PERGAMON

Computers, Environment and Urban Systems
□ (□□□□) □–□

Computers, Environment and Urban Systems

www.elsevier.com/locate/compenvurbsys

Using neural networks and GIS to forecast land use changes: a Land Transformation Model

Bryan C. Pijanowski[a,*], Daniel G. Brown[b], Bradley A. Shellito[c], Gaurav A. Manik[d]

[a] *Department of Zoology, 207 Manly Miles Building, 1405 South Harrison Road, Michigan State University, East Lansing, Michigan 48824, USA*

[b] *School of Natural Resources and Environment, G520 Dana Building, 430 East University Drive, University of Michigan, Ann Arbor, Michigan 48109-1115, USA*

[c] *Department of Geography, 314 Natural Science Building, Michigan State University, East Lansing, Michigan 48824, USA*

[d] *Department of Electrical and Computer Engineering, 2120 Engineering Building, Michigan State University, East Lansing, Michigan 48824, USA*

Abstract

The Land Transformation Model (LTM), which couples geographic information systems (GIS) with artificial neural networks (ANNs) to forecast land use changes, is presented here. A variety of social, political, and environmental factors contribute to the model's predictor variables of land use change. This paper presents a version of the LTM parameterized for Michigan's Grand Traverse Bay Watershed and explores how factors such as roads, highways, residential streets, rivers, Great Lakes coastlines, recreational facilities, inland lakes, agricultural density, and quality of views can influence urbanization patterns in this coastal watershed. ANNs are used to learn the patterns of development in the region and test the predictive capacity of the model, while GIS is used to develop the spatial, predictor drivers and perform spatial analysis on the results. The predictive ability of the model improved at larger scales when assessed using a moving scalable window metric. Finally, the individual contribution of each predictor variable was examined and shown to vary across spatial scales. At the smallest scales, quality views were the strongest predictor variable. We interpreted the multi-scale influences of land use change, illustrating the relative influences of site (e.g. quality of views, residential streets) and situation (e.g. highways and county roads) variables at different scales. © 2001 Published by Elsevier Science Ltd. All rights reserved.

Keywords: Land use change; Artificial neural networks; Geographic information systems; Land transformation model

* Corresponding author. Fax: +1-517-432-1054.
E-mail address: pijanows@msu.edu (B.C. Pijanowski).

0198-9715/01/$ - see front matter
PII: S0198-9715(01)00015-1

1. Introduction

Changes in land use result from the complex interaction of many factors including policy, management, economics, culture, human behavior, and the environment (Dale, O'Neill, Pedlowski, & Southworth, 1993; Houghton, 1994; Medley, Okey, Barrett, Lucas, & Renwick, 1995; Richards, 1990; Vesterby & Heimlich, 1991; Wilder, 1985). An understanding of how land use changes occur is critical since these anthropogenic processes can have broad impacts on the environment, altering hydrologic cycles (Steiner & Osterman, 1988), biogeochemical dynamics (Flintrop et al., 1996), size and arrangements of natural habitats such as forests (Dale et al., 1993) and species diversity (Costanza, Kemp, & Boynton, 1993). Changes to land use can also affect local and regional economies (Bingham et al., 1995; Burchell, 1996). This paper illustrates how combining geographic information systems (GIS) and artificial neural networks (ANNs) can aid in the understanding the complex process of land use change.

A GIS-based Land Transformation Model — LTM (Pijanowski, Gage, Long, & Cooper, 2000) was developed to forecast land use change over large regions. This model can be configured to use a variety of socioeconomic, political and environmental inputs. The LTM can link changes in land use to ecological process models, such as groundwater flow and solute transport (Boutt et al., 2001) and forest cover change (Brown, Duh, & Drzyzga, 2000; Brown, Pijanowski, & Duh, 2001). It can also provide local land use planners and regional resource managers with information about the potential effects of land use change on the environment.

ANNs are powerful tools that use a machine learning approach to quantify and model complex behavior and patterns. ANNs are used for pattern recognition in a variety of disciplines, such as economics (Fishman, Barr, & Loick, 1991), medicine (Babaian, Miyashita, Evans, Eshenbach, & Ramimrez, 1997), landscape classification (Brown, Lusch, & Duda, 1998), image analysis (Fukushima, Miyake, & Takayuki, 1983), pattern classification (Ritter, Logan, & Bryant, 1988), climate forecasting (Drummond, Joshi, & Sudduth, 1998), mechanical engineering (Kuo & Cohen 1998), and remote sensing (Atkinson & Tatnall, 1997). The use of neural networks has increased substantially over the last several years because of the advances in computing performance (Skapura, 1996) and the increased availability of powerful and flexible ANN software.

