Modeling Past Vegetation Change Through Remote Sensing and G.I.S: A Comparison of Neural Networks and Logistic Regression Methods

Abdolrassoul S. Mahiny, Brian J. Turner

School of Resources, Environment and Society, The Australian National University, Canberra 0200, Australia
Telephone: +61 (0)2 6125 3519
Fax: +61 (0)2 6125 0746
Email: Rassoul.Mahiny@anu.edu.au

Abstract

Past change in remnant vegetation patches was modeled using remotely-sensed MSS and TM imagery and G.I.S. The images covering 27 years from 1973 to 2000 were used in detecting change in vegetation through post-classification comparison and modeling it through neural networks and logistic regression methods. Physical factors, image-based layers and landscape metrics provided the 19 predictor variables used in the modeling. The area of study is the catchment of the Boorowa River in New South Wales, Australia, around 110 kms northwest of Canberra.

Modeling decrease in vegetation patches over the past 27 years through neural networks and logistic regression methods proved successful. Relative operating characteristic (ROC), and a modified version of multi-resolution goodness of fit (MGF) tests together with visual comparison were used to assess success of the modeling approaches. Also, the relative effect of the 19 predictor variables were evaluated through ROC and MGF methods using 19 reduced-variable models and the full model.

Overall, the neural networks method performed slightly better than the logistic regression. A surrogate layer for agricultural intensity (MAXNDVI) and the ratio of MSS band 4 to NDVI (MSS4toNDVI) were found the most important predictor variables of change in vegetation. This was also the case with the logistic regression method where additionally the slope and Para (perimeter to area ratio of the patches) parameters were demonstrated to be the other important predictor variables. The neural network method was found to be more sensitive to inclusion or exclusion of the variables as compared to the logistic regression method. The increased number of variables also created a somewhat dispersed pattern of modeled vegetation as opposed to the naturally more clumped vegetation pattern.

1. Introduction

Past decades have witnessed a rapid development in technologies related to the exploitation of natural resources. This in turn has resulted in vast and unexpected increase in human abilities to access and use ever more remote areas for those resources. Coupled with the premises, population increase has also created more demands for goods and services derived directly from natural resources. As one of the consequences, vegetation cover has shrunk in search of more wood products and agricultural lands. Hence, the plant and animal species associated with the vegetation have been threatened by extinction.

Attempts to offset these impacts on flora and fauna have mainly focused on redirecting human activities to the less vulnerable lands, minimizing impacts through better designs and improved practice of development plans, and rehabilitating degraded areas. As nature’s functions and responses are normally manifested in a longer time span and usually beyond the administrative boundaries, it is needed to address the effects on these phenomena on a similar scale. In other words, rather than assessing impacts at one point in time – the old static way – or one scale, and designing the relevant rehabilitation practices for that single time frame or scale, a more dynamic approach should be undertaken. This is more consistent with the natural phenomena and associated responses to various activities. This is why nowadays, modeling change in natural elements in the past and tracking those changes in future, is considered a necessary part of most development plans either for exploitation or protection of natural resources. These models can also be used to explore the interactions of natural elements or to evaluate the proposed management plans (Baker, 1989).

Modeling vegetation change in the past and forecasting it in the future has become more common in recent years because of easy access to remote sensing imagery and developments in the capabilities of Geographic Information Systems (G.I.S). These developments have made it possible to formulate and test different methods of change detection and modeling through time. As a result of flourishing investigations of the subject matter over the past decades, there are now many methods available to researchers and interested practitioners. For example Lambin (1997) gives a general overview of the most common modeling approaches used in landscape ecology, citing Markov chain analysis, cellular automata, and statistical techniques such as logistic regression as the common methods. Also, in recent years there has been a surge in the use of artificial intelligence such as neural networks in ecological applications and the related disciplines. Eastman (2002) suggests that Markov chain analysis is a convenient tool for modeling land-use changes when changes and processes in the landscape are difficult to describe. He further elucidates that a Markov process is simply one in which the future state of a system can be modified purely on the basis of the immediately preceding state.

Weng (2002) used Markov chain analysis to model land use change dynamics. He concluded that using Markov analysis was beneficial in describing and analyzing land use change processes. Also, Hall and others (1991) used Markov chain analysis together with a series of MSS imagery to develop the transition rates of change between forest ecological states. The study area was the Superior National Forest in the USA and the analysis at the regional scale proved to be producing acceptable results.

However, Eastman (2002) also points to the weakness of the Markov analysis for modeling change in land-use/covers too. He states that the Markov analysis does not hold any sense of geography, so no spatial component will be produced in the modeling outcome. In other words, the results from the analysis hold true on a per category basis, but as there is no understanding of the co-occurrence and juxtaposition of the different categories in a landscape, the final results might not be accurate as desired. Cellular automata can add the spatial character to the results of Markov chain analysis. Cellular automata was first introduced by Von Neumann and Ulma (Wolfram, 1986) as simple models in which to study biological processes such as self-reproduction. Cellular automata is a system of dynamic modeling where the future state of a cell is determined deterministically using its current state and that of its neighbours. Jenerette and Wu (2001) used a combination of the Markov-cellular automata methods to simulate land use change in the central Arizona. They were able to effectively capture the pattern of land use change associated with urbanization.

Logistic regression is a statistical method to evaluate the relationship between a set of independent variables and a categorical binomial dependent variable. In logistic regression, the dependent variable must be binary in nature and can only take two values (0 and 1). As such, to test and describe the possible relationship between one or more continuous independent variables and the binary dependent variable, logistic regression is used with the assumption that the probability that the dependent variable takes the value of 1, follows the logistic curve and its value can be estimated with the following formula (Eastman, 2002):

$$P(y = 1 | X) = exp(SBX) / 1+exp(SBX) \qquad (1)$$

where P is the probability of the dependent variable being 1, X is the independent variable and B is the estimated parameter. To linearise the above model, and to remove the 0/1 boundaries for the original dependent variable, the following transformation is usually applied:

$$P^{'} = log_e (P/(1+P)) \qquad (2)$$

This transformation is referred to as the logit or logistic regression transformation. By performing the logit transformation on both sides of the above logit regression model, we obtain the standard linear regression model:

$$log_e (P/(1+P)= b_0 + b_1 * x_1 + b_2 * x_2 + \ldots + b_k * x_k + error\ term \qquad (3)$$

Logit transformation of the dichotomous data ensures that the dependent variable is continuous and the predicted probability is continuous within the range from 0-1 (Clark and Hosking, 1986). A logistic regression done on pixels of an image as above produces a layer containing real scores from zero to 1. In order to produce the final simulated map showing the changed areas, the modeled layer is ranked in ascending order and exactly the same number of pixels detected as change in the real change map is selected from those having the highest scores. Using three sets of environmental factors, spatial factors and patch attributes, Hsu (2000) was able to model the change in land-use in Maolin township in Taiwan through logistic regression.

Artificial neural networks (ANN) are powerful tools that use a machine learning approach to quantify and model complex behaviour and patterns (Pijanowski *et al*. 2001). Neural networks are capable of predicting new observations on specific variables from other observations on the same or other variables after executing a process of learning from existing data (StatSoft, 2002). Being one of several data mining techniques, neural networks consist of several interconnected layers of processing elements or nodes. A neural network has an input layer, one or more hidden layers and an output layer (Figure 1). Each layer contains a certain number of nodes where each of them is associated with an independent variable. Each of the nodes in the output layer is associated with a class or state in the output layer (McCormick, 1999). Where there is a suspected non-linear relationship between dependent and independent variables and the nature of the relationship between cause and effects is vague, neural networks offer a promising tool. This tool becomes even more important when the diversity of data used in the modeling is high (Schultz *et al*., 2000).

