

Modelling deforestation using GIS and artificial neural networks

J.F. Mas ^{a,*}, H. Puig ^b, J.L. Palacio ^a, A. Sosa-López ^c

^a Instituto de Geografía, Universidad Nacional Autónoma de México (UNAM), Sede Morelia, Aquiles Serdán 382, Colonia Centro, 58000 Morelia, Michoacán, Mexico

^b Laboratoire d'Ecologie Terrestre, UMR 5552 (CNRS/UPS) 13, Av. du Colonel Roche, B.P. 4072, 31029 Toulouse Cédex 4, France

^c Centro de Ecología, Pesquería y Oceanografía del Golfo de México (EPOMEX), Universidad Autónoma de Campeche A.P. 520, 24030 Campeche, Camp, Mexico

Received 29 April 2002; received in revised form 26 February 2003; accepted 10 July 2003

Abstract

This study aims to predict the spatial distribution of tropical deforestation. Landsat images dated 1974, 1986 and 1991 were classified in order to generate digital deforestation maps which locate deforestation and forest persistence areas. The deforestation maps were overlaid with various spatial variables such as the proximity to roads and to settlements, forest fragmentation, elevation, slope and soil type to determine the relationship between deforestation and these explanatory variables. A multi-layer perceptron was trained in order to estimate the propensity to deforestation as a function of the explanatory variables and was used to develop deforestation risk assessment maps. The comparison of risk assessment map and actual deforestation indicates that the model was able to classify correctly 69% of the grid cells, for two categories: forest persistence versus deforestation. Artificial neural networks approach was found to have a great potential to predict land cover changes because it permits to develop complex, non-linear models.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Deforestation; Land use/land cover change; Spatial modelling; Artificial neural networks; Geographic information system

1. Introduction

Tropical deforestation, as an important factor in global change, is a topic that has received considerable attention recently. It has been shown to have a negative influence on regional hydrology, large-scale and long-term climate systems, global biogeochemical cycles and biodiversity lost (Fontan, 1994; IGBP, 1993; IPCC, 1996; Puig, 2000). Despite its importance, accurate statistics on deforestation rates are not available in most countries (Grainger, 1993). In Mexico, various authors reported deforestation rates that range between 0.3% and 5% by year (Cortina Villar et al., 1999; Díaz Gallegos et al., 2001; Bocco et al., 2001; Turner et al., 2001; Velázquez et al., in press). A recent study (Mas et al., 2002; Velázquez et al., 2002) estimates the nation-wide rate of deforestation in about 0.3% and 0.8% by year for temperate and tropical forests, respectively, which represents

a total loss of 84,000 km² of forest cover between 1976 and 2000. The main agents driving deforestation are well known, even though it is difficult to assess their relative contribution to deforestation and there is not a clear understanding on how these factors interact. The simulation of land use/cover changes is important for a variety of management and planning issues as well as for academic research. In the case of deforestation, the development of models is motivated by several potential benefits: (1) to provide a better understanding on how driving factors govern deforestation, (2) to generate future scenarios of deforestation rates, (3) to predict the location of forest clearing and, (4) to support the design of policy responses to deforestation (Lambin, 1994).

This study aims at developing a simple spatial model that is able to predict the location of deforestation using an artificial neural networks (ANNs) approach.

* Corresponding author.

E-mail address: jfmash@igiris.igeograf.unam.mx (J.F. Mas).

2. Background

2.1. Land use change models

The projection of land use changes can be performed using two main categories of models (Lambin, 1997; Stéphenne and Lambin, 2001): (1) empirical models based on an extrapolation of the patterns of change observed over the recent past, with a limited representation of the driving forces of these changes, and (2) simulation models based on the thorough understanding of the processes of change. The spatial prediction of land use/cover changes can be obtained by models which belong to the first category. Several studies have achieved a good projection of likely patterns of land use/cover change, based on multivariate models representing the interactions between environmental variables that are controlling the changes. Such spatial models attempt to identify explicitly the proximate causes of land use/cover change using statistical approaches such as regression or weight of evidence (Soares-Filho et al., 2002; Schneider and Pontius, 2001; Almeida et al., 2003). They are based upon the assumption that the relationships between the changes and the proximate causes of these changes remain the same over time and, therefore, can only provide short-range projections (5–10 years at most) due to the dynamic character of land use/cover change processes.

Many challenges must be faced when attempting to develop models of land use/cover change processes. The site and magnitude of land-use changes are the results of human decisions. The main difficulty is the high degree of complication, the large number of conditions, and the complexity of the interactions between human and environmental factors (Lambin, 1994). For example, Sader and Joyce (1988) examined forest area change associated with factors such as slope and transportation networks for Costa Rica. He found a strong relationship between forest clearing and proximity to road. He observed also that forest clearing on shallow slope before 1961 was low because the 0–5% slopes of the Atlantic low land were inaccessible. The increased clearing on shallow slopes occurred in an area previously less accessible and recently opened to development by the construction of a highway. This very simple example shows that forest clearing is not the result of the sum of the effects of each factor in an independent form but rather the combination of them. Taking into account environmental, socio-economic and cultural variables aspects of deforestation processes, relationships between the variables may be very complex. Statistical method, such as logistic regression, may have some limitations when variables interact on a complex way. They are invalid when spatial variables correlate with each other and have difficulties in handling poor and noisy data (Li and Yeh, 2002). Because ANNs are able to directly take

into account any non-linear complex relationship between the explicative variables and deforestation, better results can be expected from this approach. A few land use/cover change model based upon ANNs and focused on simulating urban growth have been recently reported in the literature (Li and Yeh, 2002; Pijanowski et al., 2002). However there is no model using ANNs aimed at predicting tropical deforestation.

2.2. Artificial neural networks

ANNs are non-linear mapping structures based on the function of the human brain. Advantages of the ANN approach include ability to handle non-linear functions, to perform model-free function estimation, to learn from data relationships that are not otherwise known and, to generalize to unseen situations. ANNs have been shown to be universal and highly flexible function approximators for any data. Therefore, ANNs make powerful tools for models, especially when the underlying data relationships are unknown (Lek et al., 1996; Lek and Guégan, 1999). In the last decade, ANNs have seen an explosion of interest and have been successfully applied across a large range of domains such as image recognition, medicine, molecular biology and, more recently, ecological and environmental sciences (Lek and Guégan, 1999; Maier and Dandy, 2000). Some recent utilizations of ANNs in environmental sciences are models predicting species distribution, abundance or diversity as a function of environmental variables (Lek-Ang et al., 1999; Manel et al., 1999), rice crop damage by flamingos (Tourenq et al., 1999), streamflow and flash floods (Kim and Barros, 2001), air quality parameters (Abdul-Wahab and Al-Alawib, 2002; Chelani et al., 2002) and ecosystem characteristics from remotely sensed data (Jensen et al., 1999; Paruelo and Tomasel, 1997).

