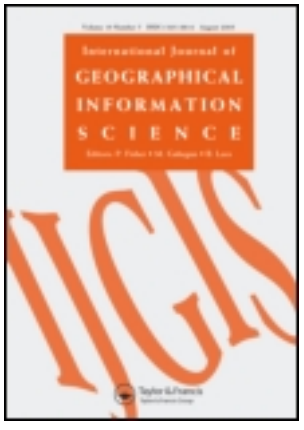
This article was downloaded by: [Vrije Universiteit Amsterdam]
On: 10 December 2011, At: 04:19
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Geographical Information Science

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgis20

Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling - a case study

Yu-Pin Lin [a], Hone-Jay Chu [a], Chen-Fa Wu [b] & Peter H. Verburg [c]

[a] Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan, ROC

[b] Department of Horticulture, National Chun Hsing University, Taichung, Taiwan, ROC

[c] Institute for Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands

Available online: 21 Oct 2010

To cite this article: Yu-Pin Lin, Hone-Jay Chu, Chen-Fa Wu & Peter H. Verburg (2011): Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling - a case study, International Journal of Geographical Information Science, 25:1, 65-87

To link to this article: http://dx.doi.org/10.1080/13658811003752332

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings,

demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

International Journal of Geographical Information Science
Vol. 25, No. 1, January 2011, 65–87

Taylor & Francis
Taylor & Francis Group

Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling – a case study

Yu-Pin Lin[a]*, Hone-Jay Chu[a], Chen-Fa Wu[b] and Peter H. Verburg[c]

[a]*Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan, ROC;* [b]*Department of Horticulture, National Chun Hsing University, Taichung, Taiwan, ROC;* [c]*Institute for Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands*

(*Received 26 February 2009; final version received 16 February 2010*)

The objective of this study is to compare the abilities of logistic, auto-logistic and artificial neural network (ANN) models for quantifying the relationships between land uses and their drivers. In addition, the application of the results obtained by the three techniques is tested in a dynamic land-use change model (CLUE-s) for the Paochiao watershed region in Taiwan. Relative operating characteristic curves (ROCs), kappa statistics, multiple resolution validation and landscape metrics were used to assess the ability of the three techniques in estimating the relationship between driving factors and land use and its subsequent application in land-use change models. The validation results illustrate that for this case study ANNs constitute a powerful alternative for the use of logistic regression in empirical modeling of spatial land-use change processes. ANNs provide in this case a better fit between driving factors and land-use pattern. In addition, auto-logistic regression performs better than logistic regression and nearly as well as ANNs. Auto-logistic regression and ANNs are considered especially useful when the performance of more conventional models is not satisfactory or the underlying data relationships are unknown. The results indicate that an evaluation of alternative techniques to specify relationships between driving factors and land use can improve the performance of land-use change models.

Keywords: auto-logistic regression; artificial neural networks; landscape metrics; empirical land-use change model

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

1. Introduction

Land-use change models have been developed to analyze the interaction between driving factors and land-use changes, and to predict land-use change patterns and variations in space and time. Apart from being learning tools for unraveling the driving forces and system dynamics, land-use change models play an important role in exploring possible future developments in the land-use system (Verburg *et al*. 2006b). The diverse modeling approaches that have evolved in recent years have motivated researchers to review and classify the different approaches (Agarwal *et al*. 2002, Verburg *et al*. 2004).

Such classification systems are based primarily on the dominant land-use change processes addressed by the model, the simulation technique used in the model or the underlying theory (Verburg *et al*. 2004). The variety of land-use change models include

*Corresponding author. Email: yplin@ntu.edu.tw

ISSN 1365-8816 print/ISSN 1362-3087 online
© 2011 Taylor & Francis
DOI: 10.1080/13658811003752332
http://www.informaworld.com

stochastic models, optimization models, dynamic process-based simulation models and empirical models (Li and Yeh 2002, Verburg *et al*. 2002, Dai *et al*. 2005, Castella *et al*. 2007, Dendoncker *et al*. 2007). Most land-use change models rely on an inductive approach in which the model specification is based on statistical associations between the land-use change of interest and a suite of explanatory variables that provide insight into the change (Overmars and Verburg 2005). Empirically parameterized models generally utilize statistical techniques to compute land-use change probabilities, indicating the likelihood of occurrence of a specific land-use type at a location (Almeida *et al*. 2008).

The relationship between land use and its drivers is often estimated using logistic regression prior to empirical land-use modeling (Lin *et al*. 2008). When using a logistical regression approach, one must be cautious about spatial autocorrelations that often exist in spatially referenced data because they may violate the assumption of the model (Hu and Lo 2007). An autocovariate can be used to correct for the effect of spatial autocorrelation. Called auto-logistic regression, this approach increases the predictive accuracy and model versatility (Dennis *et al*. 2002, Svenning and Skov 2002, Koutsias 2003, Boll *et al*. 2005, Betts *et al*. 2006, Piorecky and Prescott 2006, Dendoncker *et al*. 2007).

Quantifying all the potential interactions between the different drivers of land use in a logistic regression model is difficult given (1) the lack of understanding of all of these factors, (2) the lack of sufficient information and (3) the restrictions of the functional form of the logistic regression model (Ojima *et al*. 1994, Lambin and Geist 2006). Artificial neural networks (ANNs) were developed to mimic the brain's interconnected system of neurons so that computers could be made to imitate the brain's ability to sort patterns and learn by trial and error, and thereby observe the relationships in data (Pijanowski *et al*. 2002). Moreover, ANNs can take any nonlinear complex relationship between the driving variables and land use into account (Pijanowski *et al*. 2002, 2005, Dai *et al*. 2005). In recent years, a number of researchers have successfully applied ANN models in land-use change modeling (e.g., Pijanowski *et al*. 2002, Mas *et al*. 2004, Dai *et al*. 2005, Pijanowski *et al*. 2005, Almeida *et al*. 2008, Liu and Seto 2008). In addition, ANNs have been integrated with other models, such as cellular automata (CA), for land-use change modeling (Li and Yeh 2001, 2002, Almeida *et al*. 2008). Apart from the different capacities of these empirical techniques to quantify the relationship between driving factors and land use, it is also important to analyze how the results of these techniques can be used in dynamic simulation models. Pijanowski *et al*. (2005) have shown that it is not always the empirical model with the best fit that provides the most accurate land-use simulation.

One commonly used model for land-use simulation is the Conversion of Land Use and its Effects (CLUE-s) model. The CLUE-s model has been successfully applied in simulating land-use changes in response to different spatial and nonspatial policies (Verburg *et al*. 2006a; Castella and Verburg 2007, Castella *et al*. 2007, Lesschen *et al*. 2007, Lin *et al*. 2007a, 2007b, 2008). In this model, logistic regression analyses are commonly used to calculate the probabilities of land-use changes prior to the allocation procedure. To our knowledge, the allocation results of CLUE-s modeling based on alternative specifications of the relationships between drivers and land use by logistic regression, auto-logistic regression and ANNs have not been compared previously.

The objective of this study is to demonstrate the abilities of logistic regression, auto-logistic regression and ANN models to quantify the relationships between land-use patterns and their drivers. Specifically, we assess the efficiency of the three models in CLUE-s modeling for simulating land-use changes in the Paochiao watershed region in Taiwan.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

2. Methods and materials

2.1. Study watershed

The Paochiao watershed is bordered to the northeast by the Tamsui River Basin in northern Taiwan (Figure 1). The study area is located at 25.00°N, 121.62°E. The area of the watershed is 98.61 km^2. The elevation distribution is between 8 and 683 m and the mean elevation is 215 m. In the model, the dimensions are 128 (rows) by 211 (columns) and each cell size is equal to 80 m. Due to the population increase in the Taipei metropolitan area, land-use patterns have changed over the last decade. The main processes of change include urbanization as a threat to forest, agricultural land and other land uses. These processes of change are representative for many rural areas in the neighborhood of metropolitan areas.

Four SPOT (Satellite Pour l'Observation de la Terre) satellite images acquired on 27 March 1990, 25 December 1993, 16 July 1998 and 2 January 2000 were selected for land-use classification in this study. The images were classified using supervised classification, performed by the ERDAS IMAGINE software with 1/5000 black and white aerial photographs provided by the Aerial Survey Office of the Taiwan Forest Bureau, with maximum likelihood and fuzzy methods (Lin *et al*. 2008). In the study, a total of 300 training areas were used to evaluate the final accuracy of each SPOT image. The training area