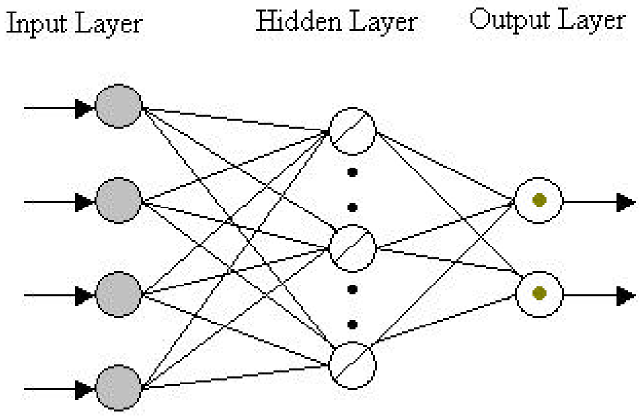

Figure 1. Schematic representation of a neural network with 1 input layer consisting of 4 nodes, 1 hidden layer with several nodes and 1 output layer with two nodes or classes (adapted from Lek and Guégan, 2000)

For instance, when modeling decrease in vegetation, a binary change layer can be used to extract information from the independent layers in the form of pattern files, and then subject them to learning and testing in neural networks. In this case, there is one input layer to the network consisting of as many nodes as there are independent layers, one or more hidden layers with a subjective number of nodes (normally three times that of the highest number of nodes in either of the network's input or output layers), and one output layer with one node. The output layer containing real numbers ranging from 0 to 1.0 depicts the likelihood of change for each pixel in the image.

Mann and Benwell (1996) used neural networks in conjunction with regression analysis to predict and monitor land condition over a larger area of Central Otago, New Zealand. Using bio-physical variables together with the management variables, they tried to predict the amount and distribution of bare ground within a pastoral system. They were successful in simulation with the neural network producing better results than the regression method.

The end result of such modeling methods is a data layer consisting of cells with continuous scores varying from zero to 1. Hence, the higher the score of a pixel the higher the probability of change for that pixel can be. After selecting the candidate pixels from among those most likely to change, the result is a boolean image containing zeros and ones.

The next step is to evaluate the goodness of fit of the simulated image to the real image. Of the possible methods to do so, the relative operating characteristics (ROC) is a technique applicable where the end result of simulation is an image containing zeros and ones. The ROC has also the advantage that: (1) it uses an index rather that percent of success, (2) it measures performance of modeling over a variety of scenarios of change, and (3) it illustrates the validation with a figure that shows clearly how a high agreement with reality differs from a low agreement (Pontius and Schneider, 2001). These qualities are not usually found in other validation methods such as overall accuracy or kappa index of agreement. When there is a perfect match between reality map and the modeled one, the ROC takes a value of 1. In the case where there is no spatial agreement between those maps the ROC value becomes 0.5.

For calculating ROC, first the simulated map is sliced into several percentile groups in a way that for example 10% of the cells that have the largest suitability values are assigned to group 1, the next 10% to group 2 and so on. Each of the percentile maps is called a scenario and is contrasted with the reality map. In each scenario, if a cell is simulated as change in both reality and the modeled map, it is considered a ‘true-positive’. If on the other hand, a cell is categorized as change in the modeled map but is actually a no-change in reality map, then it is called a ‘false-positive’. Plotting ‘false-positive’ against ‘true-positive’ produces the ROC curve (Pontius and Schneider, 2001). The area under the curve that connects two points in the curve is the ROC statistic. If the arrangement of cells in the simulated map is totally random then the ROC statistic will be 0.5. Increasing similarity between the two maps makes the statistic closer to 1. The ROC technique assesses the success of modeling on a cell-by-cell basis. In doing so, it loses sight of the spatial arrangement of the modeling result when comparing it with the reality map. As such, the ROC should be supplemented with visual comparison and additional measures of association that accounts for spatial patterns (Costanza, 1989). The multi-resolution goodness of fit (MGF) is one such method.

The basis for the MGF is that the lack of fit between two maps can be partitioned into ‘registration’, ‘resolution’ and residual components. So two maps with exactly the same number of pixels but slightly different position or resolution can produce a low similarity index using the conventional accuracy estimation methods. In other words, it is believed that measurements at one particular resolution will always be insufficient to describe complex natural phenomenon (Costanza, 1989). So, by using increasingly larger windows in the comparison process, it is possible to tackle the resolution and mis-registration problem. In doing so, increasingly larger window sizes are used to match the simulated map with the reality map through the following formula (Costanza, 1989). Hence:

$$F_w = \frac{\sum_{s=1}^{t_w} \left[1 - \frac{\sum_{i=1}^{p} \left| a_{ki} - a_{ji} \right|}{2w^2} \right]}{t^w} \tag{4}$$

where F_w is the fit for sampling window size w , w is the dimension of one side of the square sampling window, a_{ki} and a_{ji} the number of cells of category i in scene k and j in the sampling window, p the number of different categories in the sampling window, s the sampling window of dimension w by w which slides through the scene one cell at a time, and t_w the total number of sampling windows in the scene for window size w. Normally, the fit between two maps increases rapidly with the first window sizes and then reaches a plateau. The total fit as a weighted average fit across different window sizes is calculated as below (Costanza, 1989).

$$F_t = \frac{\sum_{w=1}^{n} F_w e^{-k(w-1)}}{\sum_{w=1}^{n} e^{-k(w-1)}} \qquad (5)$$

where F_t is the weighted average of fits over all window sizes, F_w is the fit for sampling windows of dimension w , k is a constant, and w the linear dimension of a sampling window. The constant k ranges from 0 to 1.0 depending on the weight one wants to put on large versus small window sizes. Obviously, larger window sizes contain more of the pattern information as opposed to the smaller ones, which emphasize exact location of the pixels. Based on the literature (Costanza, 1989) and findings of this study, a k of 0.1 was used in the present study.

In an attempt to assess the performance of the artificial neural networks and logistic regression in modeling change in vegetation, these methods were used and compared through visual, the relative operating characteristics (ROC) and a modified version of the MGF validation methods (Costanza, 1989). The decrease in vegetation over 27 years of time was detected through post-classification comparison of classified MSS and TM images. Then, modeling change in vegetation was conducted by using 19 predictor variables and the binary change map as dependent variable showing areas of decrease as ‘1’ and other areas as ‘0’. The predictor variables were from physical variables, image-based information and landscape metrics. The relative effects of 19 predictor variables on the modeling success were also assessed through the above methods. The assessment was conducted through dropping one variable at a time and calculating the ROC and MGF for the new images.

2. Area of Study

The area of study is the catchment of the Boorowa River in NSW, Australia, around 110 kms northwest of Canberra (Figure 2).

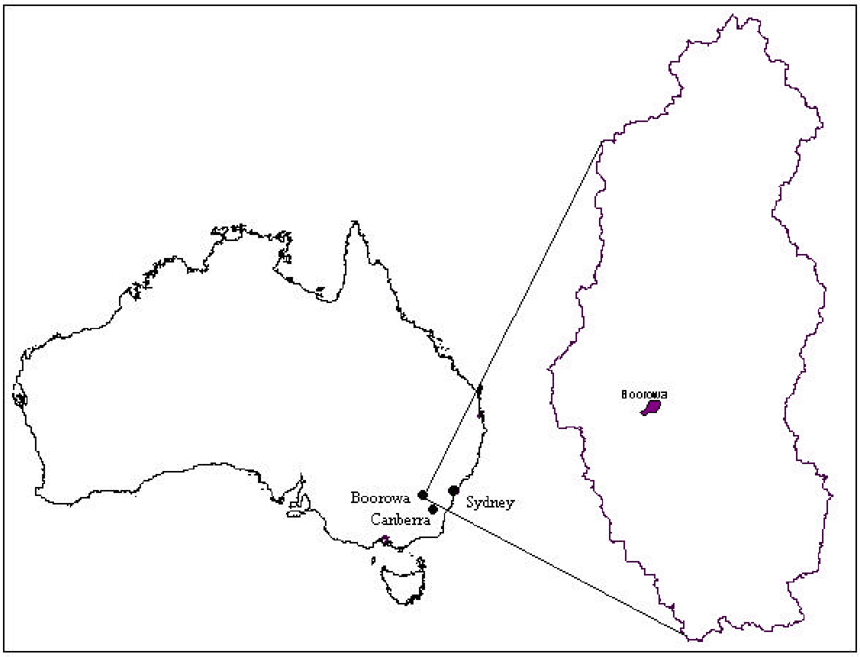

Figure 2. The area of study showing the Boorowa catchment location

The catchment covers an area of about 220 thousand hectares and is relatively flat with undulating hills on the east, north and west. Pastures, farmlands, remnant woodlands and small towns are the major land-cover/land-uses covering the area. The average annual rainfall ranges from 570 mm to 770 mm (Hird, 1991). The area has a warm climate with long summers and cool to cold winters. The maximum temperature recorded is 42.8 ^{0} C and the minimum -8.9 ^{0} C. Studies show that the pre-European native vegetation of the area was composed of the following categories (Hird, 1991):

1) Dry Sclerophyll Forest;
2) Woodlands; and
3) Grasslands.