The multi-layer feed-forward neural network or multi-layer perceptron (MLP), is one of the most popular ANN architecture in use today (Bishop, 1995; Atkinson and Tatnall, 1997; Lek and Guégan, 1999). The MLP is based on the supervised procedure, i.e. the network constructs a model based on examples of data with known outputs. The training is done solely from the examples presented, which are together assumed to implicitly contain the information necessary to establish the relation. A MLP is a powerful system, often capable of modelling complex relationships between variables. It allows prediction of an output object for a given input object or a set of input objects.

The architecture of the MLP is a layered feed-forward neural network, in which the non-linear elements (neurons) are arranged in successive layers, and the information flows unidirectionally, from the input layer to the output layer, through the hidden layer(s): when the network is executed, the input variable values are placed in the input units, and then the hidden and output

layer units are progressively executed. Each of them calculates its activation value by taking the weighted sum of the outputs of the units in the preceding layer. The activation value is passed through the activation function to produce the outputs of the neuron. When the entire network has been executed, the outputs of the output layer act as the output of the entire network (Fig. 1).

The learning procedure is based on a relatively simple concept: if the network gives the wrong answer, then the weights are corrected so the error is lessened so future responses of the network are more likely to be correct. The best-known example of a neural network-training algorithm is backpropagation. Modern second-order algorithms such as Levenberg–Marquadt are faster. Levenberg–Marquardt is an advanced non-linear optimization algorithm. However, its use is restricted to small, single output networks using sum-squared error function (Bishop, 1995).

As the network is trained to minimize the error on the training set, a major issue is over learning or over fitting. A network with more weights models a more complex function, and is therefore prone to over-fitting (Bishop, 1995; Foody and Arora, 1997). In order to avoid over learning cross-validation is used: some of the training cases (verification set) are not actually used for training but to keep an independent check on the progress of training. As training progresses, the training error naturally drops. If the verification error stops dropping, or starts to rise, this indicates that the network is starting to over fit the data, and training should cease.

In a classification problem, an output unit's task is to output a strong signal if a case belongs to the class, and a weak signal if it does not. The activation value may be also considered as a fuzzy membership (Civco, 1993; Foody, 1995), which can be perceived as a measure of certainty to belonging at the class. When the ANNs are used to model deforestation as a function of explanatory variables, these activation values express the propensity (activation values cannot be properly considered as probabilities) of deforestation.

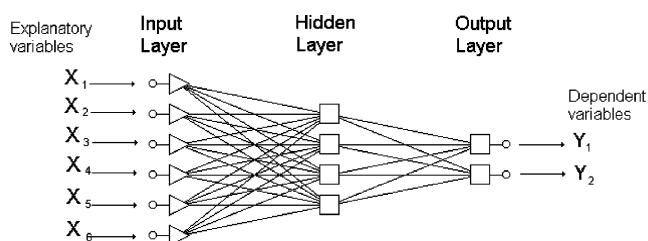


Fig. 1. Schematic illustration of a three-layered perceptron, with one input layer, one hidden layer and one output layer. In this example X_1, X_2, \dots, X_6 are six input (explanatory) variables (e.g. factors of deforestation); Y_1 and Y_2 are two output variables (e.g. deforestation and regrowth).

3. Study area

The study area is situated in the State of Campeche, south east of Mexico between $18^{\circ} 00'$ and $18^{\circ} 55'$ North and $90^{\circ} 55'$ and $92^{\circ} 06'$ West (Fig. 2) and covers about $12,400 \text{ km}^2$. Soils are dominated by gleysol; there are also solonchak and rendzina soils. The study area consists of extensive areas of savannah and a mosaic of mangroves, pastures and remnant tropical forests. A large proportion of the land surrounding the Lagoon of Términos has been deforested. Major causes of deforestation are government-directed colonization schemes, cattle ranching and, during the 1980s, rice farming (Isaac-Márquez, 1993; Mas and Puig, 2001). The main town in the study area is Ciudad del Carmen (84,000 inhabitants). There are also several towns and hundreds of villages in the study area and its surrounding. In 1994, an important part of the area has been declared as nature reserve.

4. Material and methods

Two Landsat MSS images (path 21, row 47) dated 15/2/1974 and 15/1/1986 and a Landsat TM image dated 3/4/1991 were obtained. The Landsat data were geometrically corrected and registered to a common UTM projection with an RMS error of less than 1.0 pixel. Digital maps of soils, road networks and human settlements were produced by manual digitisation of the National Institute of Geography, Statistics and Informatics (INEGI) 1:250,000-scale maps and the satellite images data. A digital elevation model generated from 10-m contour lines was used to create a slope map.

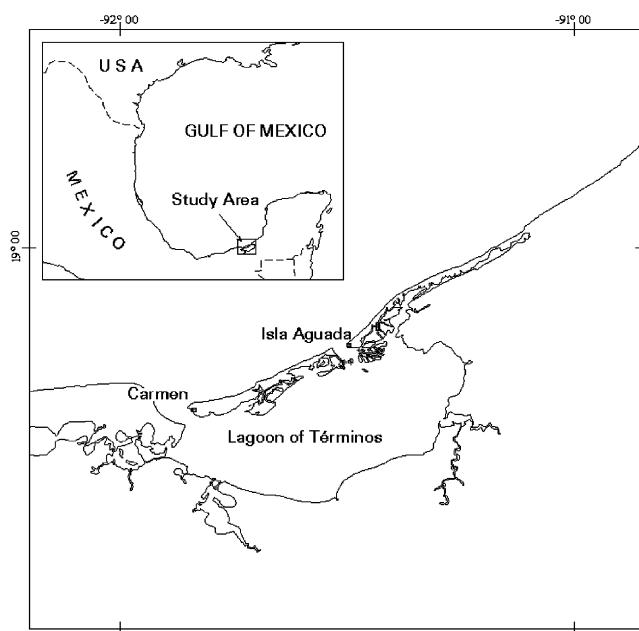


Fig. 2. Localization of the study area.

As shown in Fig. 3, the model discussed in this paper follows five sequential steps: (1) elaboration of maps of deforestation obtained by overlaying maps of forest cover from more than one point in time, (2) quantification of the relationships between deforestation and proximate causes and selection of the “best” explanatory variables, (3) calibration of the model (training of the ANN), (4) simulation (elaboration of a map of propensity to deforestation which predicts deforestation for the following period) and (5) assessment of the model performance (comparison between actual and predicted deforestation).