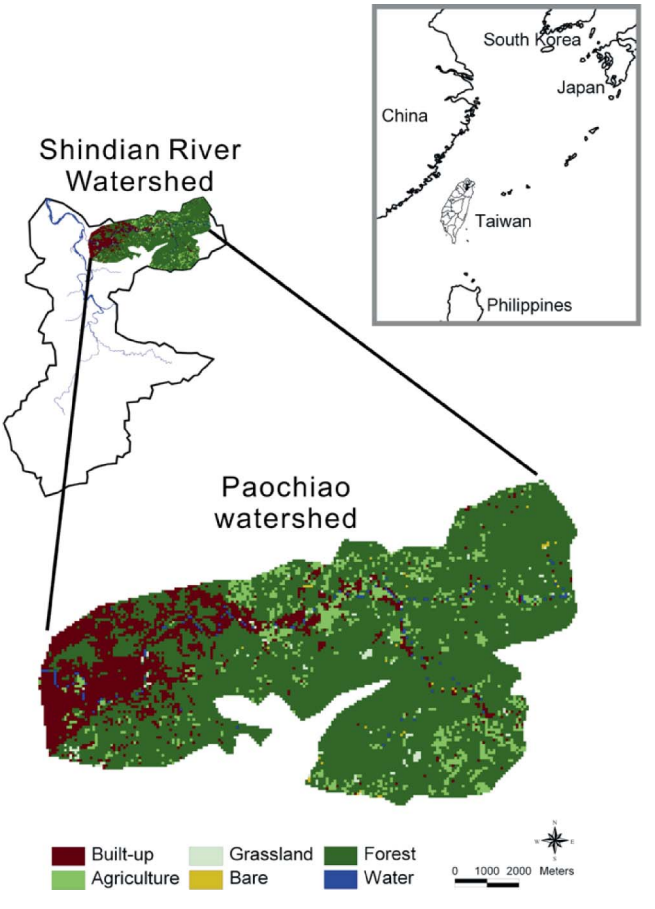


Figure 1. The location of the Paochiao watershed and the land-use distribution pattern in 2000.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

numbers per class are 133 training areas for forest, 78 training areas for built-up land, 43 training areas for agricultural land, 19 training areas for grassland, 12 training areas for bare land and 15 training areas for water. The land-use classes distinguished include forest, built-up land, agricultural land, grassland, water and bare land. The land uses covered forest (74.8%), built-up land (15.3%), agricultural land (8.2%), grassland (0.8%), water (0.7%) and bare land (0.2%) of the area in 2000. The kappa values for classification of the images for 1990, 1993, 1998 and 2000 were 0.86, 0.85, 0.86 and 0.84, respectively. The land-use classes of forest, built-up land, agricultural land and grassland had high classification accuracy rates (92–96%, 85–96%, 73–90% and 63–90%, respectively), but the classification accuracy rate for bare land was low (42–60%).

The socio-economic and biophysical driving factors of land-use allocation in the study area were selected based on the knowledge of the study area and factors frequently identified in land-change studies (Geist and Lambin 2002). A more detailed analysis of the driving factors for the same case study area is presented by Lin *et al*. (2008). The factors used include distance to transport, elevation, slope, distance to river, distance to built-up land, distance to urban planning area, soil drainage, soil erosion coefficient and population density. Data for all types of land-use and driving factors were inputs to the logistic regression, auto-logistic regression, and ANN models to calculate probability maps for the occurrence of each land-use type. Based on the land-use change probabilities calculated by the above models, the land-use change model (CLUE-s) simulated backward land-use patterns in 1993 and 1998 for the validation. A variety of methods such as kappa and multi-resolution validation were used to compare the simulation results for 1993 and 1998 with independent observations of land cover in those years. Finally, a simulation for the period 2001–2015 was made for which the change in landscape patterns by landscape metrics was analyzed with the Patch Analyst extension for ArcView (McGarigal and Marks 1995, Elkie *et al*. 1999).

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

2.2. Quantification of land use and drivers

2.2.1. Logistic regression

The logistic regression provides the probability of the presence/absence of each land use at each location based on their drivers (Verburg *et al*. 2004). The logistic regression quantifies the relationships between different types of land use and their drivers, which is specified by

$$p_{il} = \frac{\exp\left(\beta_{0l} + \sum_{j=1}^{k} \beta_{jl} x_{ji}\right)}{1 + \exp\left(\beta_{0l} + \sum_{j=1}^{k} \beta_{jl} x_{ji}\right)}, \tag{1}$$

where p_{il} is the probability of the occurrence of land-use class l at grid cell (pixel) i, k is the number of driving factors, x_{ji} is the value of cell i for the driving factor j, β_{0l} is the estimated constant and β_{jl} is the coefficient of driving factor j for land use l in the logistic model.

2.2.2. Auto-logistic regression

Auto-logistic models account for spatial autocorrelation through the addition of an auto-covariance variable, which is calculated for a specific neighborhood size (Piorecky and Prescott 2006). The formula used to express auto-logistic regression is as follows:

$$p'_{il} = \frac{\exp\left(\beta_{0l} + \sum_{j=1}^{k} \beta_{jl} x_{ji} + c_l COV_{il}\right)}{1 + \exp\left(\beta_{0l} + \sum_{j=1}^{k} \beta_{jl} x_{ji} + c_l COV_{il}\right)}, \quad (2)$$

where p'_{il} is the probability of the occurrence of land use l that accounts for spatial autocorrelation at pixel i; β_{0l}, β_{jl}, and c_l are estimated coefficients for each land use and the autocovariate (COV_{il}) is calculated by

$$COV_{il} = \frac{\sum_{j \in N_i} w_{ij} y_{jl}}{\sum_{j \in N_i} w_{ij}}, \quad (3)$$

where N_i represents the neighborhoods around the pixel i, w_{ij} is the weight factor determined by the inverse of the Euclidean distance between i and j within the neighborhoods and y_{jl} represents the occurrence of land-use l at pixel i.

In this study, geostatistics approaches determine the specific neighborhood sizes for each land use. A relatively consistent set of best-fit models with minimum RSS (model reduced sum of squares) and maximum r^2 values were generated by least-squares model fitting of indicator semivariance models in GS+ (geostatistical software; Gama Design, 1995). The ranges of the exponential indicator semivariance models of forest, built-up land, agricultural land, grassland and bare land were 800, 670, 750, 506 and 160 m, respectively. These ranges of indicator semivariograms were set as the neighborhood sizes for the autocovariate calculation.

2.2.3. *Artificial neural networks*

An ANN consists of an interconnected group of artificial neurons that process information using a connectionist approach to computation. Neural networks are hierarchically arranged layers of interconnected units that process information in a highly parallel processing fashion (Pijanowski *et al*. 2002). Each unit, called a node (analogous to a neuron), is connected to other units in network by a weighted connection so that each unit receives input from many nodes in the previous layer (Pijanowski *et al*. 2002). The most common neural network is a multilayer perceptron (MLP), which contains three types of layers: input, hidden and output (Pijanowski *et al*. 2002). A number of neurons are arranged in an input layer, one or more hidden layers and an output layer. There are two neural processing phases: learning and recall. The learning process adapts the connection weights in an ANN to produce the desired output. Then, the recall process attempts to retrieve information based on the weights derived by the learning process, and predict output data of the new example. The ANNs in this article are feedforward networks with one input layer, one hidden layer and one output layer. Although there are several ways to construct the ANNs, back-propagation networks appear to be the most widely used in practice (Pijanowski *et al*. 2002). The neural network is designed to have a flexible number of inputs depending on the number of predictor variables presented to it, as well as an equal number of hidden units as input units, and a single output.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

All input grids that existed in Arc/Info Grid format were normalized in the range 0.0 to 1.0 and converted to ASCII representations (called a pattern file). The pattern file contained information from the nine inputs, which were the driving factors. The output of the ANNs represents the likelihood of a land-use type at each pixel. The ANNs predict outputs based on the weights derived by the learning process. The functions can be written as follows:

$$O_{j1} = f\left(\sum_i O_i w_{i,j1} - b_{j1}\right) \tag{4}$$

$$O_{jn} = f\left(\sum_{jn-1} O_{jn-1} w_{jn-1,jn} - b_{jn}\right) \tag{5}$$

$$O_k = f\left(\sum_{jn} O_{jn} w_{jn,k} - b_k\right) \tag{6}$$

where O_i denotes the input; O_k denotes the output; O_{jn}, O_{jn-1}, …, O_{j1} denote the hidden units in n th, $(n-1)$th, …, 1st hidden layer; i, j and k denote the input unit, hidden unit and output unit, respectively; w denotes a connected weight; b represents the bias value and f is a transfer function.

There is a three-layer feedforward network in the study. The neural networks are designed to have nine inputs based on the number of driving factors presented to it. The transfer functions in the model are a hyperbolic tangent sigmoid function for a hidden layer and a linear function for the output layer. The total available data have been divided into two sets, training and validation set: the cells (samples) of the entire map in 2000 were used to train the ANN, and the cells of the entire map in 1993 and 1998 were used to calculate goodness of fit for validation. The training stopping criterion is MSE = 10^{-5}. If the criterion is not met, the ANN training algorithm will continue.

2.3. *CLUE-s – an empirical land-use change model*

An empirical land-use model, called the Conversion of Land Use and its Effects (CLUE-s) model, was employed to simulate future land-use patterns in the study area (Verburg *et al*. 2002, Verburg and Overmars 2009). The model combines an empirical specification of relationships between driving factors and land-use allocation with the dynamic simulation of competition among land uses due to changes in regional-level demand for the different land-use types. Conversion elasticities and specific conversion rules are implemented to account for the costs of land-use conversion and to restrict unlikely conversions. Details of the model can be found in Verburg and Overmars (2009).

In this study, regional-level land-use demands in the CLUE-s model were set according to simulation results obtained by the SLEUTH (Slope, Land use, Exclusion, Urban extent, Transportation, Hillshade) model (Clarke *et al*. 1997) for 2001–2025 made in another study (Lin *et al*. 2008). It was assumed that the area of the water body remained constant during the simulation period. Areas with slope $>$ 21% were defined as a restricted area (Lin *et al*. 2008). Land-use transition rules allow forested land, agricultural land, grassland, and bare land to be converted to any of these land-use classes or to built-up land.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

2.4. Comparison of results

Relative operating characteristics (ROCs), kappa statistics, a multi-resolution validation procedure and landscape metrics were used to compare, validate and analyze the behavior of the three different model specifications.

2.4.1. Relative operating characteristic

The area under the ROC curve was calculated to measure the explanatory power of each model (Boll *et al*. 2005). The ROC curve is constructed by calculating the sensitivity and specificity of the resulting classification for each possible classification. The ROC is a measure for the goodness of fit of a logistic regression model similar to the r^2 statistic in ordinary least-squares regression. ROC values above 0.7 are generally considered good while values exceeding 0.9 are considered to indicate an excellent model fit. Since the ROC method is considered a proper measure to evaluate the goodness of fit (Swets 1986, Manel *et al*. 2001), we applied it to assess the fit of the logistic regression, auto-logistic regression and ANN models.