Eucalypt forests are dominant on the hilly country of the eastern edge of the catchment where the average annual rainfall exceeds 640 mm (Hird, 1991). The woodlands are mostly found on the lower slopes and plains. Also minor areas of grasslands can occur in small pockets in the woodland environment. The area has been subject to land reclamation in the form of woodland conversion to agricultural lands for nearly two centuries. Almost all native communities comprised of eucalypt species are thought to have been cleared or modified to some extent by agriculture or grazing (Yates and Hobbs, 1997). As such, woodland patches have shrunk and remain mainly in the hilly eastern areas (Figure 3). Areas of poor agricultural quality also harbor small patches of woodland (Newham, 1999).

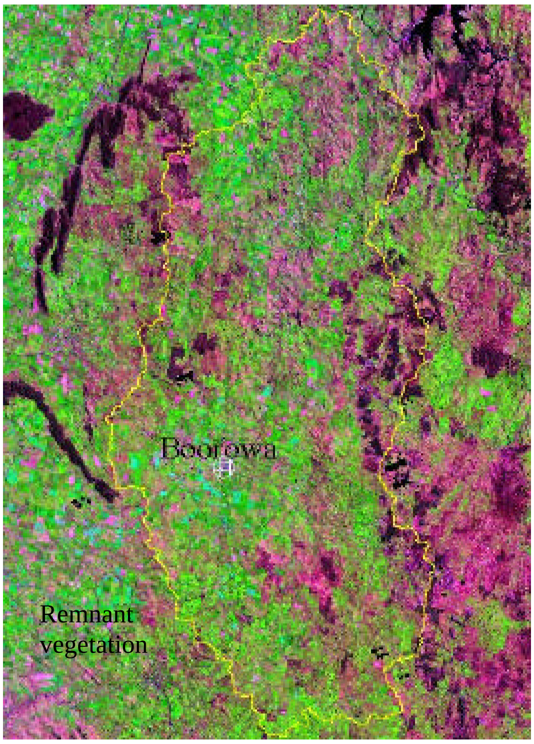

Figure 3. False color composite image of the Boorowa region. Landsat TM Bands 2,3,4 for year 1991. Darker areas are remnant forest

3. Data and Methodology

The present study was conducted in five steps. First the needed data were acquired and change in vegetation over the past 27 years was detected in a post-classification comparison method. Nineteen variables as predictors of change were defined and prepared in the next step. Then the changed areas were used as the dependent variable in the neural networks and logistic regression techniques with the 19 independent variables described below. Afterwards, performance of the two methods was assessed through the visual, ROC and MGF methods. Finally, the influence of different independent variables was evaluated using the cited validation methods.

3.1Change Detection

For the change detection study, a Landsat MSS scene for the year 1973 and a TM scene for 2000 were obtained (Table 1). This was done to provide the most extensive time frame for change detection in the area of study. Also, the choice of imagery was influenced in part by the availability of cloud free images.

Table 1. Image data sets used for post-classification change detection

Imagery	Date	Path / Row	Sun Zenith	Sun Azimuth
Landsat MSS	18 Jan.1973	91/84	22.91^0	301.90^0
Landsat TM	22 Oct.2000	91/84	38.91^0	61.80^0

The images were individually classified using past and newly sampled training areas. The training areas were regrouped into two classes of woody and non-woody areas. Also, in order to achieve a classification as accurate as possible, training pixels were "purified" using the method described by McKormick (1999). To do so, first the bands of MSS and TM images were subjected to the ISOCLASS clustering technique with a maximum of 100 clusters in Erdas Imagine (Version 8.5). Then the training samples were intersected with the image containing the clustering results for each date. Finally, only unique clusters or those accommodating a high proportion of one training class and a low proportion of the other were selected as training areas for those classes.

Applying the purified training samples and neural network classification with only two classes of vegetated and non-vegetated areas resulted in a highly accurate thematic map for the two dates. The results were directly used in post-classification comparison to detect change. The post-classification comparison method for change detection is one of the common methods least affected by difference in image characteristics used in the change detection. It is also rather straightforward for implementation producing relatively acceptable results (Brogaard and Prieler, 1998; Caccetta *et al*., 1998; Foody, 2001; Sing, 1989).

3.2 Derivation of independent variables

There have been quite a few studies focusing on flora and fauna of the Boorowa catchment or adjacent areas (Austin *et al*., 2000, Newham, 1999, Cox and Goldney, 1997, NSW National Parks and Wildlife Service, 2002). However, none of the cited studies entail modeling change. So, it was not clear initially which variables might play an important role in making areas prone to deforestation. Hence, as many predictors of change as possible were collated, in order to both test them and to maximize the discriminating power of the modeling approaches used (Table 4). Considering the above facts and the result from a review of the literature on the relevant modeling approaches for other areas led to the definition of the following groups and member predictor variables.

3.2.1 Physical variables

Of the physical variables, elevation, slope, and aspect were created using the base digital terrain model (DTM) with 25 meter resolution and the Idrisi software (version 32). Also, to investigate the possible effects of remnant patches location on the deforestation process, 11 components of topographic features including peak, ridge, saddle, flat, ravine, pit, convex hillside, saddle hillside, slope hillside, concave hillside, and inflection hillside were determined from the same DTM layer. Additionally, the geology layer was acquired from another source (Newham, 1999), edited and used as another physical layer.

3.2.2 Image-based parameters

The normalised difference vegetation index (NDVI) was calculated using the MSS bands 2 and 4. Then the layer was subjected to pattern analysis resulting in three new data layers viz., *Number of Different Classes* (NDC), *Fragmentation Index* (FI), and *Dominance Index* (DI). In the *Number of Different Classes*, the distribution of different NDVI indices over the window filter, which in this instance was chosen to be 5*5 is assessed. Idrisi software (version 32) was used for the calculations.

$$NDC = \text{number of different classes in each 5 x 5 neighborhood} \qquad (6)$$

The *Fragmentation Index* measures the number of different classes in a filter window such that:

$$FI = (n-1)/(c-1) \qquad (7)$$

where n is the number of different classes present in the filter window, and c the number of cells considered. The *Dominance Index* is a measure of the difference in maximum diversity in a window and the total diversity in the image.

$$DI = Hmax – H \qquad (8)$$

where H is diversity. The diversity itself is calculated using the below formula:

$$H = -? \ (p_* ln(p)) \qquad (9)$$

here the summation is over all classes in the entire image, p is the proportion of each class in the kernel and ln is the natural logarithm.

The above layers were produced for the base NDVI layer generated from the MSS 73 in order to extract as much information on the vegetation quality and quantity of the area as possible and subject it to modeling methods.

Agricultural activities and their intensity around the remnant vegetation patches in the form of crop and wool production were postulated as having effects on the quality and the quantity of the vegetation patches. These activities normally entail application of fertilizers, cultivation of non-native species, and sheep and cattle grazing in and around remnant vegetation patches. As there was no suitable land-use map available depicting agricultural intensity, a surrogate layer was constructed using the maximum NDVI of 7 MSS and TM images over the past 27 years (Table 3) on the assumption that the maximum NDVI over the cited period reflects to some degree the agricultural activities and their intensity. This idea was based on the fact that high crop production is reflected in the NDVI layer with proportionally higher indices. Covering the whole 27 years time span, the individual NDVI layers with the woody areas masked out would capture highly cultivated areas in different times of the year. As such, the intersection of those NDVI layers showing the highest index for each pixel can be taken as a reflection of the agriculture intensity in the area.