4.1. Deforestation monitoring

In a previous study several techniques of change detection, such as image differencing, selective principal components analysis, direct multiday unsupervised classification and post-classification image comparison, have been tested in order to determinate the optimal procedure to assess deforestation in this area. Post-classification comparison was found to be the more accurate technique because it was less sensitive to radiometric variations due to the difference in soil moisture and vegetation phenology between scenes from different dates (Mas, 1999). Therefore, images were classified independently using the maximum likelihood algorithm into the following land cover classes: “undisturbed” tropical forest, secondary tropical forest, mangroves and non-forest vegetation. Classification accuracy was evaluated by calculating global accuracy and Kappa coefficient using an

independent sample of 106 reference points obtained from 1:50,000-scale aerial photographs interpretation. Areas of forest were calculated for the three dates and annual rates of forest clearing were estimated. Classified images were simplified by grouping tropical forest (both undisturbed and secondary) into a single “forest” class and all other land cover classes in another “no-forest” class. Thus, the final images, after completing this processing, were binary-coded. In order to reduce file size, pixels were aggregated using cells of 120×120 m. A majority filter that replaced cells based upon the majority of their contiguous neighbouring cells was applied.

As a following step, images were overlaid in order to produce a digital map of deforestation that represents changes in forest cover. Some change combinations such as “no-forest/no-forest” or “no-forest/forest” (regrowth) were coded as “no data” because the study focus on modelling the deforestation process only. Therefore, the deforestation maps present only two classes: forest persistence (forest in both dates) and deforestation (e.g. forest in 1974 deforested in 1986) coded 0 and 1, respectively.

4.2. Integration of potential proximate causes of deforestation in the GIS database

An attempt was made to determine the relationship between deforestation and environmental and socio-economic factors considered a priori as elements that could influence deforestation such as distance from

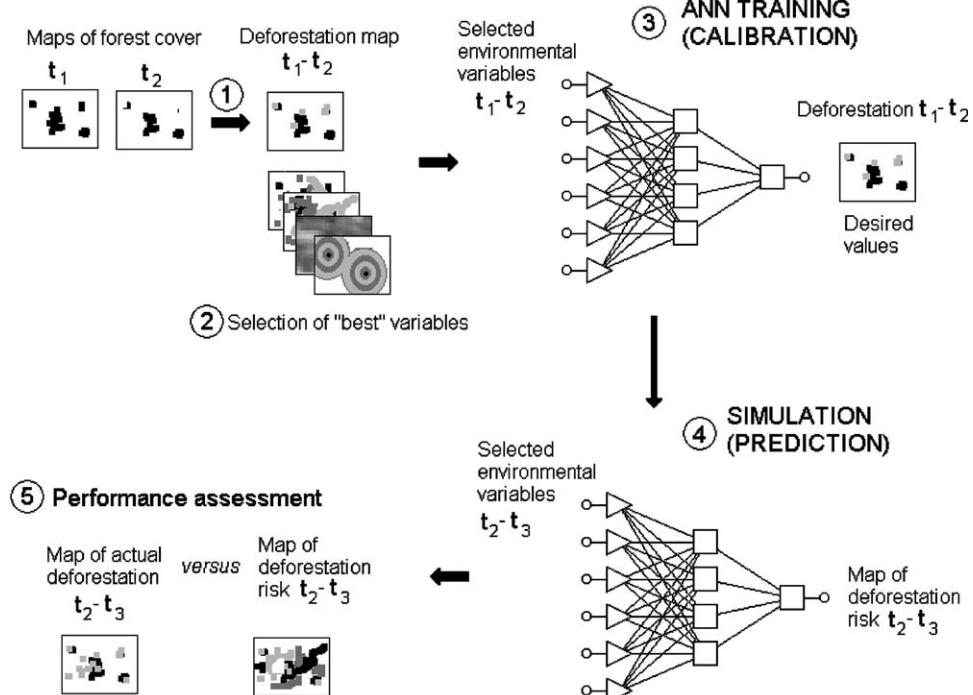


Fig. 3. Flowchart of the neural-network based model for simulating deforestation.

roads, distance from settlements, elevation, slope, soils and forest fragmentation patterns.

Several spatial explanatory variables describing potential proximate causes of deforestation were generated:

1. Elevation.
2. Slope.
3. Type of soil.
4. Shortest distance to the nearest road.
5. Shortest distance to the nearest settlement: all the settlements surrounding the study area were included in this calculation.
6. Shortest distance to the nearest forest/non-forest edge.
7. Spatial fragmentation of the forest cover in the immediate surroundings of each location. Forest fragmentation was estimated using two indices (a) the forest cover index and, (b) the Matheron index. The forest cover index, expressed in %, represents the proportion of forest pixels calculated in 3×3 , 9×9 and 15×15 pixels windows. The Matheron index, calculated in 3×3 pixels windows, is defined as (Matheron, 1970):

$$M = \frac{N_{F-NF}}{\sqrt{N_F} \cdot \sqrt{N}} \quad (1)$$

where N_{F-NF} is the number of boundaries between forest and non-forest pixels, N_F is the number of forest pixels and N is the total number of pixels. The numerator measures the number of pairs of adjacent pixels classified as forest and no-forest (i.e. the length of the perimeter line of forest pixels) and the denominator normalizes this count by the size of the forest and entire area (Mertens and Lambin, 1997).

Variables based upon distances and forest fragmentation were generated for 1974 and 1986. Each 1974 spatial variable was then overlaid on the 1974–1986 digital deforestation map in order to establish the relationship between deforestation rates and variables. The overlay operation allows constructing a tabular database which indicates, for each pixel, the value of the explanatory variables and the value of deforestation.

4.3. Selection of the input variables

The correlation coefficient of Pearson was calculated for each pair of variables in order to avoid using highly correlated variables (coefficient of correlation over 0.80) and to reduce the effects of multi-collinearity. When classifying data, class separability rises initially with an increase in the number of discriminating variables used at a point beyond which the addition of additional data has either no significant effect or results in a decrease in classification performance. This effect, sometime

referred as the Hughes phenomenon, can have a significant effect on ANNs (Shahshahani and Landgrebe, 1994; Bishop, 1995; Yang et al., 1996; Benediktsson and Sveinsson, 1997). A key step is then the determination of the optimal subset of variables to use in the ANN training. In order to reduce dimensionality, a subset of variables presenting the greatest divergence between deforestation and forest persistence classes was identified by means of the Bhattacharyya distance calculation (Mausel et al., 1990; Landgrebe and Biehl, 2000).

4.4. Multi-layer feed-forward neural network training (calibration)

The variables selected in the previous step were used to elaborate the ANN training data. Data were divided into three sections: the training set, verification set and test set using proportion of 2:1:1. Training algorithms do not use the verification or test sets to adjust network weights. The verification set is used to track the network's error performance and to stop training if overlearning occurs. The test set is not used in training at all, and is designed to give an independent assessment of the network's performance when an entire network design procedure is completed.