2.4.2. Kappa statistics

For the calculation of a kappa statistic, the simplest assessment is to measure the proportion of agreement between two observed maps (the observed and simulated data) accounting for the proportion of observed agreement (p_0) and the proportion of chance agreement (p_e). The kappa statistics can be calculated as (Congalton and Mead 1983, Jensen 1996, Sousa *et al*. 2002, Pijanowski *et al*. 2005, Saito *et al*. 2005)

$$\kappa = \frac{p_0 - p_e}{1 - p_e} = \frac{\sum_{i=1}^{c} p_{ii} - \sum_{i=1}^{c} p_{iT} p_{Ti}}{1 - \sum_{i=1}^{c} p_{iT} p_{Ti}} \tag{7}$$

where $p_0 - p_e$ is the difference between the proportion of observed agreement and that of agreement by chance, while $1 - p_e$ is interpreted as the maximum possible correct classification beyond that expected by chance (Cook 1998); c is the number of categories; p_{iT} indicates the proportion of cells in category i of observed change, taken from the marginal totals of the last column of the contingency matrix; p_{Ti} indicates the proportion of cells in category i of the simulation, taken from the marginal totals of the last row of the contingency matrix and p_{ii} indicates the proportion of cells in the same category, i, on both observed changes and simulation results, taken from the diagonal elements of the contingency matrix.

2.4.3. Multi-resolution validation procedure

Another method to compare the observed and simulated results is the multiple resolution procedure. This method compares the maps over a range of resolutions in order to account for both small location errors and larger errors in a differentiated manner. The details of the method are described in Costanza (1989) as well as in Castella and Verburg (2007). The fit at a particular sampling window size (F_w) is calculated as follows:

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

$$F_w = \frac{\sum_{j=1}^{t_w} \left[1 - \sum_{i=1}^{p_c} a_{i,S} - a_{i,R} \right]_j}{t_w} \tag{8}$$

where F_w is the fit for sampling window size w; w is the dimension of one side of the sampling window; $a_{i,S}$ is the number of cells of category i in the simulations in the sampling window; $a_{i,R}$ is the number of cells of category i in the references in the sampling window; p_c is the number of different categories in the sampling windows; j is the sampling window of dimension w by w, which slides through the scene one cell at a time and t_w is the total number of sampling windows in the scene for window size w.

To use these measures to determine an overall degree of fit between two maps, a weighted average of the fit over a range of window sizes is calculated. This can be done by giving exponentially less weight to the fit at lower resolution, shown as follows:

$$F_t = \frac{\sum_w F_w e^{-D(w-1)}}{\sum_w e^{-D(w-1)}} \tag{9}$$

where F_t is a weighted average of the fits over all window sizes and D is a constant that determines how much weight is to be given to small vs. large sampling windows ($D = 0.2$ in this study)

2.4.4. Landscape metrics

In order to test the influence of the different model specifications on the development of spatial patterns of land use, an additional comparison was made based on landscape metrics for the validation period (1990, 1993 and 1998) and the simulation for 2001–2015. Landscape metrics were calculated using the Patch Analyst in the GIS software ArcView 3.2a (Elkie *et al*. 1999), which is designed to compute a wide variety of landscape metrics for categorical map patterns. In this study, the following metrics were used (Table 1): the number of patches (NP), mean patch size (MPS), total number of edges (TE), mean shape index (MSI), mean nearest neighbor (MNN) and interspersion and juxtaposition index (IJI) (McGarigal and Marks 1995). Given the relatively small changes in landscape pattern during the validation period, we only compared the behavior of the models in terms of landscape metrics for the period 2001–2015.

3. Results

3.1. Logistic regression, auto-logistic regression and ANN models

The logistic regression model was used to predict the probabilities for all land-use classes and measure the coefficients between land uses and their driving factors (Table 2). For the forest class in the watershed, the distance to major roads, elevation, slope, distance to a built-up land and soil drainage were positively associated with the likelihood to find forest at the location, while the distance to a river and population density negatively correlated with the probability. Population density positively is correlated with the occurrence of built-up land, but the distance to a built-up land is negatively correlated with the probability for built-up land. The obvious correlation of current built-up land with the distance to built-up land is

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

Table 1. Landscape metrics.

Name	Equation	Note
Number of patches (NP)	$NP = n_i$	Patch size metrics
Mean patch size (MPS)	$MPS = \frac{1}{n_i}\sum_{j=1}^{n_i} a_{ij}$	Patch size metrics
Total edge (TE)	$TE = \sum_{k=1}^{m} e_{ik}$	Edge metrics
Mean shape index (MSI)	$MSI = \frac{\sum_{j=1}^{n_i} \frac{0.25 p_{ij}}{\sqrt{a_{ij}}}}{n_i}$	Shape metrics
Mean nearest neighbor (MNN)	$MNN = \frac{\sum_{j=1}^{n_i} h_{ij}}{n_i}$	Diversity metrics
Interspersion and juxtaposition index (IJI)	$IJI = \frac{-\sum_{k=1}^{m} \frac{e_{ik}}{\sum_{k=1}^{m} e_{ik}} \ln\left(\frac{e_{ik}}{\sum_{k=1}^{m} e_{ik}}\right)}{\ln(m-1)}(100)$	Diversity metrics

Note: n_i is the number of patches in land-use class i ; a_{ij} is the jth patch area (m^2) in land-use class i ; $m$is the total number of patch classes; e_{ik} is the total length (m) of the edge between patch classes i and k; p_{ij} is the jth patch perimeter (m) in land-use class i ; h_{ij} is the distance (m) from the jth patch to the nearest neighboring patch of the same class i, based on the edge-to-edge distance.

Table 2. Logistic regression model for land-use classes.

Variable[a]	Forest	Built-up land	Agricultural land	Grassland	Bare land
DTransport	0.164***	–	–0.205***	–	–0.546*
DEM	0.267***	–	0.239***	–	–
Slope	0.958***	–	–0.470***	–0.469***	–
DRiver	–0.122***	–	–	0.189*	–
DBuild	1.326***	–190.711*	–0.207***	–	–
DZone	–	–	0.119***	0.273**	0.533***
SDr	0.082***	–	–	–	–
SErosin	0.088**	–	0.314***	0.349**	0.491*
PDens	–0.063*	0.052	–	–	0.196***
Constant	1.875	–174.469	–2.577***	–4.915	–6.349
ROC[b]	0.866	1	0.653	0.647	0.715

Note: '–' represents variable is not selected; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$;.
[a]DTransport represents the minimum distance to a major road, DEM represents the altitude, Slope represents the slope, DRiver represents the minimum distance to a river, DBuild represents the minimum distance to a built-up land, DZone represents the minimum distance to an urban planning area, SDr represents the soil drainage, SErosin represents the soil erosion coefficient and PDens represents the population density.
[b]ROC represents the area under the ROC curve.

directly responsible for the perfect model fit. To predict the probability of agricultural land, the fitted logistic model used three positive coefficient factors (elevation, distance to urban planning area and soil erosion) and three negative coefficient factors (distance to major roads, slope and distance to built-up land). To predict the probability of grassland, three positive coefficient factors (distance to river, distance to urban planning area and soil

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

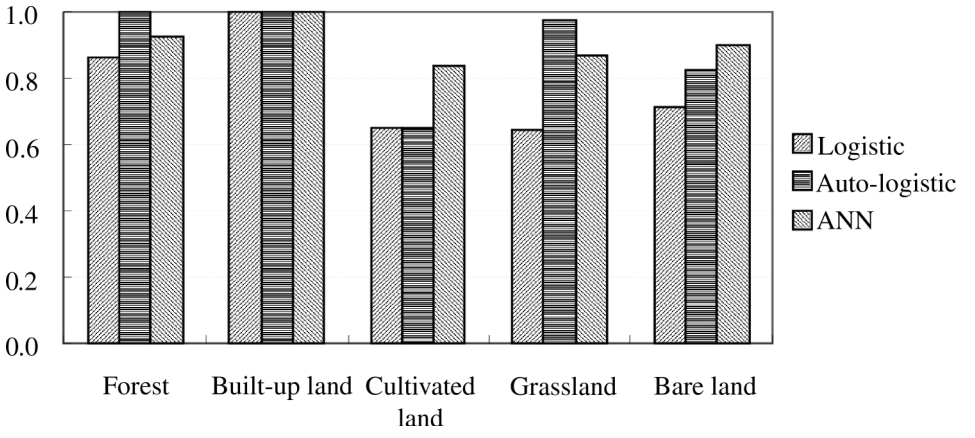


Figure 2. Relative operating characteristics (ROCs) for the three models.

erosion) and one negative coefficient factor (slope) were used to fit the logistic model. Finally, the distance to an urban planning area, soil erosion and population density positively correlated with the probability for bare land, while the distance to major roads negatively correlated with the probability for bare land in the watershed. In addition, the ROC values for the logistic regression model are in the range 0.647–1 (Figure 2).

In Table 3, the auto-logistic regression model for predicting forest areas in the watershed has five positive coefficients for the driving factors (i.e., distance to major roads, distance to river, distance to built-up land, soil drainage and forest autocovariate) and four negative coefficient factors (i.e., elevation, slope, soil erosion and population density). For agricultural land, the driving factors were elevation, distance to urban planning area, soil erosion and the agricultural land autocovariate, each of which has positive coefficient in the auto-logistic regression for agricultural land. The distance to major roads, slope and distance to

Table 3. Auto-logistic regression model for land-use classes.