Table 3. Image data sets used for NDVI and agricultural density layer calculation

IMAGERY	DATE	SUN ZENITH	SUN AZIMUTH
Landsat MSS	18 Jan.1973	22.91^0	301.90^0
Landsat MSS	05Feb.1980	45.73^0	78.90^0
Landsat MSS	23 Feb.1985	45.08^0	67.98^0
Landsat MSS	16 Feb.1988	44.43^0	71.39^0
Landsat TM	08 Feb.1991	44.74^0	76.15^0
Landsat TM	22 Oct.1997	38.91^0	61.80^0
Landsat TM	2 Sep.2000	48.39^0	76.80^0

Studies on vegetation characteristics in the state of Queensland have revealed that the proportion of TM band 5 to NDVI plays a major role in detecting different densities of foliage projective cover (Kuhnell *et al.*, 1999). As such, band 4 of the MSS image, closest to band 5 of TM, was used and a data layer of the above ratio named MSS4toNDVI was produced.

The variance of pixel scores can entail information on the distribution and structure of the interested elements. Therefore, bands 2,3 and 4 of the MSS image were subjected to texture analysis and the resulting layers were also used in the subsequent modeling.

3.2.3 Landscape metrics

Patch and distance metrics were calculated as the landscape metrics. Vegetation patches exhibit different attributes in terms of landscape metrics. In order to test the effect of these metrics on the likelihood of change, the ratio of perimeter to area (*Para*), and the shape, size, and fractal dimension were calculated for each patch using the Fragstats (Version 2) software (McGarigal and Marks, 1995).

Distance measures have proven to be of relatively high importance in bringing about change in vegetation. Hence, distance from the edge of remnant patches (*Edge*), distance to the towns and villages, distance to the surface water network and to the roads were also included in the modeling. The idea was to include these layers as surrogates of socio-economic factors affecting the fate of the remnants.

Table 4. Summary of the independent variables tested / used in the modeling process

VARIABLE GROUPS	VARIABLE NAME	ABBREVIATION	USED IN THE MODELING
Bio-Physical	Elevation	Elev.	Yes
	Slope	Sl.	Yes
	Aspect	Asp.	No
	Topographic Features	Topof.	Yes
	Geology	Geo.	Yes
Image-Based	Normalised Difference Vegetation Index	NDVI	Yes
	NDVI Number of different classes	NDVINDC	Yes
	NDVI Fragmentation index	NDVIFRAG	Yes
	NDVI Dominance index	NDVIDOM	Yes
	Ratio of MSS band4 to NDVI	MSS4NDVI	Yes
	Texture of bands 2,3, and 4	Texture2, 3, 4	Yes
	Maximum of NDVI over the past 27 years	MAXNDVI	Yes
Landscape Metrics	Ratio of perimeter to area	Para	Yes
	Shape	Shape	Yes
	Size	Size	Yes
	Fractal Dimension	Fract.	Yes
	Distance to the edge of patches	Edge	Yes
	Distance of patches from each other	Dist	Yes
	Distance to roads	Roadist	No
	Distance to towns	Towndist	No
	Distance to surface drainage	Waterdist	No

4. The modeling process

The changed areas amounting to 1392 hectares (22272 pixels) were used as the dependent layer along with the above independent layers to model the change over the years 1973 to 2000. All the independent layers were normalised between 0.1 to 0.9 as this is a critical criterion especially when using neural networks methods.

4.1 Neural Networks

For the neural network modeling approach, the Stuttgart Neural Network Simulator or SNNS (Zell *et al.*, 1996) software (Version 4.2) was used. In order to diminish the auto-correlation effects on the modeling results and also decrease the size of pattern files in the neural network, thinning was carried out using the ‘contract’ module in the Idirsi software (Version 32). This was applied both on the dependent and independent variables. The thinned out dependent file showing the decrease in vegetation was used as the sampling points for dependent layer. Pixels in the file were randomly assigned to the training and test sampling sets containing 70% and 30% of the original pixels respectively. Using the sets, training and test pattern files were created through a script in Arc/Info AML format (Laffan, 1996) such that each line in the pattern files represented the scores for each pixel of the 19 predictor variables and finally a 0 or 1 for the dependent layer depending on the state of the former layer in that pixel (Figure 4).

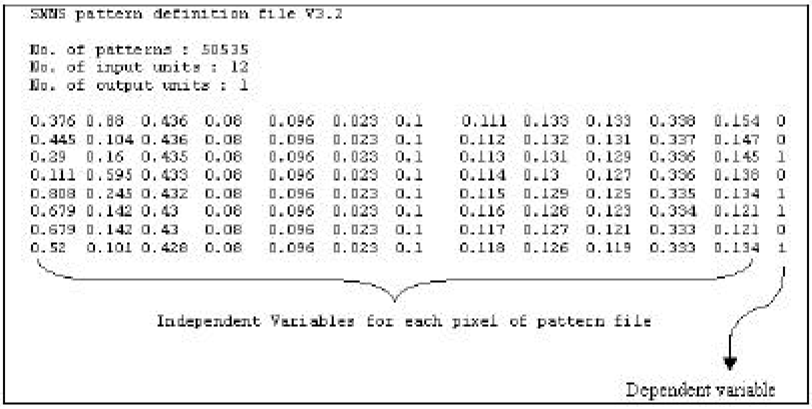

Figure 4. An excerpt from a typical pattern file

The network was designed with 1 input layer, 1 hidden layer, and 1 output layer. The input layer consisted of 19 nodes equal to the number of predictor variables and the hidden layer was assigned 45 nodes. The output layer comprised one node showing the areas prone to deforestation based on the predictor layers. A feed-forward multi-layer perceptron (MLP) with the back-propagation algorithm was used for training and prediction. Training parameters were set as 0.1 for the learning rate, 0.02 for the maximum error and 10000 cycles for training and testing. Pixel thinning carried out on the input layer using Idrisi software (Version 32) also helped avoid over-training. The SNNS accomplished the job of testing and training with 2750 cycles on the thinned out pattern files. Then the resulting network file from the SNNS was applied to the whole grid layers using another AML script in Arc/Info (Laffan, 1996). The final product which was in the form of an image containing figures between 0 to 1.0, showed in ascending order, the pixels most likely to change (from vegetated to non-vegetated). The resultant grid was then ranked in ascending order and exactly the same number of pixels detected in the post classification comparison for vegetation decrease (22272 pixels) was selected as changed areas.

4.2 Logistic Regression

Logistic regression was undertaken using the same variables within the Idrisi software. The real map of decreased areas of vegetation was used as the dependent layer with the 19 predictor variables cited above as independent layers. The vegetation cover in the year 1973 was used as the mask in performing the logistic regression. The mask defines the areas of image in which to perform the regression analysis, controlling for effects of other parts of the image in results. The result of modeling was in the form of an image containing figures in the range of 0.0 to 1.0. Selection of pixels of vegetation decrease was done in a similar manner to the SNNS method.

4.3 Comparison of the modeling methods

Results of the modeling through neural nets and logistic regression were compared using direct visual evaluation and the ROC and MGF methods. In order to compare performance of the two methods through the MGF validation procedure, an AML was written in Arc/Info (Version 8.2) and the final boolean images produced by each method were subjected to a comparison with the real decrease map. The AML script was written in a way that it allowed comparison of two boolean images containing only zero and one with the vegetation coverage as a mask. In this case, the window size was started from 1 x 1 and increased steadily in 3 cell intervals up to a window size of 30 x 30.