As the model developed in this study presents a single output (propensity to deforestation) both backpropagation and Levenberg–Marquadt training algorithms were alternatively tested in the training process. A key design decision is the question of how many hidden units to include in the network. The network configuration was determined empirically by testing various possibilities and selecting the one that provides the best compromise between bias and variance (Geman et al., 1992; Bishop, 1995).

4.5. Elaboration of a map of propensity to deforestation

Using the GIS layers, the model, trained on past data (1974–1986), was used to create a deforestation risk assessment map that indicates the areas with the highest propensity for deforestation during the following period (1986–1991). The output of the MLP is an activation value which express, for each pixel, the propensity to deforestation. The result is then a fuzzy deforestation map that portrays gradations of the likelihood of being deforested.

4.6. Model performance assessment

The map of propensity for deforestation for 1986 forests was compared with the actual deforestation during 1986–1991. Performance of the model was evaluated using two methods in order to compare it with other spatial models published elsewhere (de Brujin, 1991; Mas

et al., 1996; Pontius et al., 2001): (1) calculating a coefficient of prediction (de Brujin, 1991) and, (2) estimating the agreement between the deforestation risk assessment map and the actual deforestation map through Kappa coefficient calculation (Pontius et al., 2001).

The coefficient of prediction is based on the comparison of the activation values (propensity for deforestation) and actual deforestation and indicates how much of the actual deforestation was predicted by the model. For example, assume a study area where 10% of the forest area has been deforested, a perfect prediction would label exactly 10% of the forest land as sensitive to deforestation and only those areas actually change. The coefficient of prediction would be 1. When changes are evenly distributed over the different risk classes (no relation between prediction and actual deforestation), the coefficient of prediction would be 0. Assume a deforestation model that would classify the land in three classes of risk of deforestation: class 1 (high risk) contains 10% of the land, but 50% of the changes; class 2 (medium risk), 30% of the land and 40% of the changes; and class 3 (low risk), 60% of the land and 10% of the changes. In a plot of cumulative change percentage against cumulative forest percentage, we would obtain the line OBCM (Fig. 4). Perfect prediction would give the line OA; no relation between prediction and deforestation would give the line OM. The coefficient of prediction is defined as the area of polygon OBCM divided by the area of polygon OAM.

The second method of performance assessment is based on the comparison of a predicted deforestation map, obtained by thresholding the risk assessment map,

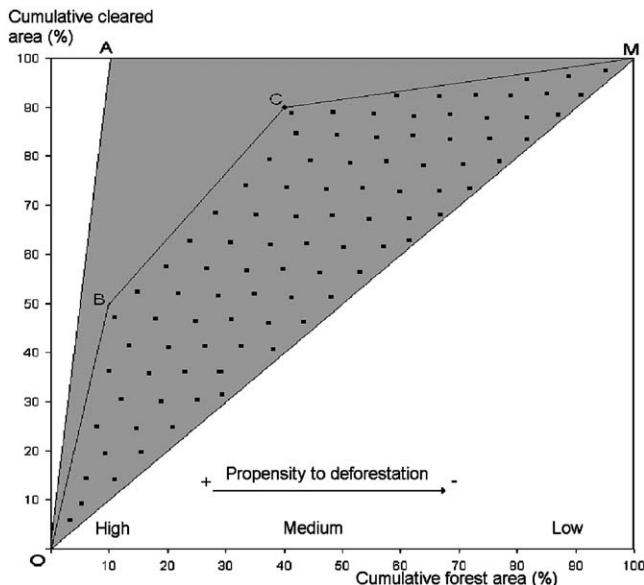


Fig. 4. Coefficient of prediction calculation. The line OA represents perfect prediction, OM no prediction. The coefficient of prediction is the area of polygon OBCM (area with black squares) divided by the area of polygon OAM (total grey shaded area).

and the actual deforestation map (Geoghegan et al., 2001). The deforestation risk assessment maps show, for each pixel, the propensity to deforestation. In the “actual” event, either a pixel was deforested or not. That is, the prediction is a propensity to change, while the actual change is a binary response (deforestation/forest persistence). Therefore, in order to compare both maps, the total number of actual pixels, n , that were deforested was first identified. Then, the pixels with highest propensity for deforestation were sequentially identified until n have been selected (i.e. the threshold value to transform the activation values of the model into binary values is set in order to obtain the “right” proportion of deforestation/forest persistence pixels). A “success” occurs when a pixel in the map of simulated deforestation matches the corresponding grid cell in the map of actual deforestation. The Kappa coefficient compares the percent success of the model to the expected percent success due to chance alone (Rosenfield and Fitzpatrick-Lins, 1986). It is important to use Kappa to evaluate the effectiveness of the modelling, because the percent due to random chance can be substantial, due in part to the fact that the model cheats by specifying the correct quantity of deforestation. Kappa administers appropriate punishment for cheating on the quantity of deforestation (Pontius et al., 2001).

5. Results

The three images were classified into undisturbed tropical forest, secondary tropical forest, mangroves and non-forest vegetation. Classification accuracy ranges between 75% and 92% (Kappa 0.6583–0.8789). Forests areas were calculated for the three dates. Results indicate that more than 42% of the tropical forest was deforested between 1974 and 1991. The deforestation rates are 2.2% and 5.3% annually for 1974–1986 and 1986–1991 periods, respectively. As a following step, all the land cover classes except tropical forest were grouped into a single class. This operation allows an improvement of accuracy by grouping classes which presented the most common spectral confusion such as primary and secondary forests. Accuracy of the binary images (forest/non-forest) ranges between 81% and 95% (Kappa 0.7076–0.9108). These images were then overlaid in order to generate the digital deforestation maps for 1974–1986 and 1986–1991 periods (Fig. 5).

The frequency of deforestation was calculated as a function of explanatory variables, successively, road proximity, settlements proximity, forest-cover fragmentation indices and proximity to a forest/non-forest edge, for both 1974–1986 and 1986–1991 periods. For both periods, deforestation rates are higher in the elevated and gently sloped areas, which are the non-flooded areas, and in more fertile soils such as solonchak. The relationship