Variable[a]	Forest	Built-up land	Agricultural land	Grassland	Bare land
DTransport	18.544	–	–0.111*	–	–0.081
DEM	–138.912	–	0.188***	–	–
Slope	–1.557	–	–0.439***	–0.753	–
DRiver	3.044	–	–	–0.342	–
DBuild	21.449	–168.948	–0.105*	–	–
DZone	–	–	0.090*	0.478*	0.423*
SDr	5.509	–	–	–	–
SErosin	–19.459	–	0.240***	–0.002	0.051
PDens	–3.873	0.131	–	–	–0.242
Autocov	372.999	10.586	0.211***	14.424	8.044
Constant	354.732	–152.401	–2.583	–6.538	–6.717
ROC[b]	1	1	0.654	0.975	0.825

Note: '–' represents variable is not selected; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.
[a]DTransport represents the minimum distance to a major road, DEM represents the altitude, Slope represents the slope, DRiver represents the minimum distance to a river, DBuild represents the minimum distance to a built-up land, DZone represents the minimum distance to an urban planning area, SDr represents the soil drainage, SErosin represents the soil erosion coefficient and PDens represents the population density.
[b]ROC represents the area under the ROC curve.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

built-up land are negatively correlated with the predicted probability for agricultural land. The distance to built-up land is negatively correlated with the probability for built-up land, while population density and the built-up land autocovariate are positively correlated with probability for built-up land. Two factors, distance to urban planning area and grassland autocovariate, are positively correlated with the probability for grassland. Three factors (slope, distance to a river and soil erosion) are negatively correlated with the probability for grassland. To predict the probability of bare land, the fitted auto-logistic model used three positive coefficient factors (distance to an urban planning area, soil erosion and the bare land autocovariate) and two negative coefficient factors (the distance to major roads and population density). Accordingly, the models' ROC values for the auto-logistic regression model were in the range 0.654–1 (Figure 2). The results demonstrate that the auto-logistic regression is superior to logistic regression.

The ANN models used in this study were a three-layer, feedforward network, and the hidden layer contained 20 neurons. The ROC values of the ANN models for forest, built-up land, agricultural land, grassland and bare land are 0.92, 1.00, 0.836, 0.873 and 0.903, respectively (Figure 2). The high ROC values indicate the very good fit of the model to the observations which may be explained by the capacity of ANNs to capture complex, non-linear relationships.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

3.2. *Model goodness of fit results*

Model goodness of fit results for both simulations between 1993 and 1998 are shown in Table 4. The overall accuracy and kappa statistics for ANN-CLUE-s are higher than the values for both Autologistic-CLUE-s and Logistic-CLUE-s. The accuracy of the ANN-CLUE-s model is best for both years and all resolutions (Figure 3). The overall agreement increases as the resolution becomes coarser for both the models because location disagreement becomes agreement as the resolution becomes coarser. Table 4 also shows the *Ft* values of three models in 1993 and 1998, which highlights that the ANN-CLUE-s model has the best fit. The *Ft* values represent a weighted average of the agreement over the window size varying between 1 pixel (80 m) and 20 pixels (1600 m). *Ft* is generally used to determine an overall degree of fit between two maps across multiple resolutions (Costanza 1989).

3.3. *Land-use change scenario for 2001–2015*

Based on the probabilities calculated by the logistic regression, auto-logistic regression and ANN models, the CLUE-s model was applied to simulate land-use patterns in the study watershed from 2001 to 2015. Figure 4 shows the land-use maps of the Logistic-CLUE-s, Autologistic-CLUE-s and ANN-CLUE-s models for 2015. In all three simulation results the

Table 4. Model goodness of fit results for the three model implementations for 1993 and 1998.

Year	Model	Overall accuracy (%)	Kappa	Multi-resolution goodness of fit (*Ft*)
1993	ANN-CLUE-s	92.3	0.80	0.947
	Autologistic-CLUE-s	89.5	0.72	0.915
	Logistic-CLUE-s	86.3	0.64	0.892
1998	ANN-CLUE-s	90.5	0.75	0.933
	Autologistic-CLUE-s	86.1	0.64	0.892
	Logistic-CLUE-s	84.6	0.60	0.887

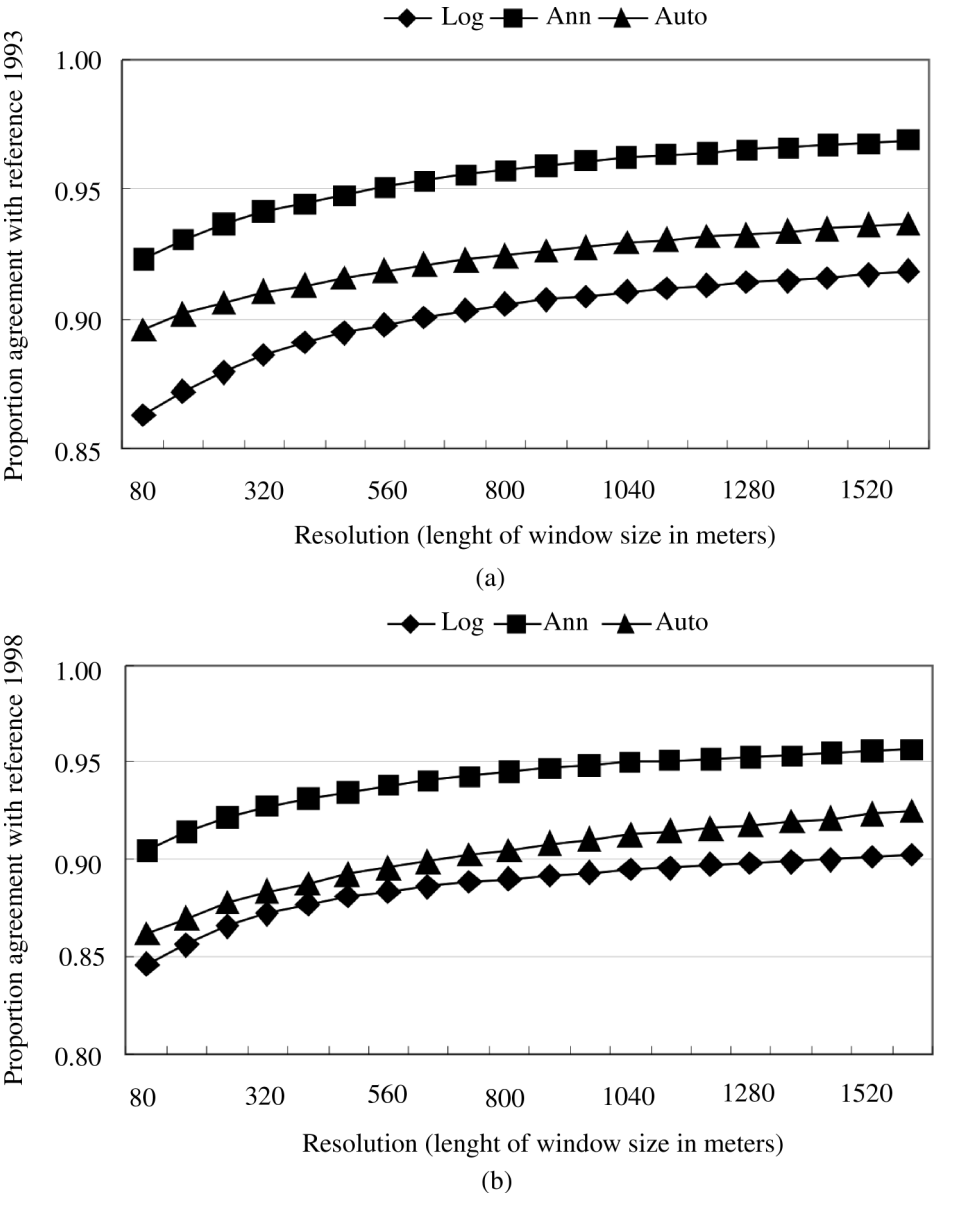


Figure 3. Proportion agreement at multiple resolutions for the three models and the reference map for (a) 1993 and (b) 1998.

built-up land gradually increases along the river while the locations of forested land, agricultural land and grassland change and their sizes decrease. The areas of agricultural land, forest land and grassland decrease by 51%, 12% and 43% respectively during the period while the built-up area increases by 87%. The maps show that the most developed and dynamic areas of the Paochiao watershed are located in the middle and downstream areas, especially in areas with low elevations.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