5. Assessing contribution of variables to modeling success

The selected independent variables are obviously of different importance in bringing about change, hence affecting the modeling success. In an effort to assess the effects these variables have on the final modeling product, namely the modeled change areas, variables were dropped one at a time from the model and the result was evaluated through the ROC and MGF methods. To do so, 20 different pattern files were created using the full and reduced-variable data set. Then the patterns were used one at a time for testing and training in SNNS with the same learning parameters as cited above. The resultant network files were then used to create 20 different simulated images of vegetation change areas. Ranking and selecting the same number of pixels was done on each image to create the boolean change image. The ROC and MGF validations were implemented using simulated and boolean images. The same procedure was followed for the logistic regression method.

6. Results and Discussion

Classification of the images was successfully carried out using the training samples acquired through literature review and the field visit (Table 5).

Table 5. Accuracy estimation of the classification for the MSS and TM images

Image : MSS73

A. Individual category accuracies (%)

ID	Producer	User	Overall	Kappa Index
Non-woody	98.84	100	99.42	91.80
Woody	100	98.85		

B. Error Matrix

Classified	Reference Non-Woody	Woody
Non-Woody	98845	0
Woody	1155	100,000

Image : TM2000

A. Individual category accuracies (%)

ID	Producer	User	Overall	Kappa Index
Non-woody	99.21	100	99.60	94.39
Woody	100	99.21		

B. Error Matrix

Classified	Reference Non-Woody	Woody
Non-Woody	99210	0
Woody	790	100,000

Post-classification comparison revealed the vegetated areas vanished from the landscape over the cited time period (Figure 5). It also provided the necessary real decrease map used as the independent layer in modeling approaches. The result of the post-classification comparison showed that over the past 27 years, around 1392 hectares of woodland and forest had been removed from the area. This amounts to around 0.63% of the total catchment area. Given that the vegetation cover within the catchment amounts to around 14375 hectares, the decrease detected is 9.6% of the vegetation cover over the past 27 years.

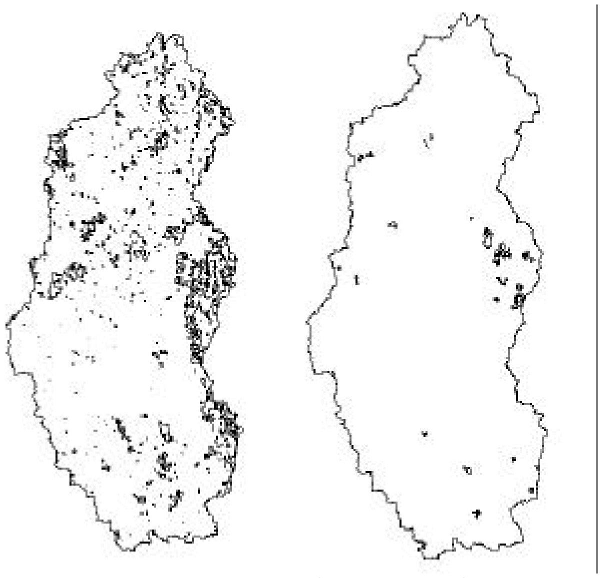

Figure 5. Vegetation cover in the catchment of Boorowa (left) and the decreased area of vegetation (right)

Initial modeling through neural networks and logistic regression methods revealed that the aspect and surface drainage network variables had very little impact on the modeling, so they were excluded from the data sets. Also, preliminary modeling results suggested that the effects of distance-to-towns and distance-to-roads were minimal, so they were excluded from the models. Running neural networks and logistic regression with the given parameters resulted in simulated images with real numbers in the range from 0 to 1 (Figure 6).

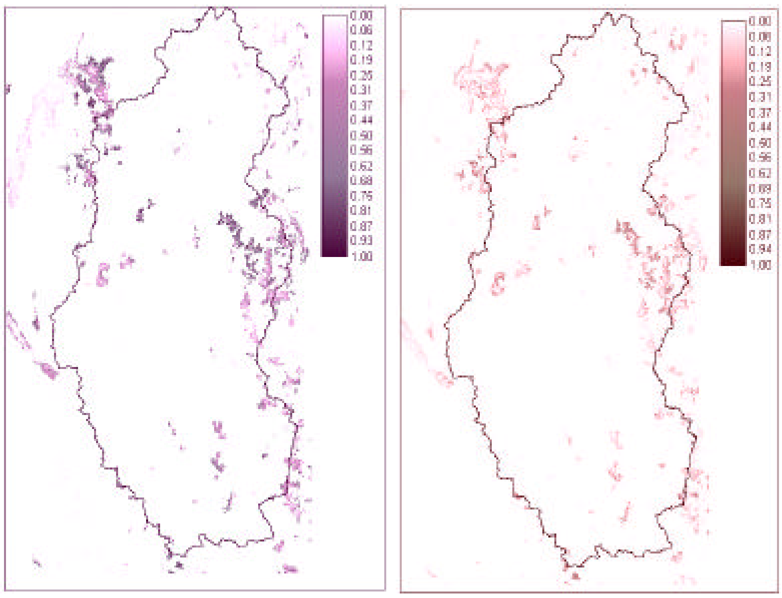

Figure 6. Resultant images from application of neural networks (left) and logistic regression methods

As mentioned before, the simulated images were ranked and exactly the same number of the detected change pixels in the real map was selected from the highest scores of the ranked images. This resulted in boolean images showing vegetation decrease areas as 1 and other areas as zero. Comparing results of the neural network and logistic regression visually with each other and the real vegetation decrease map revealed that the neural networks performs better than logistic regression (Figure 7).

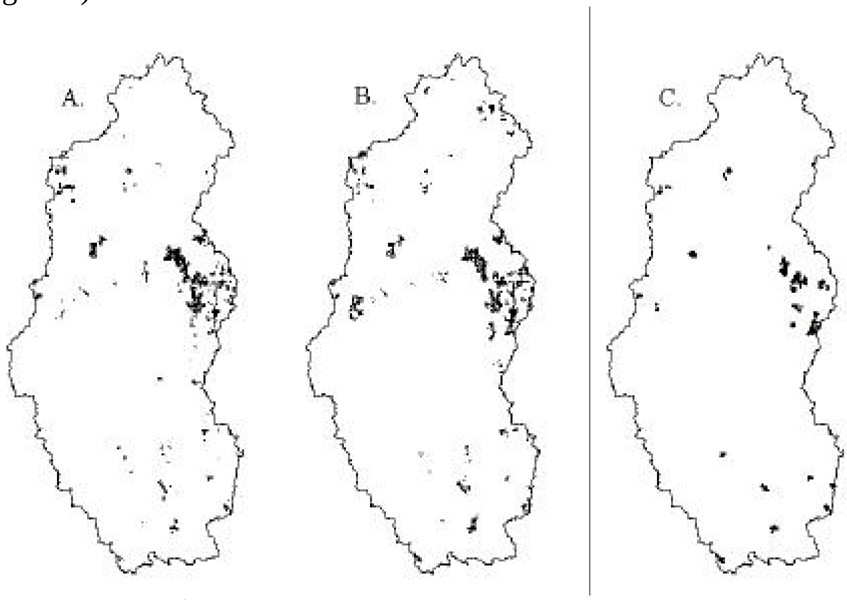

Figure 7. Visual comparison of modeling results through neural networkss (A), logistic regression (B), and the reality map (C)

Adherence to the reality map, compaction of the change pixels, and especially visual accuracy in the north and west of the area are readily apparent in the neural networks' simulated map as compared to the logit map.

The reality map and the simulated ones were subjected to the ROC statistic estimation using 10 equal-interval thresholds. Results depicted in Figure 8 below shows that the ROC estimate for the neural networks method is slightly higher than that of logit one. Also, the ROC curve showed a more rapid increase in neural networks in the first few scenarios as compared to the logit result. However, fluctuations were seen in the neural networks curve when reaching the plateau, as opposed to the logit curve showing a relatively smoother trend towards the end.