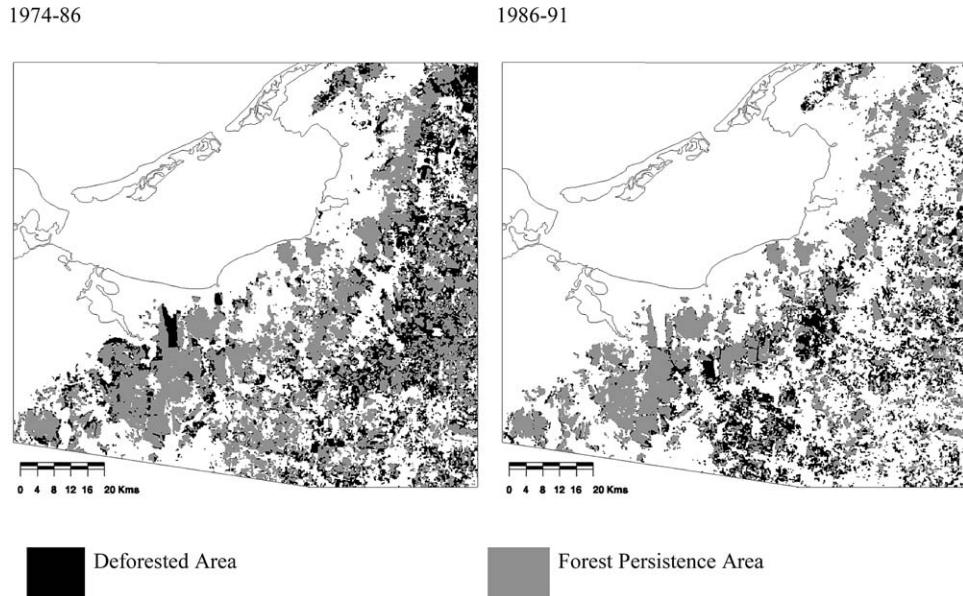


Fig. 5. Deforestation digital maps.

with the proximity to roads shows that deforestation rates decrease rapidly when moving away from roads. A similar pattern was observed with distance from settlements. The relationship with Matheron indice and local forest cover indices shows a much higher occurrence of deforestation for fragmented forest covers. This confirmed that forest openings attract forest clearing. The relationship with the proximity to forest/non-forest edges shows also a much higher occurrence of deforestation near forest/non-forest boundaries. A more detailed discussion about the relationship between the explanatory variables and the deforestation rates in this region can be found in [Mas and Puig \(2001\)](#).

The correlation coefficient of Pearson between explanatory variables was calculated. Some variables such as distances from roads and from settlements or the different indices of fragmentation are redundant variables (highly correlated). The Bhattacharyya distance was calculated for all combinations of 2, 3, 4 and so on variables. Combination based on three variables provided most of the available separability: adding more inputs did not increase significantly the separability between forest persistence and deforestation classes ([Table 1](#)).

Therefore, the best combination of three variables was used. These variables are the forest cover index calculated in a 3×3 pixels window, the elevation and, the distance from human settlements. Further analysis was done using only these three (non-correlated) variables. The data were divided randomly into three parts: the training, test and verification sets using 2000, 986 and 985 cases, respectively. Several network structures and training algorithms allowed satisfactory classification accuracy ([Table 2](#)). As accuracies were very similar, we

considered the simplest structure as the optimal network ([Fig. 6](#)). This MLP correctly classified 61% of the pixels into the two classes (deforestation and forest persistence).

The MLP trained with the 1974–1986 data was run with 1986 data as inputs. The activation values were used to generate a deforestation risk assessment map that indicates the propensity for deforestation for each forest pixel from the 1986 digital forest map ([Fig. 7](#)). Activation values and actual changes were compared in order to calculate the coefficient of prediction. The deforestation risk assessment map was thresholded in order to assess its agreement with the actual 1986–1991 deforestation digital map. The threshold value was set to 0.3138 in order to obtain the same quantity of deforestation pixels than the 1986–1991 deforestation map. The model was able to classify correctly 68.6% of the pixels. The coefficient of Kappa was 0.34 and the coefficient of prediction was 0.49 which indicates a reasonable performance of the model: as comparison, coefficients of prediction obtained by urban growth models ([de Brujin, 1991](#)) ranged between 0.25 and 0.59. Previous deforestation models based on regression techniques using a different and the same data reached a coefficient of prediction of 0.4 and 0.45, respectively ([Mas et al., 1996, 2000](#)). Kappa coefficient obtained by predictive deforestation maps derived from a model of deforestation in Costa Rica ranges from 0.31 to 0.53 ([Pontius et al., 2001](#)).

[Fig. 8](#) shows the location of the model's successes and errors. By visually inspecting this figure and by calculating characteristics of correctly and incorrectly predicted areas ([Table 3](#)), some intuition is gained on what could be potential problems and limitations of the current model. The areas incorrectly predicted as forest persist-

Table 1

Best measures of separability (Battacharyya distance) and variable combination for subset with increasing number of variables

| Number of variables | Separability value | Best variables combination |
|---------------------|--------------------|---|
| 1 | 5.99 | (DFS) |
| | 0.13 | (FCI3 × 3) |
| | 0.02 | (MATH) |
| | Less than 0.02 | All others combinations |
| 2 | 6.12 | (DFS, FCI3 × 3) |
| | 6.00 | (MATH, DFS), (SOIL, DFS), (DISTFE, DFS), (ELEV, DFS), (SLOPE, DFS) |
| | 5.99 | (FCI9 × 9, DFS), (FCI15 × 15, DFS), (DFR, DFS) |
| 3 | Less than 0.16 | All others combinations |
| | 6.14 | (DFS, FCI3 × 3, ELEV) |
| | 6.13 | (DFS, FCI3 × 3, SOIL), (DFS, FCI3 × 3, DISTFE), (DFS, FCI3 × 3, SLOPE), (DFS, FCI3 × 3, MATH) |
| 4 | Less than 6.13 | All others combinations |
| | 6.15 | (DFS, FCI3 × 3, ELEV, DISTFE) |
| | 6.17 | (DFS, FCI3 × 3, ELEV, DISTFE, SOIL) |
| 5 | 6.18 | (DFS, FCI3 × 3, ELEV, DISTFE, SOIL, MATH) |
| 6 | 6.19 | (DFS, FCI3 × 3, ELEV, DISTFE, SOIL, MATH, FCI9 × 9) |

DFS: Distance from settlements; DFR: distance from roads; FCI3 × 3: forest cover index (3 × 3 window); ELEV: elevation; DISTFE: distance from forest edge; SLOPE: slope; SOIL: soil type; MATH: Matheron index; FCI9 × 9: forest cover index (9 × 9 window); FCI15 × 15: forest cover index (15 × 15 window). The combinations using more than seven variables did not present separability improvement.

Table 2

Three best networks: structure and performance

| Number of hidden units ^a | Training algorithm | Number of epochs used for training | Classification accuracy of test data (%) |
|-------------------------------------|--------------------|------------------------------------|--|
| 2 | Levenberg–Marquadt | 99 | 61.1 |
| 4 | Backpropagation | 119 | 63.1 |
| 8 | Backpropagation | 257 | 62.4 |

^a The number of input and output units is the same (three and one) for all the networks.