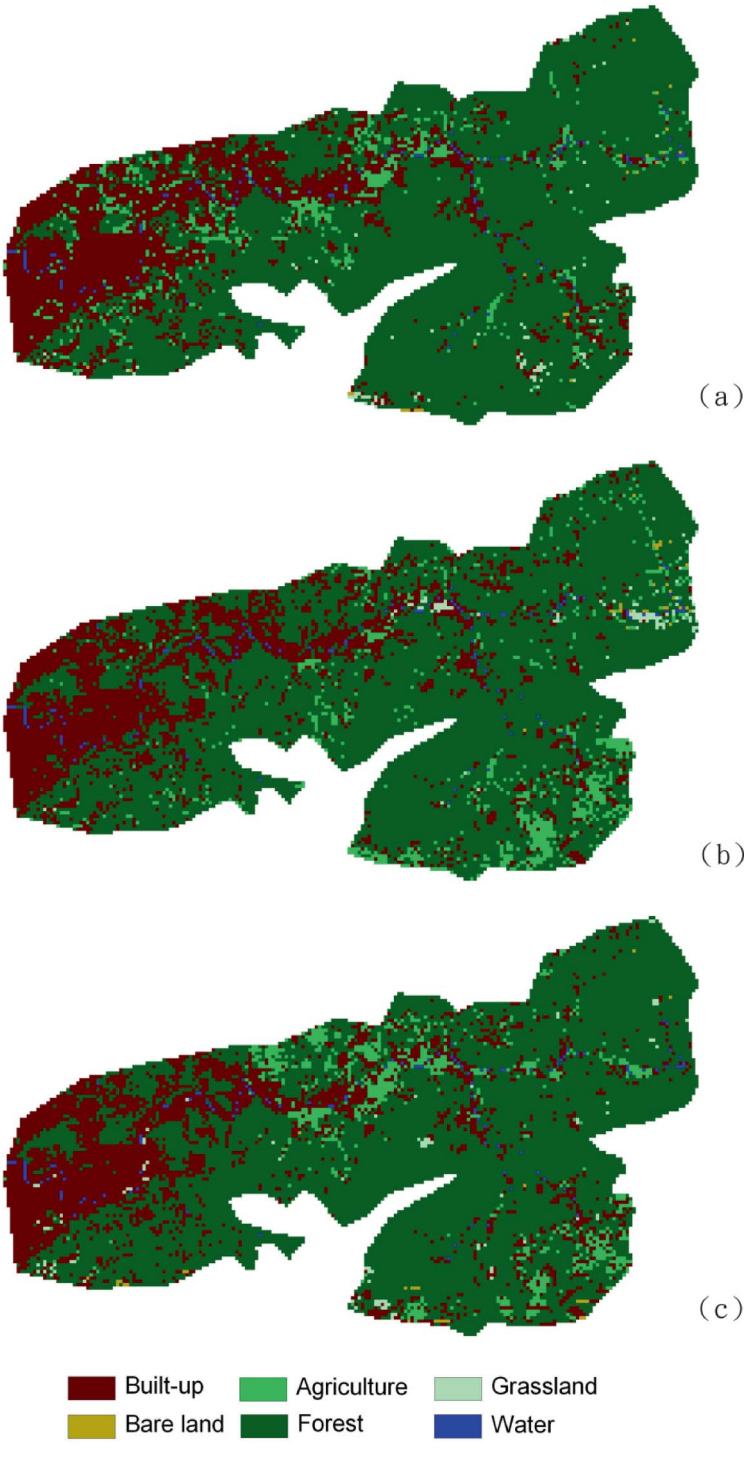


Figure 4. Spatial land-use distribution in the watershed simulated by CLUE-s based on (a) logistic regression, (b) auto-logistic regression and (c) ANN in 2015.

Figure 5 shows the land-use change between land uses in 2000 and 2015. Each model provides various forecasting results, especially with regard to spatial pattern of land uses. The land-use change maps show that land-use changes simulated by Logistic-CLUE-s are spread over the entire study watershed. Simulated built-up and agricultural lands were spread over the downstream and mid-stream areas, and agricultural lands were interspersed among the built-up lands (Figure 5a). Clustered agricultural land areas simulated by the Autologistic-CLUE-s model were mostly (66.46%) located in the upstream area, with some areas (20.32%) in the mid-stream area (Figure 5b). The ANN-CLUE-s simulated agricultural land clusters in the mid-stream area (44.48%), and some in the upstream area (36.27%) (Figure 5c).

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

3.4. Landscape metrics of simulated land-use patterns

Figure 6 shows the landscape metrics at the landscape level of the observed land-use patterns in 1990, 1993, 1998 and 2000, as well as the land-use patterns simulated by the Logistic-CLUE-s, Autologistic-CLUE-s and ANN-CLUE-s models for the period 2001–2015. During the validation periods (1990, 1993 and 1998), the landscape metrics for the number of patches (NP: Figure 6a) and the mean shape index (MSI: Figure 6c) associated negatively with the mean patch size (MPS: Figure 6b) and the total number of edges (TE: Figure 6d), respectively. For the period 2001–2015, the trend of the MPS values in Autologistic-CLUE-s simulated land-use patterns is similar to that of ANN-CLUE-s patterns. Moreover, the trends of the mean nearest neighbor (MNN: Figure 6e) at the landscape level are similar in the three models. However, the trend of the interspersion and juxtaposition index values (IJI: Figure 6f) for the simulated land-use patterns based on auto-logistic regression is higher than that of the other models. The reason is that because each land-use pattern, especially that for agricultural land, is more dispersed in the Autologistic-CLUE-s model (Figure 4), juxtaposed agricultural lands with other land-use classes yield high IJI values.

Figures 7–9 show the landscape metrics at the class level for built-up areas, agricultural land and forested land, respectively. The MNN trends for built-up areas and agricultural land are similar in the three models from 2001 to 2015 (Figures 7e and 8e). During the same period, the trends of the TE values for simulated agricultural land derived by the three models are similar (Figure 8d); for simulated built-up areas, the trends of the MPS and TE values derived by ANN-CLUE-s and Logistic-CLUE-s are similar (Figure 7b and d); and the MSI trends of ANN-CLUE-s and Autologistic-CLUE-s are similar (Figure 7c). For simulated forest lands, with the exception of IJI, the trends of the landscape metrics derived by the three models are different during the simulation period (Figure 9f). Even so, analysis of the landscape metrics derived by the different methods implies the same phenomenon overall: the size of forest patches decreases as built-up patches sprawl.

4. Discussion

4.1. Model goodness of fit by logistic regression, auto-logistic regression and ANNs

Results from the model fit indicate that ANNs are superior to the logistic and auto-logistic models in this particular case study. The ROC and kappa values of the ANNs in predicting agricultural land are better than those of the logistic and auto-logistic regression, implying a complex relationship between agricultural land-use pattern and its drivers. The ROC and kappa values of the auto-logistic regression and ANNs in predicting bare lands are significantly greater than that of the logistic regression, implying that the pattern of bare land

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

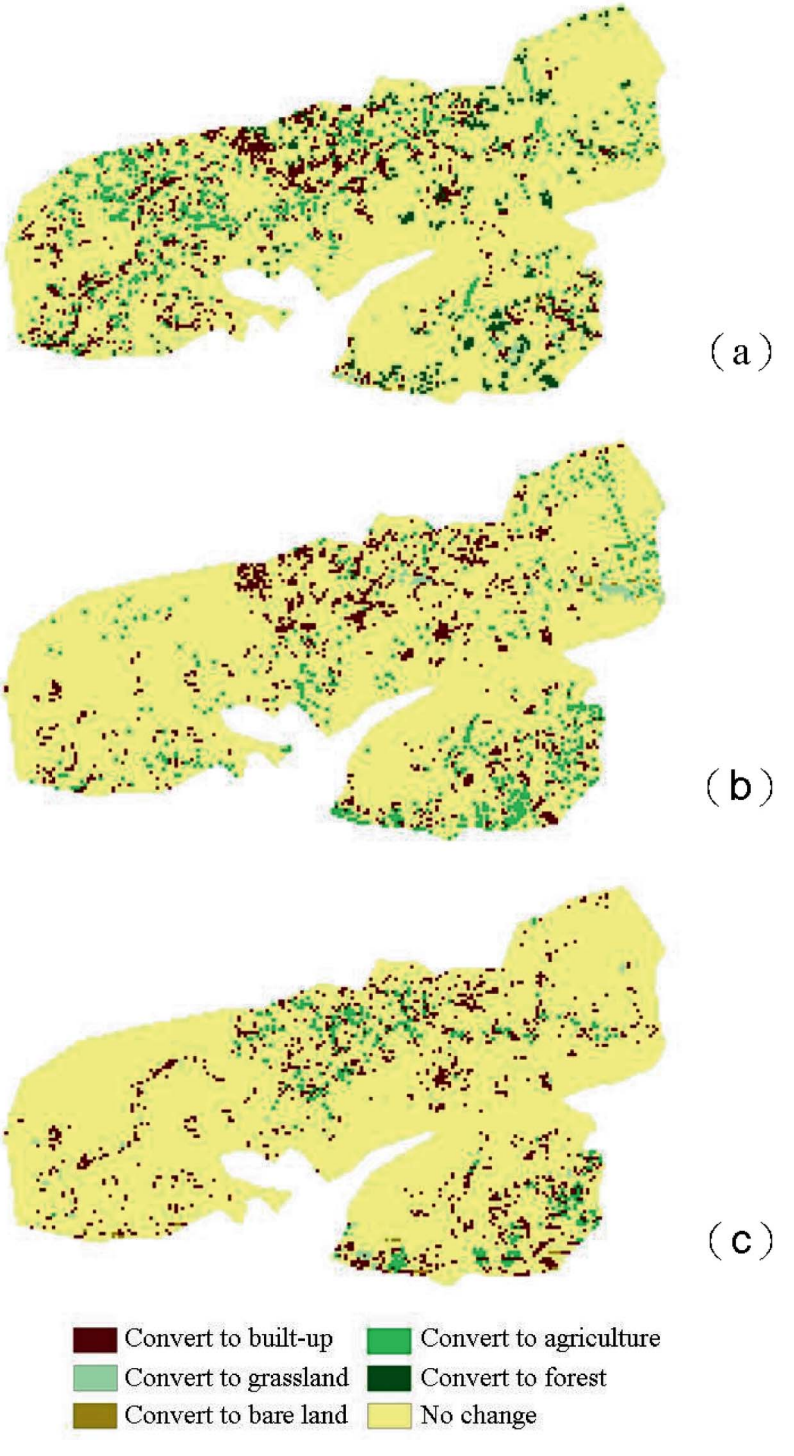


Figure 5. Land-use changes between 2000 and 2015 using (a) logistic regression, (b) auto-logistic regression and (c) ANN.