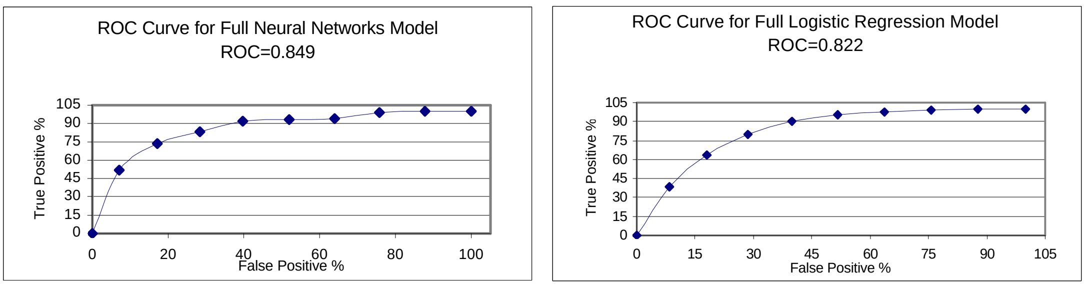

Figure 8. Comparison of the ROC curve and ROC statistic for the two modeling methods

When comparing results of the two methods with a totally random map in a combined plot, it becomes readily apparent that the methods have performed acceptably well in increasing the simulation success (Figure 9).

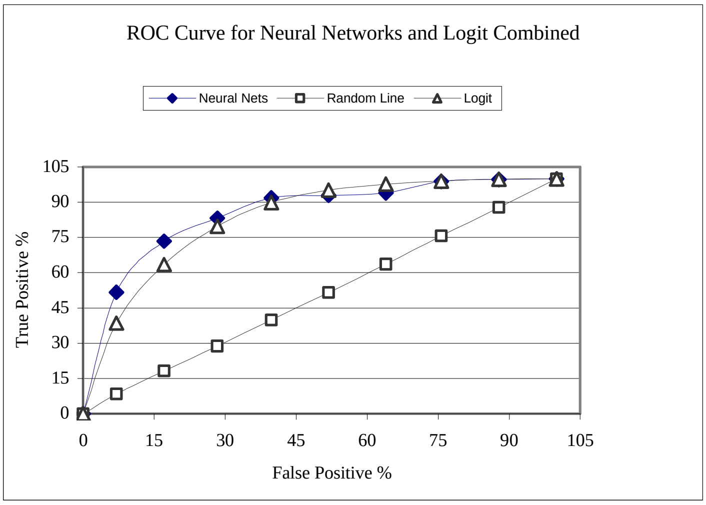

Figure 9. Comparison of the ROC curve for the two modeling methods with that of a random image

Although the increasing trend of goodness of fit with increasing window size calculated for the neural networks result is sharper than that of the logit, the difference is in fact slight. The maximum goodness of fit across all window sizes with a constant k=0.1 was found to be 0.76 for neural networks and for the logit 0.71. This result suggests that with the MGF validation method, the results of the two modeling methods are nearly the same across all window sizes (Figure 10).

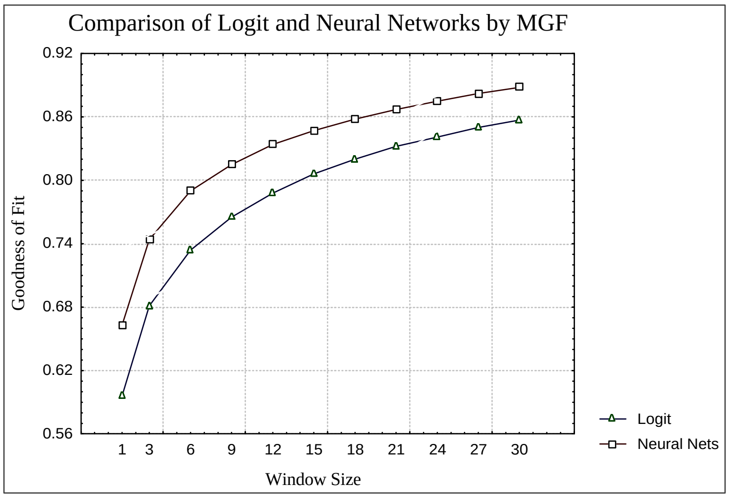

Figure 10. Comparison of the MGF curve for the two modeling methods

Examination of Figure 11 reveals the relative effect of predictor variables in the neural networks method through the MGF validation. It is apparent that the surrogate layer for agricultural density (MAXNDVI) and ratio of MSS band 4 to NDVI (MSS4toNDVI) are playing highly important roles in modeling success. Also, slope, NDVI, texture4 and texture3 are important variables since their exclusion lowers the simulation accuracy. It is additionally clear that exclusion of the layer of number of different densities of NDVI (NDVINDC) actually does benefit the modeling success.

The model without NDVINDC was considered the best one and results of other models were subtracted from it and ranked in ascending order. The result shown in figure 11, depicts the importance of the variables and also the fact that the full model effect is inferior to the other reduced-variable models in upgrading modeling success.

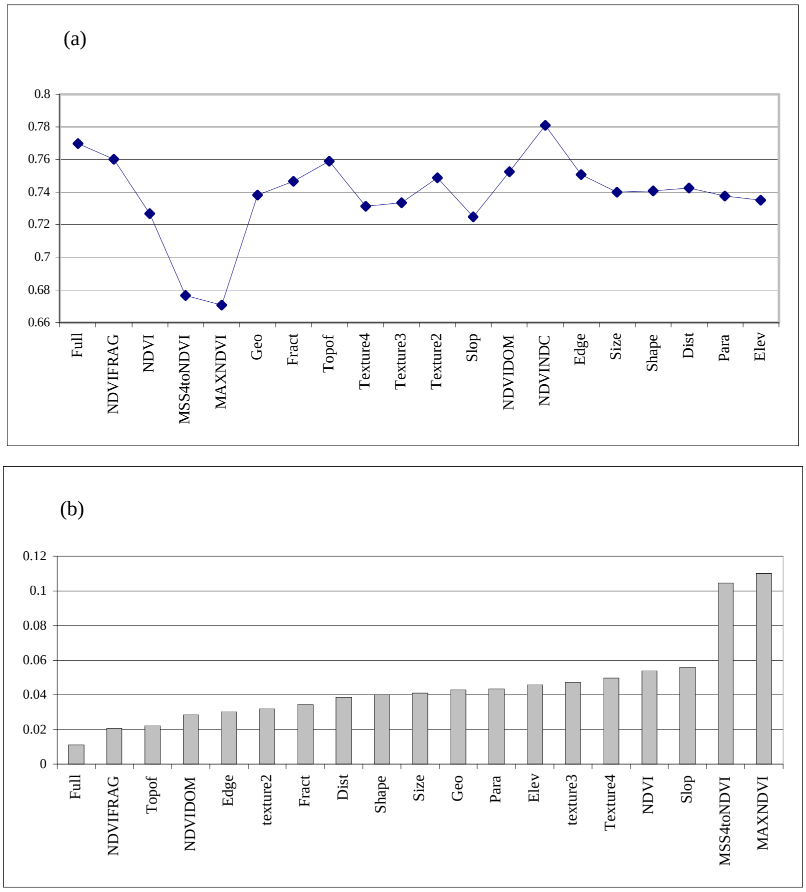

Figure 11. Effects of predictor variables on neural networks results through MGF validation (a) and the ranked order of importance of variables (b)

Shown in Figure 12 are the relative effects of 19 predictor variables on the logistic regression modeling tested by the MGF method. The figure demonstrates the importance of the perimeter to area ratio of the patches (Para) together with slope, MAXNDVI, MSS4toNDVI and the NDVI layers in the modeling success. It is also apparent that exclusion of the fractal dimension and shape layers does benefit the modeling accuracy. Figure 12 also shows the ranked order of the predictor variables for the model with the fractal dimension excluded. This demonstrates the importance of other parameters such as topographic features, texture bands, NDVI dominance index and distance of the remnant patches to each other.