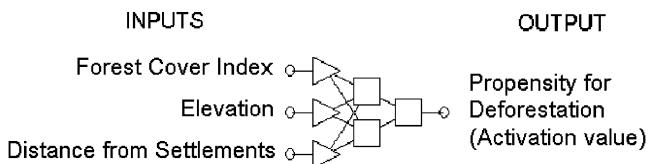


Fig. 6. MLP structure. It consists of three inputs, one hidden layer with two nodes and as output the activation value which indicates the propensity for deforestation (gradient of belonging to the deforestation class).

ence are large patches of tropical forest in low elevation and correspond mainly to pasture land and rice cultivation. The areas incorrectly predicted as deforestation are upland fragmented forest. The model over-estimated deforestation of fragmented upland forest and was not able to predict some large patches of deforestation. These large forest clearing are due mainly to government policy aimed at increasing rice production in the State of Campeche during the 1980' decade (Isaac-Márquez, 1993).

6. Discussion and conclusion

ANNs are able to directly take into account any non-linear relationship between the explanatory and dependent variables. According to the Kolmogorov's theorem, the use of $2n + 1$ hidden neurons (with n the number of input neurons) can guarantee the perfect fit of any continuous functions (Bishop, 1995). Experiments indicate that $2n + 1$ hidden neurons may be too many in applications and that $2n/3$ hidden neurons can generate almost similar results (Wang, 1994; Li and Yeh, 2002). However, this ability may become a disadvantage because the model could be over-specific to the training data. Over-specificity reduces the generalization capacities of the ANN. In this study, networks more complex, using more input variables, were able to obtain better performance when classifying the training data than the network we selected. But they were not able to predict deforestation of the following period. It seems that they were very specific to particular patterns of deforestation of the training period, lost the general

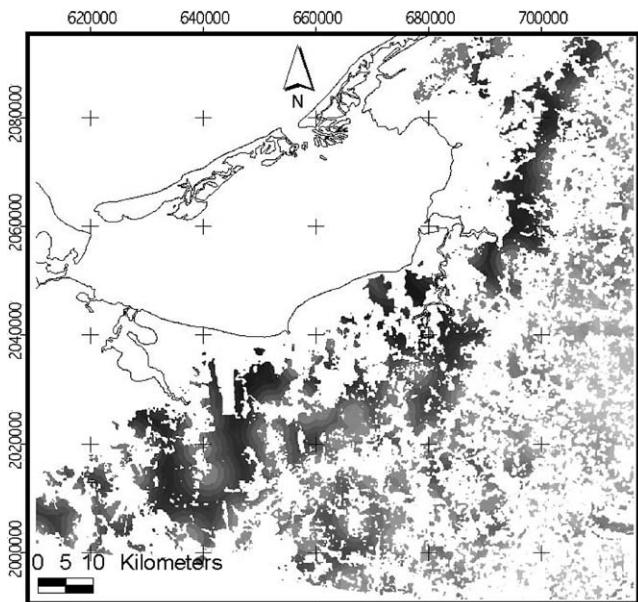


Fig. 7. Deforestation risk assessment map. Activation values which indicate the propensity to deforestation are scaled from black (low propensity to deforestation) to white (high propensity to deforestation).

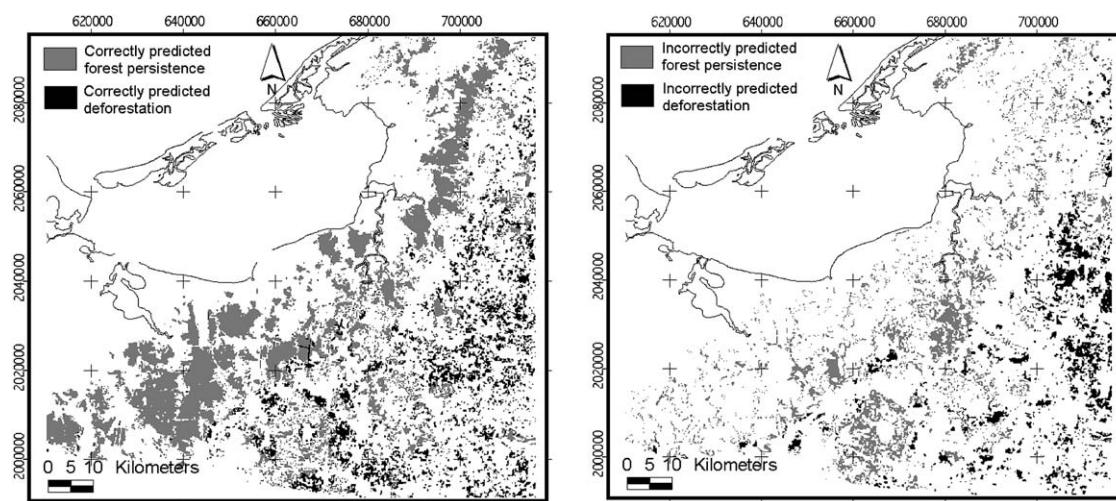


Fig. 8. Spatial distribution of correctly and incorrectly predicted deforestation/forest persistence areas.

Table 3
Characterization of predicted areas grouped by type of error/success

| Type of error/success | AV | | ELEV (m) | | DFS (km) | | FCI3 × 3 (%) | |
|--|------|------|----------|------|----------|------|--------------|------|
| | Mean | S.D. | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| Correctly predicted deforestation | 0.51 | 0.11 | 41 | 24 | 0.9 | 1.5 | 77.2 | 19.9 |
| Correctly predicted forest persistence | 0.25 | 0.09 | 10 | 9 | 2.7 | 2.1 | 96.0 | 10.2 |
| Incorrectly predicted deforestation | 0.50 | 0.09 | 46 | 25 | 1.1 | 1.7 | 85.7 | 18.2 |
| Incorrectly predicted forest persistence | 0.29 | 0.07 | 14 | 10 | 2.0 | 1.4 | 89.9 | 14.9 |

AV: Activation value; ELEV: elevation; DFS: distance from settlement; FCI3 × 3: forest cover index; S.D.: standard deviation.

trends of deforestation processes and therefore were not able to predict correctly deforestation for the following period.

The loss of generalization capacity may be overcome when the number of input variables introduced to the network is reduced (Bishop, 1995; Rosin and Fierens, 1995; Foody and Arora, 1997; Kavzoglu and Mather, 2000). Variable selection can, therefore, play a very important role in the reduction of number of inputs in order to obtain relevant and uncorrelated inputs. Variable selection may be obtained through separability measures (Kavzoglu and Mather, 2000), PCA, discriminant analysis and decision boundary feature extraction (Benediktsson and Sveinsson, 1997). The main difficulty is then to find the trade-off between a model with high performance when training but little capacity of prediction and a simpler model which presents lower performance with training period but a better prediction power.

A limitation in the use of ANNs to model deforestation is that ANNs provide a “black box” approach to the description of the relationship between two sets of variables. Even an ANN able to make perfect predictions would tell us nothing about the functional form of the relationship between the input variables and the output

layer. The weight matrices of the network do not have any direct meaning. Numerical experimentation with the ANNs may help, however, to identify the most important variables and the functional form of the relationship between them and the output(s).

