locations that can be allocated to a reserve network. If a location has a value of 1, it will be included in the initial reserve system (but may not be part of the final result). If a location has a value of 0, it may be chosen in the initial reserve system, depending on the value indicated for the "starting proportion" parameter. A value of 2 is given for a fixed reserve system (such as the current reserve network), and a value of 3 is given for locations that are excluded from selection, such as particular land use types (urban areas, agriculture, roads, etc.).

The planning unit tenure map used in this analysis is derived from the land use map of Bolivia, the map of protected areas and the road map.

The map FU_TENURE has values of 0 for all available lands, values of 2 for the currently protected areas network, and a value of 3 for all roads and disturbed (agriculture/urban) land cover classes. This map will be used in the second part of this exercise.

The map FU_TENURE_PAASSESS is a modified tenure map created for evaluating the success of the current reserve network. In this case, current protected areas are assigned a value of 2 and all other locations are assigned a value of 3. This map will be used in the first part of this exercise.

Land cost layer: This layer specifies the cost of including the planning unit in the reserve system (for example, the cost of purchasing the land). This map is optional and if it is not included, the cost will be proportional to the planning unit size. For this exercise, we will not include this map.

Along with the input images, Marxan requires the following parameters.

**Target:** The target indicates how much of the species range needs to be protected and is specified in the number of cells.

**Species penalty factor (SPF):** This is a value given to a particular species or group of species to indicate its importance for inclusion in the reserve network. The higher the value, the more likely that species' target is met. There is no fixed rule on how to determine this value. The Marxan Tutorial recommends that you run Marxan with the specification of a uniform value for all species first, then evaluate the results. If with that particular SPF, all targets are not met, increase the SPF by a factor of two until all targets have been met. When that point is reached, lower the SPF slightly to see if all targets are still met. Once the base SPF is set up in this way, relative values can be applied to each species based on their ecological significance, vulnerability, rarity, etc.

**Boundary length file:** If reserve compactness is important and you want to consider this for reserve selection, select the checkbox.
Determining whether the current reserve network is protecting Bolivia’s endemic diversity

One of the uses of Marxan is to determine whether existing protected areas are fulfilling conservation objectives. For the purpose of this tutorial, we will specify a species conservation target to protect at least 50% of the range of distribution of Bolivia’s endemic species.

Marxan: Input and Output

b) Open Land Change Modeler and go to the Planning tab. From the Planning tab, open the Marxan Input and Output panel. Specify BOLIVIA PU as the planning unit layer. For the species distribution layers, specify the raster group file BOLIVIA_ENDEMICS.RGF. This contains the distribution of 73 species of mammals, birds and amphibians endemic to Bolivia. You will notice that the name column of the species grid will populate. Leave the default Type of 1 as we will first select a uniform SPF and target for all of them. In IDRISI, this can be accomplished automatically. All species that will have the same target and SPF should have the same type number. Then, in the Target % input box (percentage of the species range that needs to be protected to meet the conservation target), specify 50 and in the Penalty Factor input box, specify 10. Then click the Autofill Spec. Type button. The SPF and Target (in number of cells) will populate the species grid automatically for all species. These values will not be important to assess current protected areas but will be important when selecting new reserve areas.

d) Next, indicate that you wish to use a Planning unit tenure layer and enter PU_TENURE_PAASSESS as the name. For this exercise, we will not be utilizing a land cost layer or boundary length file. Specify an output prefix of ASSESS_CURRENT_PA. Click the Continue button and the Marxan: Parameters panel will open.

Marxan: Parameters

j) In the Marxan: Parameters panel, specify 1 in the Repeat runs input box. We are using a low number because we are not allocating new areas; we are just evaluating the current protection network. For the Species missing if proportion of target lower than input box, specify 0.95. This means that with a conservation target of 50%, the target will be considered met if the reserve protects 47.5% of the range or more (0.95 x 50 = 47.5). For the Run Mode, select the Use only the heuristic option and specify Greedy as the Heuristic type. Since we are only assessing current protected areas, we are choosing the fastest method. We will not utilize the Cost threshold nor will we specify a random seed. Set the Starting proportion to zero and use the default Clumping rule. Partial clumps do not count.

k) Click the Run Marxan button.

Results

For the evaluation of current reserves, the generated maps are not significant since we are not allocating new areas. We are interested only in the text outputs.

l) When Marxan finishes running, it will display two images and a log file. Close the two images.