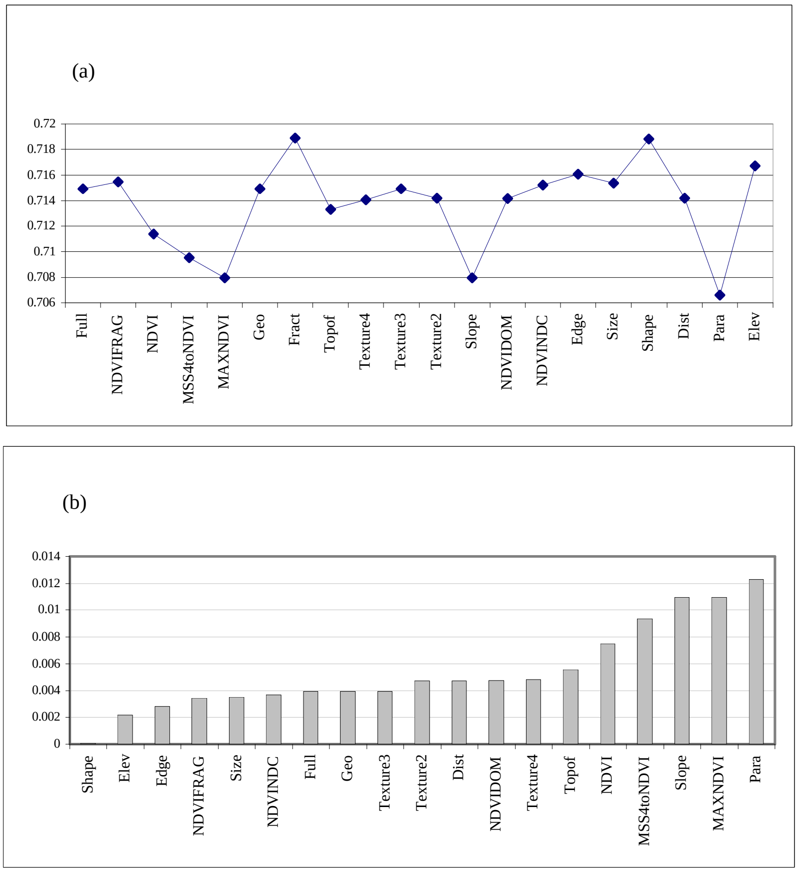

Figure 12. Effects of predictor variables on logistic regression results through MGF validation (a) and the ranked order of importance of variables (b)

When compared, it becomes evident that the neural networks method is relatively superior to logistic regression (Figure 13). Also, the neural networks method appears to be more sensitive to the inclusion or exclusion of the predictor variables. This fact is further confirmed when comparing the methods with the ROC and MGF methods in one graph. Figure 13 clearly shows that the neural networks method is generally superior to logistic regression and also more sensitive to the inclusion and exclusion of the variables.

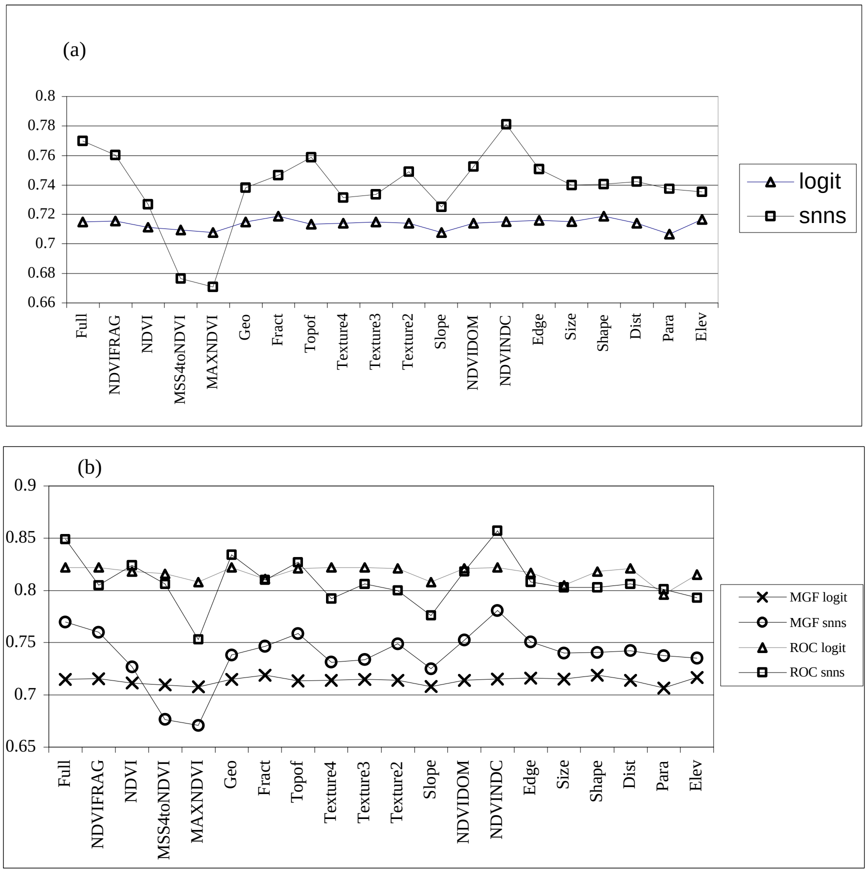

Figure 13. Comparison of the two modeling methods by MGF in a combined graph (a). Also, ROC and MGF have been compared for the two modeling methods (b)

The possible effects of the number of variables on the models' results were also examined in the neural networks method through creating a nine-variable pattern file. After training and testing, the resultant network file was applied to the selected variables and the final image was ranked and used for the boolean image creation. The latter image showing pixels most likely to change state from vegetated to non-vegetated was compared visually and through the MGF methods to the full models in neural networks and logistic regression methods. The nine-variable model, although of lower quality than the full models when MGF was considered (full model = 0.76, nine-variable model = 0.73), was of higher visual quality. This was especially evident in the boolean image which showed a clumped and compacted pattern for the clusters of pixels detected as change similar to the reality map. Adding to the number of predictor variables involved in the modeling seems to add to the confusion while slightly increasing the accuracy. The full models from both methods created a somewhat dispersed pattern for the pixels of change (Figure 14).

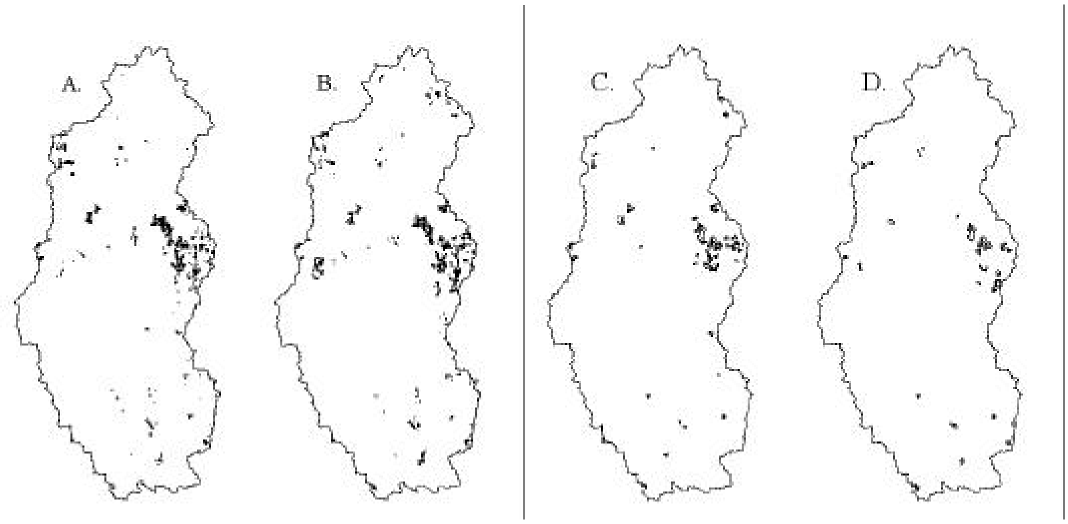

Figure 14. Visual examination of the full neural network (A), and logistic regression (B), with that of nine-variable neural network model (C), and the reality map (D)

7. Conclusion

Around 10% of the vegetation cover in the study area has vanished over the past 27 years. This was successfully shown through image classification and post-classification comparison. A review of the relevant literature and available information on the area of study resulted in defining 19 predictor variables used in the modeling.