More generally, the limitation of the model are the following: (a) the model is designed to predict only the location where land-use change is likely to occur, it does not predict the quantity of change and (b) the model does not take into consideration the regrowth of forest vegetation. Training an ANNs with data derived from a land cover change data which present such classes of change could allow to simulate the different types of changes. However, such modelling will need a more complex network (with more outputs), which may result in a loss of generalization capacity.

More general additional limitations of empirical deforestation models are that land-cover changes and their driving forces are not necessarily found at the same location and the intervention of unpredictable factors which change deforestation processes. Therefore, the model is not helpful in providing explanations about the processes of change and long-range projections are not reliable. These limitations affect the models independently of the approach used to establish the relationships between land use/cover change and explanatory variables. In the study area, main causes of deforestation were cattle ranching in the first period and rice cultivation in the second. The effect of the explanatory variables such as human settlement proximity, elevation and forest fragmentation remains similar from a period to another (Mas and Puig, 2001). However, variation of the relative importance of these explanatory variables over time leads to errors in predicting the location of deforestation. Accuracy may be gained developing prediction for a shorter time and developing models at a coarser scale (Pontius et al., 2001) that will not predict the exact location of future clearing but rather the region where exists a high risk of deforestation.

In this study, the model based on a very simple MLP structure presents a better performance, expressed as the power of prediction of deforestation location in the following period, than a model elaborated with the same data by means of a logistic regression, although improvement was not very significant. These results are in agreement with literature data, where performance of ANN has been repeatedly reported to overpass those of more traditional method such as regression models (Lek et al., 1996; Lek-Ang et al., 1999; Paruelo and Tomasel, 1997). ANNs constitute a powerful alternative in spatial land-cover change processes modelling, when more conventional models obtain poor performance. However, it is probably impossible to develop models of deforestation processes, and more generally land use change processes, which present a high power of prediction because these processes depend upon very diverse fac-

tors from environmental to socio-economic and cultural that are changing over time. In the present study, the location of forest clearing was predicted with a reasonable precision and the model could be used by environmental planners and managers to develop policies aimed at controlling the adverse ecological and social effects of deforestation.

References

- Abdul-Wahab, S.A., Al-Alawib, S.M., 2002. Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. *Environmental Modelling and Software* 17 (3), 219–228.
- Almeida, C.M., Batty, M., Monteiro, A.M.V., Câmara, G., Soares-Filho, B.S., Cerqueira, G.C., Pennachin, C.L., 2003. Stochastic cellular automata modelling of urban land use dynamics: empirical development and estimation. *Computers, Environment and Urban Systems* 27 (5), 481–509.
- Atkinson, P.M., Tatnall, A.R.L., 1997. Neural networks in remote sensing. *International Journal of Remote Sensing* 18 (4), 699–709.
- Benediktsson, J.A., Sveinsson, J.R., 1997. Feature extraction for multi-source data classification with artificial neural networks. *International Journal of Remote Sensing* 18, 727–740.
- Bishop, C.M., 1995. In: *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, 482.
- Bocco, G., Mendoza, M., Masera, O., 2001. La dinámica de cambio del uso del suelo en Michoacán. Una propuesta metodológica para el estudio de los procesos de deforestación. *Investigaciones Geográficas* 44 (18), 38.
- Chelani, A.B., Chalapati Rao, C.V., Phadke, K.M., Hasan, M.Z., 2002. Prediction of sulphur dioxide concentration using artificial neural networks. *Environmental Modelling and Software* 17 (2), 159–166.
- Civco, D.L., 1993. Artificial neural network for land cover classification and mapping. *International Journal of Geographical Information Systems* 7 (2), 173–186.
- Cortina Villar, S., Macario Mendoza, P., Ogneva-Himmelberger, Y., 1999. Cambios en el uso del suelo y deforestación en el sur de los estados de Campeche y Quintana Roo, México. *Investigaciones Geográficas* 38, 41–56.
- De Brujin, A.C., 1991. Spatial factors in urban growth: towards GIS models for cities in developing countries. *ITC Journal* 4, 221–231.
- Díaz Gallegos, J.R., García Gál, G., Castillo Acosta, O., March Mifret, I., 2001. Uso del suelo y transformación de selvas en un ejido de la Reserva de la Biosfera Calakmul, Campeche, México. *Investigaciones Geográficas* 44, 39–53.
- Fontan, J., 1994. Changements globaux et développement. *Nature—Sciences—Sociétés* 2 (2), 143–152.
- Foody, G.M., 1995. Land cover classification by an artificial neural network with ancillary information. *International Journal of Geographical Information Systems* 9 (5), 527–542.
- Foody, G.M., Arora, M.K., 1997. An evaluation of some factors affecting the accuracy of classification by an artificial neural network. *International Journal of Remote Sensing* 18 (4), 799–810.
- Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. *Neural Computation* 4, 1–58.
- Geoghegan, J., Villar, S.C., Klepeis, P., Mendoza, P.M., Ogneva-Himmelberger, Y., Chowdhury, R.R., Turner, B.L., Vance, C., 2001. Modelling tropical deforestation in the southern Yucatan peninsula region: comparing survey and satellite data. *Agriculture, Ecosystems and Environment* 85, 25–46.
- Grainger, A., 1993. Rates of deforestation in the humid tropics: estimates and measurements. *The Geographical Journal* 159 (1), 33–44.
- International Geosphere Biosphere Program (IGBP), 1993. Relating