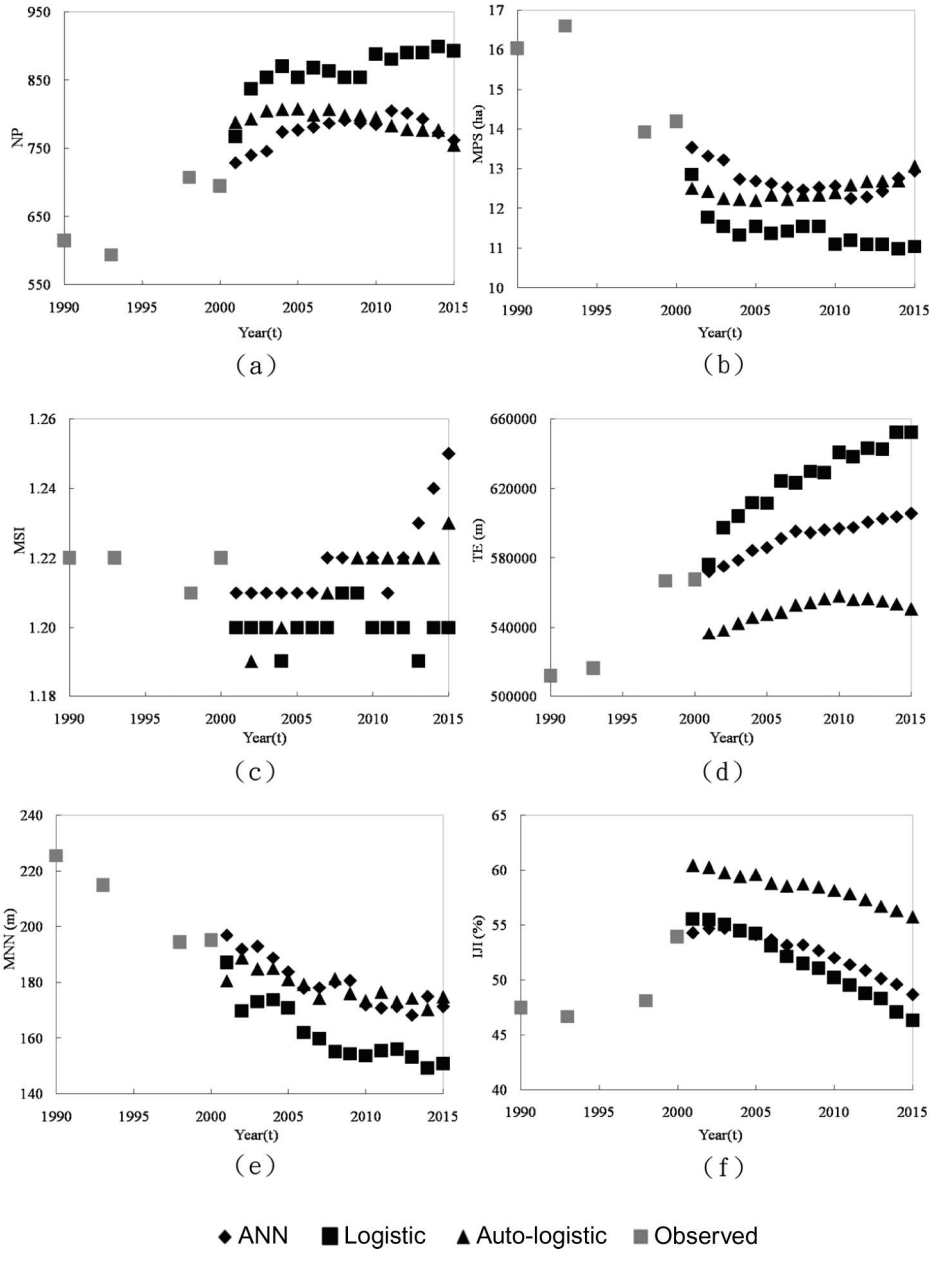


Figure 6. Results of the simulated landscape metrics at landscape level – (a) NP, (b) MPS, (c) MSI, (d) TE, (e) MNN and (f) IJI (see Table 1) – based on observed data and simulation results of the CLUE-s model with different specifications.

not only exhibits spatial correlation but also shows a complex relation with its drivers. The validation results also confirm that ANNs have the potential to produce good models of urban change (Pijanowski *et al*. 2005) and constitute a powerful alternative for modeling spatial land-cover change processes when the performance of conventional models is not satisfactory (Mas *et al*. 2004) or the underlying data relationships are unknown (Dai *et al*.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

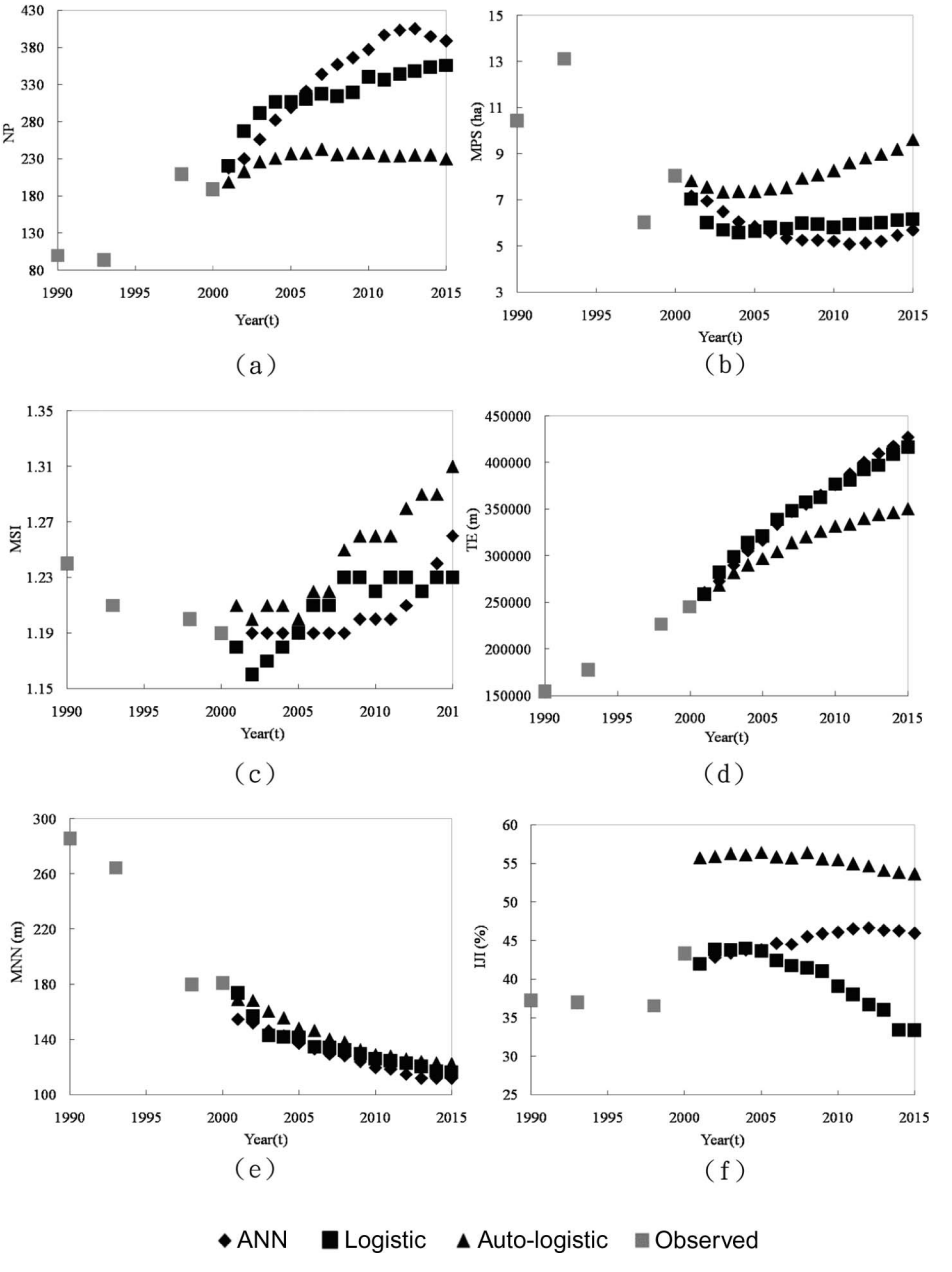


Figure 7. Results of the simulated landscape metrics for built-up areas – (a) NP, (b) MPS, (c) MSI, (d), TE, (e) MNN and (f) IJI (see Table 1) – based on observed data and simulation results of the CLUE-s model with different specifications.

2005). While ANNs can account for nonlinear complex relationships between the driving variables and changes in land use, ANNs do not directly provide insights in the relationships between dependent and independent variables. However, with additional techniques, it is possible to obtain some insight in the associations between driving variables and land use (Bishop 1995, Pijanowski *et al*. 2002, Chu and Chang 2009).