The log file provides information on the total area of final reserves, existing reserves and newly added reserves as well as information on the species that are not protected under this reserve network. For each conservation feature (species) not protected, the log file provides the feature name, the target (amount of the range that we sought to protect), amount held in the network of protected areas, occurrences held (the number of reserves the species is present), and if the target was met (yes or no). The other options of occurrence target, separation target and separation achieved are not applicable in this implementation of Marxan and have values of zero. At the end of the log file is the number of species that have not
met the target with the current protected areas network.

This information is also included in the file called ASSISE_CURRENT_PA.MULTI.txt, saved in the Working Folders. This file is comma-delimited and can be viewed in IDRISI with the Edit module.

From the Target met column in the output text file, we can extract the following information (with the help of a calculator or spreadsheet program):

From the 73 endemic species in Bolivia (16 mammals, 21 birds and 36 amphibians), the protection target of 50% of range is fulfilled (target met) for only 18 species. Two mammals, 1 bird and 16 amphibians meet the target, representing 12.5% of the endemic mammals, 4.76% of the endemic birds and 41.67% of the endemic amphibians.

Select new protected areas to meet target

In this section, we will run Marxan to identify new protected areas that meet specified targets.

Marxan: Input and Output

We will use the same planning units layer BOLIVIA PU, as well as the same group file of species distribution layers BOLIVIA ENDEMIC RGF. In the Target % input box, specify 50 and in the Penalty Factor input box, specify 10. Then click the AutoFill Spec. Type button. Indicate that you wish to use a Planning unit tenure layer and specify the file PU_TENURE. We will not utilize a land cost layer. Indicate that you wish to use a Boundary length file in order to generate more compact reserves. Specify the output prefix as NEW_PA. Click Continue and the Marxan: Parameters panel will open.

Marxan: Parameters

In the Marxan: Parameters panel, specify 1000 in the Repeat runs input box and set the Boundary Length Modifier (BLM) to 2. The boundary length modifier determines how much emphasis should be placed on maximizing reserve compactness. It can utilize any positive value greater than zero, the larger the value, the more compact the reserve network. Since the BLM value depends on the study area, the user can try different values to achieve desired compactness. For the Species missing if proportion of target lower than input box, specify 0.95. For the Run Mode, use the default method Apply annealing followed by the iterative improvement algorithm. The default settings for the Annealing controls and Iterative improvement type also will be used.

For the Cost threshold, indicate that you wish to enable the threshold. This will generate reserves with costs less than the threshold value, or area (when no cost layer is used). Set 1600 as the Threshold (1600 pixels ~ 8000 km²). The penalty factor (cost threshold penalty) applies a penalty to the objective function if the cost (or area) of the selected reserve is greater than the threshold. The penalty factor A determines the size of the penalty. The higher the value, the larger the penalty for exceeding the threshold. A lower value for penalty factor A allows the threshold to move slightly above. Penalty factor B determines how gradually the penalty is applied. The higher the value for penalty factor B, the longer it will take for the penalty to be applied (e.g., applied to later iterations). Set Penalty Factor A to 9 and Penalty Factor B to 2. Set the Starting proportion to zero, do not specify a random seed and use the default Clipping rule. Partial clumps do not count.

Click the Run Marxan button.

Results

When Marxan finishes, it displays two images and a log file. For each run, Marxan generates a reserve network solution. The SUMMEDSOLUTION map provides for each planning unit the selection frequency across all runs. The larger the
value, the more likely those reserves are required in the reserve system to meet the conservation targets. The best solution map shows the solution for the run with the best objective value. Although it is called best solution, the Marxan User Manual states that this should only be seen as a very good solution, not as the best possible reserve system.

The log file here provides information on the total area of each reserve, existing reserves and newly added reserves as well as information on the species that are not protected under this reserve network. With this generated reserve network, the target of 50% protection was not met for 7 species.

This information is also in the file called NEW_PA_MVBEST.TXT, saved under the Working Folder.

From the Target met column of the output text file, we can extract the following information:

From the 73 endemic species in Bolivia (16 mammals, 21 birds and 36 amphibians), the new conservation system would allow the protection of 50% of ranges (target met) for 65 species. Twelve mammals, 19 birds and 34 amphibians met the target, representing 75% of the endemic mammals, 90.5% of the endemic birds and 94.4% of the endemic amphibians.

References