- Land Use and Global Land Cover Change. In: Turner II, B.L., Moss, R.H., Skole, D.L. Report no. 24.
- Intergovernmental Panel on Climate Changes (IPCC), 1996. Climate Change 1995, second assessment report. Cambridge University Press, Cambridge, UK.
- Isaac-Márquez, R., 1993. Evaluación del cultivo de arroz como fuente de contaminación e impacto ambiental para la Laguna de Términos, Campeche, México. Tesis profesional, Escuela de Biología, Universidad Autónoma de Guadalajara, Mexico.
- Jensen, J.R., Qiu, F., Ji, M., 1999. Predictive modelling of coniferous forest age using statistical and artificial neural network approaches applied to remote sensor data. *Ecological Modelling* 20 (14), 2805–2822.
- Kavzoglu, T., Mather, P.M., 2000. Using feature selection techniques to produce smaller neural networks with better generalisation capabilities. In: Proceedings of IGARSS 2000, vol. 7., pp. 3069–3071.
- Kim, G., Barros, A.P., 2001. Quantitative flood forecasting using multisensor data and neural networks. *Journal of Hydrology* 246 (1-4), 45–62.
- Lambin, E.F., 1994. Modelling deforestation processes, a review. EUR 15744 EN, TREES series B: Research Report No. 1. Joint Research Centre, Institute for Remote Sensing Applications; European Space Agency, Luxembourg, Office for Official Publications of the European Community, p128.
- Lambin, E.F., 1997. Modelling and monitoring land-cover change processes in tropical regions. *Progress in Physical Geography* 21, 375–393.
- Landgrebe, D., Biehl, L., 2000. In: An Introduction to MultiSpec. Purdue Research Foundation, West Lafayette, 168.
- Lek, S., Guégan, J.F., 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling* 120, 65–73.
- Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., Aulanier, S., 1996. Application of neural networks to modelling non-linear relationships in ecology. *Ecological Modelling* 90, 39–52.
- Lek-Ang, S., Deharveng, L., Lek, S., 1999. Predictive models of collembolan diversity and abundance in a riparian habitat. *Ecological Modelling* 120, 247–260.
- Li, X., Yeh, A.G.-O., 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science* 16 (4), 323–343.
- Maier, H.R., Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling and Software* 15 (1), 101–124.
- Manel, S., Dias, J.M., Ormerod, S.J., 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with Himalayan river bird. *Ecological Modelling* 120, 337–347.
- Mas, J.F., 1999. Monitoring land-cover changes: a comparison of change detection techniques. *International Journal of Remote Sensing* 20 (1), 139–152.
- Mas, J.F., Puig, H., 2001. Modalités de la déforestation dans le Sud-ouest de l'Etat du Campeche, Mexique. *Canadian Journal of Forest Research* 31 (7), 1280–1288.
- Mas, J.F., Sorani, V., Alvarez, R., 1996. Elaboración de un modelo de simulación del proceso de deforestación. *Investigaciones Geográficas* 5, 43–57.
- Mas, J.F., Sosa-López, A., Palacio, J.L., Puig, H., 2000. Modelling deforestation in the region of the Lagoon of Términos, South East Mexico. In: Proceedings of the ASPRS (American Society of Photogrammetry and Remote Sensing) Annual Convention, Washington, DC, May 22–26, 2000.
- Mas, J.F., Velázquez, A., Díaz, J.R., Mayorga, R., Alcántara, C., Castro, R., Fernández, T., 2002. Assessing land use/cover changes in Mexico. In: Proceedings of the 29th International Symposium on Remote Sensing of Environment (CD), Buenos Aires, Argentina, April 8–12, 2002.
- Matheron, G., 1970. La théorie des variables généralisées et ses applications. *Les Cahiers du Centre de Morphologie Mathématiques de Fontainebleau*, Fascicule, 5.
- Mausel, P.W., Kramber, W.J., Lee, J.K., 1990. Optimum band selection for supervised classification of multispectral data. *Photogrammetric Engineering and Remote Sensing* 56 (1), 55–60.
- Mertens, B., Lambin, E.F., 1997. Spatial modelling of deforestation in southern Cameroon, spatial disaggregation of diverse deforestation processes. *Applied Geography* 17 (2), 143–162.
- Paruelo, J.M., Tomasel, F., 1997. Prediction of functional characteristics of ecosystems: a comparison of artificial neural networks and regression models. *Ecological Modelling* 98, 173–186.
- Pijanowski, B.C., Brown, D.G., Shellito, B.A., Manik, G.A., 2002. Using neural networks and GIS to forecast land use changes: a land transformation model. *Computers, Environment and Urban Systems* 26 (6), 553–575.
- Pontius, R.G., Cornell, J.D., Hall, C.A.S., 2001. Modelling the spatial pattern of land-use change with GEOMOD2: application and validation in Costa Rica. *Agriculture, Ecosystems and Environment* 85, 191–203.
- Puig, H., 2000. Diversité spécifique et déforestation: l'exemple des forêts tropicales humides du Mexique. *Bois et Forêts des tropiques* 268 (2), 41–55.
- Rosenfield, G.H., Fitzpatrick-Lins, K., 1986. A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing* 52 (2), 223–227.
- Rosin, P.L., Fierens, F., 1995. Improving neural network generalization. In: Proceedings of IGARSS 1995, vol. 2., pp. 1255–1257.
- Sader, A.S., Joyce, A.T., 1988. Deforestation rates and trends in Costa Rica, 1940 to 1983. *Biotropica* 20 (1), 11–19.
- Schneider, L.C., Pontius, 2001. Modeling land-use change in the Ipswich watershed, Massachusetts, USA. *Agriculture Ecosystem and Environment* 85 (1-3), 83–94.
- Shahshahani, B.M., Landgrebe, D.A., 1994. The effect of unlabelled samples in reducing the small sample size problem and mitigating the Hughes phenomenon. *IEEE Transactions on Geoscience and Remote Sensing* 32 (5), 1087–1095.
- Soares-Filho, B.S., Pennachin, C.L., Cerqueira, G., 2002. DINAMICA—a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling* 154 (3), 217–235.
- Stéphenne, N., Lambin, E.F., 2001. A dynamic simulation model of land-use changes in Sudano-sahelian countries of Africa (SALU). *Agriculture, Ecosystems and Environment* 85, 154–161.
- Tourenq, C., Aulagnier, S., Mesléard, F., Durieux, L., Johnson, A., Gonzalez, G., Lek, S., 1999. Use of artificial neural networks for predicting rice crop damage by greater flamingos in the Camargues, France. *Ecological Modelling* 120, 349–358.
- Turner, B.L., Villar, S.C., Foster, D., Geoghegan, J., Keys, E., Klepeis, P., Lawrence, D., Mendoza, P.M., Manson, S., Ogneva-Himmelberger, Y., Plotkin, A.B., Pérez Salicrup, D., Chowdhury, R.R., Savitsky, B., Schneider, L., Schmook, B., Vance, C., 2001. Deforestation in the southern Yucatan peninsula region: an integrative approach. *Forest Ecology and Management* 154, 353–370.
- Velázquez, A., Duran, E., Ramírez, I., Mas, J.-F., Ramírez, G., Bocco, G., Palacio, J.-L. Environmental conversion processes in highly biodiverse areas: the case of Oaxaca, Mexico. *Global Environment Change*, in press.
- Velázquez, A., Mas, J.F., Palacio-Prieto, J.L., Bocco, G., 2002. Land cover mapping to obtain a current profile of deforestation in Mexico. *Unasylva (FAO)* 210 (53), 37–40.
- Wang, F., 1994. The use of artificial neural networks in a geographical information systems for agricultural land suitability assessment. *Environment and Planning A* 26, 265–284.
- Yang, G., Collins, M.J., Gong, P., 1996. Multisource data integration with neural networks: optimum selection of net variables for lithologic. In: Proceedings of IGARSS 1996, vol. 4., pp. 2068–2070.