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

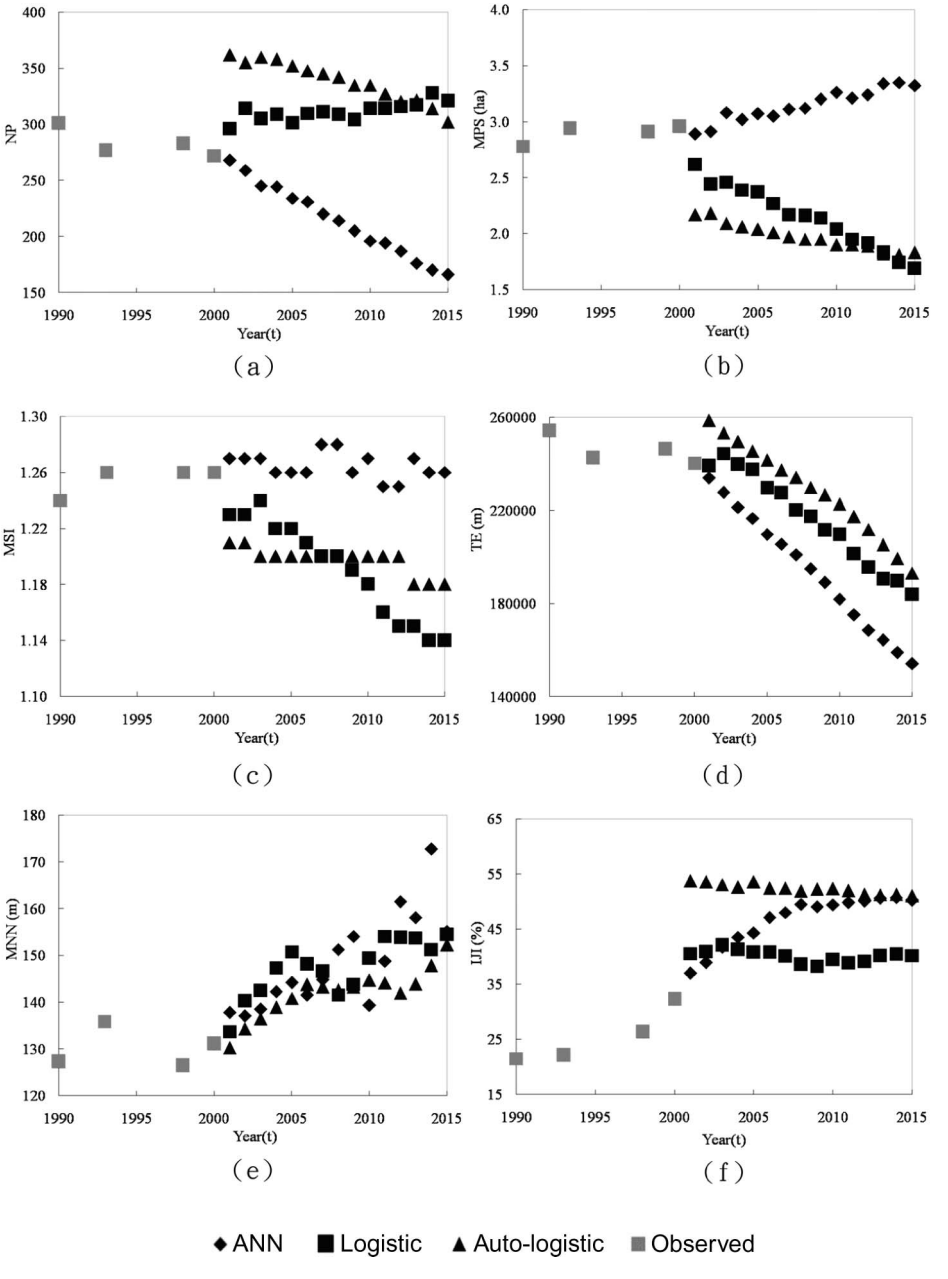


Figure 8. Results of the simulated landscape metrics for agricultural land – (a) NP, (b) MPS, (c) MSI, (d), TE, (e) MNN and (f) IJI (see Table 1) – based on observed data and simulation results of the CLUE-s model with different specifications.

Auto-logistic regression models include one or more neighborhood variables that explicitly account for spatial autocorrelation in the data; hence, this allows us to assess the effects of the different neighborhood variables while statistically considering for the effect of spatial autocorrelation (Svenning and Skov 2002, Boll *et al*. 2005). In most cases land-use patterns

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

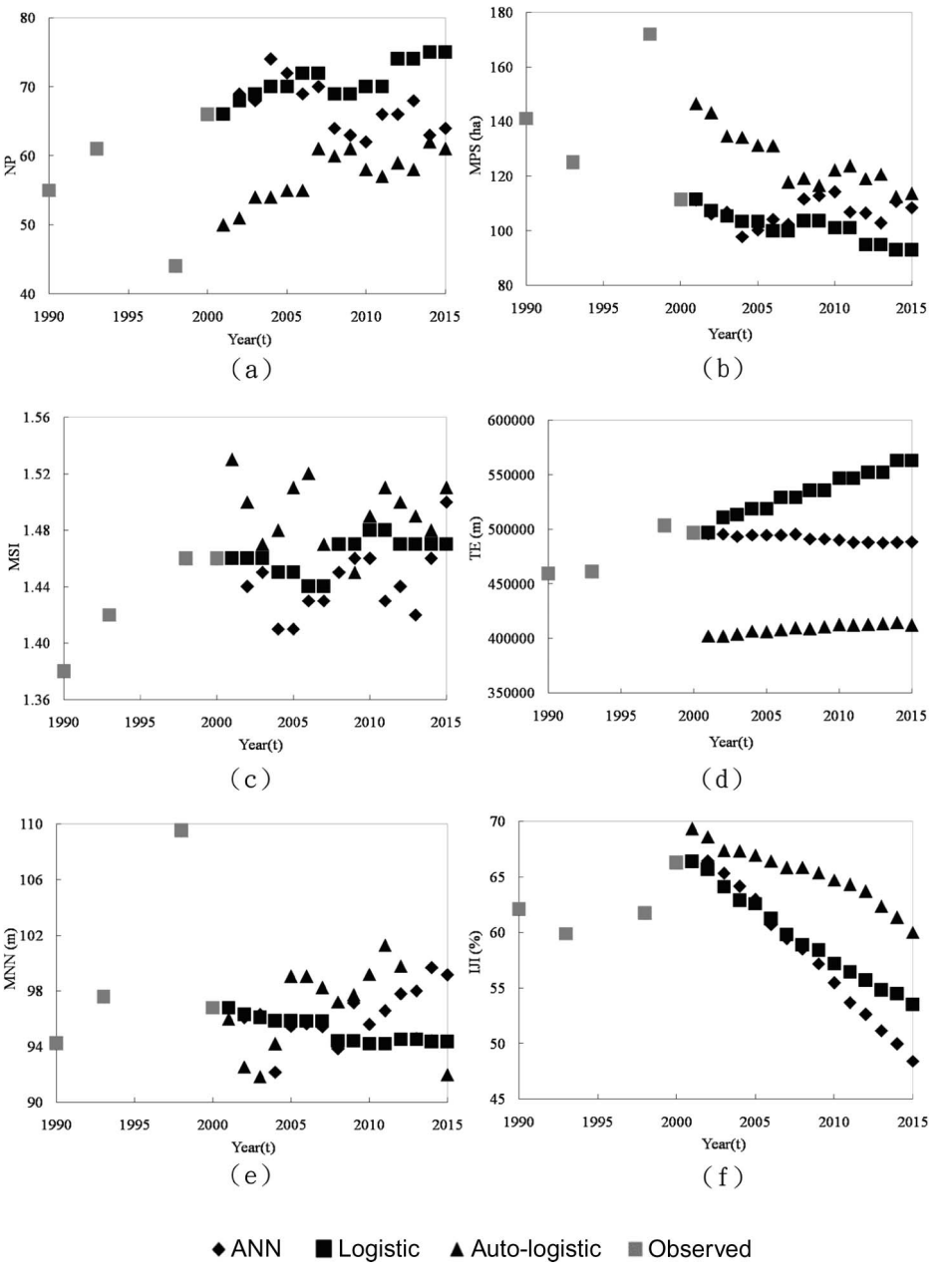


Figure 9. Results of the simulated landscape metrics for forest areas – (a) NP, (b) MPS, (c) MSI, (d), TE, (e) MNN and (f) IJI (see Table 1) – based on observed data and simulation results of the CLUE-s model with different specifications.

exhibit spatial autocorrelation (Verburg *et al*. 2006b). This is mainly due to the clustered distribution of landscape features and gradients in environmental conditions that are important determinants of the land-use pattern (Verburg *et al*. 2006b). The average prediction ability of auto-logistic regression (average ROC value = 0.89) is slightly less than that of the

ANNs (average ROC value = 0.91). Moreover, the ability of the auto-logistic regression model to predict forested areas, built-up areas and grasslands is better than that of the logistic regression model and the ANN models (Figure 2).

The results of the model comparison made in this study are only valid for the specific case study at the chosen spatial and temporal scales. In different landscapes, different processes and different drivers of land-use change may cause different methods to be most suited for land-use simulation. Therefore, case-specific testing of alternative methods is preferable above choosing a method based on arbitrary criteria or habit. Although logistic regression has been widely used to predict land-use changes, ANNs and auto-logistic regression should be considered when land-use patterns show complex relations or are spatially autocorrelated.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

4.2. Landscape metrics of simulated land uses

Landscape metrics enhance a way to characterize and quantify land-use composition and configuration both for analyzing the change in landscape composition and patterns as well as for informing land-use planning. Mapping time-varying spatial metrics is beneficial for analyzing urban growth as they provide a comprehensive method for describing a process, comparing cities and making comparisons with theory (Herold *et al*. 2003, Lin *et al*. 2008). They can also be used to assess the performance of land-use models (Herold *et al*. 2003, Pijanowski *et al*. 2005, Lin *et al*. 2008). Increasing the number of patches (NP), total number of edges (TE) and interspersion and juxtaposition index (IJI) associated with decreasing mean patch size (MPS), mean shape index (MSI) and mean nearest neighbor (MNN) show that land-use patterns in the watershed tend to fragment, and had regular shapes and interspersion patterns (Lin *et al*. 2008). Similar to the land-use patterns for the 1990–1998 period, all CLUE-s-simulated land-use patterns tended to fragment and had regular shapes and interspersion patterns; however, they were relatively less isolated and less interspersed from 2001 to 2015 compared with the land-use pattern in 1990. Such developments may be enhanced by the autocorrelation in the independent variables the model is using in its allocation procedure and the lack of information on local scale conditions such as tenure and the objectives of individual land managers.