The real change map was used as the dependent variable in the modeling process. Modeling the decrease in vegetation over the cited time period through neural networks and logistic regression proved successful. This was confirmed by using the MGF and ROC validation tests. The MGF and ROC validation methods were applied to the results of models where it was shown that the models were able to capture most of the deriving forces of change through the predictor variables.

The neural networks method was found to perform slightly better than the logistic regression method. It was also shown that the neural networks method is more sensitive to inclusion or exclusion of predictor variables as this resulted in fluctuations in the MGF or ROC curves. Increasing the number of predictor variables involved in the neural networks modeling caused confusion in the model, as a more dispersed pattern of vegetation change was detected. This was true for visual assessment of the result but not for the MGF and ROC tests.

Of the 19 predictor variables involved in the modeling, the maximum NDVI over the past 27 years (MAXNDVI), ratio of MSS band 4 to NDVI (MSS4toNDVI), and slope were found commonly of high importance in both modeling approaches. Given that the MAXNDVI reflects agricultural activity in and around remnant vegetation patches, it can be said that the fate of the vegetation in the area is dependent to a large extent on the agricultural activities and their associated intensities. MSS4toNDVI has been shown to be related to the foliage projective cover, hence the density. So, vegetation density is also very important in making the remnant patches of vegetation vulnerable to change. Slope, a variable related to ecological quality of the site and its suitability for development is also shown to be of importance in bringing about change.

Assuming the availability of both methods of analysis, either can produce equally acceptable results, but the neural networks method might be preferred where there is suspected non-linear relationships between variables. Also, in situations where the cause and effects relationship between variables and the phenomenon under study is vague, neural networks offer a promising tool. However, it seems that the greatest benefit to the modeling procedure can be obtained by careful selection of the subset of variables chosen for the final model.

8. References

Austin, M. P., Cawsey, E. M., Baker, B. L., Yialeloglou, M. M., Grice, D. J. and Briggs, S. V. 2000. Predicted vegetation cover in the central Lachlan Region. Final report of the Natural Heritage Trust Project AA 1368.97 (Canberra: CSIRO Wildlife and Ecology).

Baker, W. L. 1989, A review of models of landscape change. *Landscape ecology*, 2, **2**, 111-133.

Brogaard, S., and Prieler, S. 1998, Land cover in the Horqin grasslands, north China: detecting changes between 1975 and 1990 by means of remote sensing. http://www.iiasa.ac.at/Publications/Documents/IR-98-044.pdf

Caccetta, P. A., Campbell, N. A., Evans, F., Furby, S. L., Kiiveri, H. T., and Wallace, J. F. 1998, Mapping and monitoring land use and condition change in the south-west of western Australia using remote sensing and other data (Canberra: CSIRO Mathematical and information sciences). URL: //www.cmis.csiro.au/rsm/research/index.htm

Clark, W. A. V. and Hosking, P. L. 1986. *Statistical methods for geograohers* (Newyork: John Wiley & Sons).

Costanza, R. 1989, Model goodness of fit: A multiple resolution procedure, *Ecological modeling*, **47**, 199-215.

Cox, S. V. and Goldney, D. C. 1997, Boorowa Shire koala survey. (Bathurst: Environmental Studies Uint, Charles Sturt University).

Eastman, J. R., 2002. Idrisi for windows user's guide ver.32. Clark labs for cartography technology and geographic analysis (Worcester: Clark University).

Foody, G. M. 2001, Monitoring the magnitude of land-cover change around the southern limits of the Sahara,. *Photogrammetric Engineering and Remote Sensing*, **67,** 841-847.

Hall, F. G., Botkin, D. B., Strebel, D. E., Woods, K. D. and Goetz, S. T. ,1991, Larger-scale patterns of forest succession as determined by remote sensing, *Ecology*, 72, **2**, 628-640.

Hird, C. 1991. Soil landscapes of the Goulbourn 1:250 000 sheet (Sydney: Soil conservation service of NSW).

Hsu, L. T. and Cheng, C. C. 2000, Modeling landscape changes using logit models. URL: //www.gisdevelopment.net/aars/acrs/2000/ts6/fore0005.shtml.Last visited: 1 April 2003.

Jenerette, G. D. and Wu, J. 2001, Analysis and simulation of land-use change in the central Arizona-Pheonix regio, USA, *Landscape ecology*, **16**, 611-626.

Kuhnell, C. A., Goulevitch, B. M., Danaher, T. J. and Harris, D. P. 1999. Mapping woody vegetation cover over the state of Queensland using landsat TM imagery (Qeeunsalnd: Department of natural resources). URL: //www.dnr.qld.gov.au/slats/

Laffan, S. 1996. Arc/Info AML scripts for SNNS simulation. Personal communication.

Lambin, E. F. 1997, Modeling and monitoring land-cover change processes in tropical regions, *Progress in physical geography*, 21, **3**, 375-393.

Lek, S. and Guégan, J. F. 2000. *Artificial neuronal networks. Application to ecology and evolution* (New York: Springer-Verlag).

Mann, S., Benwell, G. L., 1996, The integration of ecological, neural and spatial modelling for monitoring and prediction of semi-arid landscapes, *Computers & Geosciences*. **22**, 1003-1012.

McCormick, N. 1999. Irish forest inventory and planning system project. (Italy: Space Applications Institute).

McGrarigal, K. and Marks, B. J. 1995. Fragstats: Spatial pattern analysis program for quatifying landscape structure. General technical report. PNW-GTR-351 (Portland, OR: U.S. Department of Agriculture, Forest Service Pacific Northwest Research Station).

Newham, L. T. H. 1999. The use of remote sensing and geographic information systems in landcare resource assessment and planning. M.Sc. thesis. The Australian National University, Canberra.

NSW National Parks and Wildlife Service (2002). The native vegetation of Boorowa Shire (Hurstville: NSW National Parks & Wildlife Service).

Pijanowski, B.C., Brown, D. G., Shellito, B. A. and Manik, G. A. 2001, Using neural networks and GIS to forecast land use changes: a Land Transformation Model, *Computers, environment and urban systems*, 26, **6**, 553-575.

Pontius, Jr., R. G. and Schneider, L. C. 2001, Landcover change model validation by a ROC method for the Ipswitch watershed, Massachusetts, USA, *Agriculture ecosystems & environment*, **85**, 239-248.

Schultz, A., Wieland, R., Lutze, G. 2000, Neural networks in agroecological modelling-stylish application or helpful tool? *Computers and Electronics in Agriculture.* **29**, 73-97.

Singh, A. 1989, Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing*, **10,** 989-1003.

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., and Macomber, S. A. 2001, Classification and change detection using landsat TM data: when and how to correct atmospheric changes? *Remote Sensing of Environment,* **75,** 230-244.

StatSoft, Inc. 2002. *Electronic Statistics Textbook* (Tulsa, OK: StatSoft). URL: http://www.statsoft.com/textbook/stathome.html.

Weng, Q. 2002, Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modeling, *Journal of Environmental Management*, **64**, 273-284.

Wolfram, S. 1986. *Theory and applications of cellular automata*, (Singapore: World Scientific)

Yates, C. J. and Hobbs, R. J. 1997. Temperate eucalypt woodlands: a review of their status, processes threatening their persistence and techniques for restoration, *Australian Journal of Botany*, **45**, 949-973.

Zell, A., Mamier, G., Vogt, M., Mache, N., Hübner, R., Döring, S., Herrmann, K., Soyez, T., Schmalzl, M., Sommer, T., Hatzigeorgiou, A., Posselt, D., Schreiner, T., Kett, B., & Clemente, G. 1996. *Stuttgart neural network simulator user manual* v4.2. [ftp.informatik.uni-stuttgart.de in /pub/SNNS/SNNSv4.1.tar.Z] (Stuttgart: Institute for Parallel and Distributed High Performance Systems, University of Stuttgart).