5. Conclusion

Land-use change models, which play an important role in exploring possible future land-use patterns, have been developed to delineate the driving factors that influence land-use changes, as well as predict land-use change patterns and variations in space and time precisely. This study investigates the ability of logistic regression, auto-logistic regression and ANN models to detect land-use changes and the relationships between land uses and their drivers for a specific case study. Specifically, it assesses the performance of the three empirical models as inputs to the dynamic land-use simulation model. In this case, the results indicate that the CLUE-s model with ANNs constitutes a powerful alternative for models based on logistic regression, especially when the performance of conventional models is not satisfactory or the underlying data relationships are unknown. Similar to ANNs, the auto-logistic regression provides explanatory power along with a relatively high degree of goodness of fit and leads to an improvement of performances. The alternative techniques such as auto-logistic regression and ANNs should be considered when land-use patterns are spatially autocorrelated or with complex relations.

This article has indicated that a comparison of alternative techniques to quantify relationships between driving factors and land use leading to different parameterizations of simulation models is useful to improve the performance of land-use models. Future studies could consider testing if similar effects are observed in other case studies representing similar and different types of landscape, landscape metrics variation, the effects of the different processes of land-use change and different drivers and the grain, level and scale in the empirical land-use change model.

Acknowledgment

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under contract nos. 96-2415-H-002-022-MY3 and 98-2621-M-002-023.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

References

Agarwal, C., et al., 2002. *A review and assessment of land-use change models: dynamics of space, time, and human Choice*. Gen. Tech. Rep. NE-297. Newton Square, PA: US Department of Agriculture, Forest Service, Northeastern Research Station, 12–27.

Almeida, C.M., et al., 2008. Using neural networks and cellular automata for modelling intra-urban land-use dynamics. *International Journal of Geographical Information Science*, 22 (9), 943–963.

Betts, M.G., Diamond, A.W., and Forbes, G.J., 2006. The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. *Ecological Modelling*, 191 (2), 197–224.

Bishop, C. M., 1995. *Neural networks for pattern recognition*. Oxford, UK: Oxford University Press.

Boll, T., Svenning, J.C., and Vormisto, J., 2005. Spatial distribution and environmental preferences of the piassaba palm *Aphandra natalia* (Arecaceae) along the Pastaza and Urituyacu Rivers in Peru. *Forest Ecology and Management*, 213 (1–3), 175–183.

Castella, J.C. and Verburg, P.H., 2007. Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam. *Ecological Modelling*, 202, 410–420.

Castella, J.C., et al., 2007. Combining topdown and bottom-up modelling approaches of land-use/cover change to support public policies: application to sustainable management of natural resources in northern Vietnam. *Land-use Policy*, 24 (3), 531–545.

Chu, H.J. and Chang, L.C., 2009. Optimal control algorithm and neural network for dynamic groundwater management. *Hydrological Processes*, 23, 2765–2773.

Clarke, K. C., Hoppen, S., and Gaydos, L., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay Area. *Environment and Planning B: Planning and Design*, 24, 247–261.

Congalton, R.G. and Mead, R.A., 1983. A quantitative method to test for consistency and correctness in photo-interpretation. *Photogrammetric Engineering and Remote Sensing*, 49 (1), 69–74.

Cook, R.J., 1998. Kappa. *In:* P. Armitage and T. Colton, eds. *The encyclopedia of biostatistics*. New York: John Wiley & Sons, Inc, 2160–2166.

Costanza, R., 1989. Model goodness of fit: a multiple resolution procedure. *Ecological Modelling*, 47, 199–215.

Dai, E., Wu, S.H., and Shi, W.Z., 2005. Modeling change-pattern-value dynamics on land-use: an integrated GIS and artificial neural networks approach. *Environmental Assessment*, 36 (4), 576–591.

Dendoncker, N., Rounsevell, M., and Bogaert, P., 2007. Spatial analysis and modeling of land-use distributions in Belgium. *Computers, Environment and Urban systems*, 31 (2), 188–205.

Dennis, R., et al., 2002. A comparison of geographical and neighbourhood models for improving atlas databases. The case of the French butterfly atlas. *Biological Conservation*, 108, 143–159.

Elkie, P.C., Rempel, R.S., and Carr, A.P., 1999. *Patch Analyst user's manual: a tool for quantifying landscape structure*. Ontario Ministry of Natural Resources Northwest Science and Technology Manual TM-002, Ontario: Thunder Bay, 4–12.

Geist, H.J. and Lambin, E.F., 2002. Proximate causes and underlying driving forces of tropical deforestation. *Bioscience*, 52 (2), 143–150.

Herold, M., Goldstein, N.C., and Clarke, K.C., 2003. The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment*, 86 (3), 286–302.

Hu, Z. and Lo, C.P., 2007. Modeling urban growth in Atlanta using logistic regression, computers. *Environment and Urban Systems*, 31, 667–688.

Jensen, J.R., 1996. *Introductory digital image processing: a remote sensing perspective*. 2nd ed. Upper Saddle River, NJ: Prentice Hall, Inc.

Koutsias, N., 2003. An autologistic regression model for increasing the accuracy of burned surface mapping using landsat thematic mapper data. *International Journal of Remote Sensing*, 24, 2199–2204.

Lambin, E.F. and Geist, H.J., eds., 2006. *Land-use and land-cover change: local processes and global impacts*. The IGBP Series. Berlin: Springer-Verlag.

Lesschen, J.P., et al., 2007. Identification of vulnerable areas for gully erosion under different scenarios of land abandonment in southeast Spain. *Catena*, 71, 110–121.

Li, X. and Yeh, A.G., 2001. Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environment and Planning A*, 33 (8), 1445–1462.

Li, X. and Yeh, A.G., 2002. Neural-network-based cellular automata for simulating multiple land-use changes using GIS. *International Journal of Geographical Information Science*, 16 (4), 323–343.

Lin, Y.P., et al., 2007a. Impacts of land-use change scenarios on hydrology and land-use patterns in the Wu-Tu watershed in northern Taiwan. *Landscape and Urban Planning*, 80 (1–2), 111–126.

Lin, Y.P., et al., 2007b. Modeling and assessing land-use and hydrological processes to future land-use and climate change scenarios in watershed land-use planning. *Environmental Geology*, 53 (3), 623–634.

Lin, Y.P., et al., 2008. Monitoring and predicting land-use changes and the hydrology of the urbanized Paochiao watershed in Taiwan using remote sensing data, urban growth models and a hydrological model. *Sensors*, 8, 658–680.

Liu, W. and Seto, K.C., 2008. Using the ART-MMAP neural network to model and predict urban growth: a spatiotemporal data mining approach. *Environment and Planning B: Planning and Design*, 35, 296–317.

Manel, S., Williams, H.C., and Ormerod, S.J., 2001. Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38, 921–931.

Mas, J.F., et al., 2004. Modelling deforestation using GIS and artificial neural networks. *Environmental Modeling and Software*, 19, 461–471.

Mcgarigal, K. and Marks, B.J., 1995. *FRAGSTATS: spatial pattern analysis program for quantifying landscape structure*. Portland, OR: US Deptartment of Agriculture, Forest Service, Pacific Northwest Research Station, 38–53.

Ojima, D., Galvin, K., and Turner, B. II, 1994. The global impact of land-use change. *BioScience*, 44, 300–304.

Overmars, K.P. and Verburg, P.H., 2005. Analysis of land use drivers at the watershed and household level: linking two paradigms at the Philippine forest fringe. *International Journal of Geographical Information Science*, 19 (2), 125–152.

Pijanowski, B.C., et al., 2002. Using neural networks and GIS to forecast land-use changes: a land transformation model. *Computers, Environment and Urban Systems*, 26, 553–575.

Pijanowski, B.C., Pithadia, S., and Shellito, B.A., 2005. Calibrating a neural network-based urban change model for two metropolitan areas of the Upper Midwest of the United States. *International Journal of Geographical Information Science*, 19 (2), 197–215.

Piorecky, M.D. and Prescott, D.R.C., 2006. Multiple spatial scale logistic and auto-logistic habitat selection models for northern pygmy owls, along the eastern slopes of Alberta's Rocky Mountains. *Biological Conservation*, 129, 360–371.

Saito, H., McKenna, S.A., and Goovaerts, P., 2005. Accounting for geophysical information in geostatistical characterization of unexploded ordnance (UXO) sites. *Environmental and Ecological Statistics*, 12 (1), 7–25.

Sousa, S., Caeiro, S., and Painho, M., 2002. Assessment of map similarity of categorical maps using kappa statistics: the case of Sado Estuary. *In: ESIG 2002*. Taguspark, Portugal: Associação de Utilizadores de Informação Geográ fica.

Svenning, J.C. and Skov, F., 2002. Mesoscale distribution of understory plants in temperate forest (Kalø, Denmark): the importance of environment and dispersal. *Plant Ecology*, 160, 169–185.

Swets, J.A, 1986. Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011

Verburg, P.H., et al., 2002. Modeling the spatial dynamics of regional land-use: the CLUE-S model. *Environmental Management*, 30 (3), 391–405.

Verburg, P.H. and Overmars, K., 2009. Combining top-down and bottom-up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landscape Ecology.* 24 (9), 1167–1181.

Verburg, P.H., Overmars, K.P., and Witte, N., 2004. Accessibility and land-use patterns at the forest fringe in the northeastern part of the Philippines. *Geographical Journal*, 170, 238–255.

Verburg, P.H., et al., 2006a. Downscaling of land-use change scenarios to assess the dynamics of European landscapes. *Agriculture, Ecosystems and Environment*, 114 (1), 39–56.

Verburg, P.H., et al., 2006b. Modeling land-use and land-cover change. *In*: E.F. Lambin and H.J. Geist, eds. *Land-use and land-cover change: local processes and global impacts*. The IGBP Series. Berlin: Springer-Verlag.

Downloaded by [Vrije Universiteit Amsterdam] at 04:19 10 December 2011