This paper provides an update on the development of the LTM (Pijanowski, Machemer, Gage, Long, Cooper, & Edens, 1995; Pijanowski et al., 2000) by illustrating an application of ANNs and GIS to model land use change. The principal objective of this paper is to demonstrate how GIS and neural network tools can be used to forecast land use changes over large regions. Specifically, we present:

1. the basic principles of ANNs as they apply to land use change modeling;
2. the use of GIS to develop spatial data layers for use as inputs to the ANNs;
3. methods for quantifying the model's predictive ability and the relative contribution of input variables across multiple scales; and

4. an application of the model that (1) develops a control run of the model within Grand Traverse County, Michigan, and (2) illustrates the extension of the county level model for predicting development patterns within the six-county Grand Traverse Bay Watershed.

2. Background

2.1. ANNs

ANNs were developed to model the brain's interconnected system of neurons so that computers could be made to imitate the brain's ability to sort patterns and learn from trial and error, thus observing relationships in data. Rosenblatt (1958) is credited with developing one of the first artificial neural networks when he created his "perceptron". The perceptron (Fig. 1A) consists of a single node, which receives weighted inputs and thresholds the results according to a defined rule. This type of simple neural machine is capable of classifying linearly separable data and performing linear functions.

The multi-layer perceptron (MLP) neural net described by Rumelhart, Hinton, and Williams (1986) is one of the most widely used ANNs. The MLP consists of three layers: input, hidden, and output (Fig. 1B) and thus can identify relationships that are non-linear in nature. ANN algorithms calculate weights for input values, input layer nodes, hidden layer nodes and output layer nodes by introducing the input in a feed forward manner, which propagates through the hidden layer and the output layer. The signals propagate from node to node and are modified by weights associated with each connection. The receiving node sums the weighted inputs from all of the nodes connected to it from the previous layer. The output of this node is then computed as the function of its input called the "activation

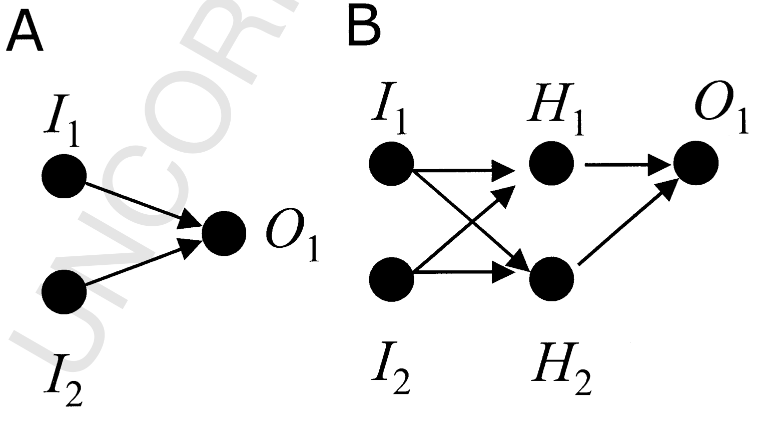

Fig. 1. A simple perceptron (A) and a multi-layer perceptron (B) illustrating input layers, hidden nodes and output layers.

function''. The data moves forward from node to node with multiple weighted summations occurring before reaching the output layer.

Weights in an ANN are determined by using a training algorithm, the most popular of which is the back propagation (BP) algorithm. The BP algorithm randomly selects the initial weights, then compares the calculated output for a given observation with the expected output for that observation. The difference between the expected and calculated output values across all observations is summarized using the mean squared error. After all observations are presented to the network, the weights are modified according to a generalized delta rule (Rumelhart et al., 1986), so that the total error is distributed among the various nodes in the network. This process of feeding forward signals and back-propagating the errors is repeated iteratively (in some cases, many thousands of times) until the error stabilizes at a low level.

2.2. Land use change models

Models of land use change serve as useful tools for (1) exploring the various mechanisms by which land use change occurs and the social, economic, and spatial variables that drive it (Batty & Longley, 1994; deKoning, Verburg, Veldkamp, & Fresco, 1999), (2) projecting potential future environmental and economic impacts of land use change (Alig, 1986; Theobald, Miller, & Hobbes, 1997), and (3) evaluating the influence of alternative policies and management regimes on land use and development patterns (Bockstael, Costanza, Strand, Boyton, Bell, & Wagner, 1995). Although some approaches focus on modeling aggregate land use amounts within areal units, like counties (Alig & Healey, 1987), models that predict the spatial patterns of land use provide more information with which to evaluate the impacts of change. To address the multi-scale nature of land use change drivers, some models, like the CLUE model of de Koning et al. (1999), have used regional and global scale drivers to determine the aggregate amounts of change and geographic and landscape scale drivers to determine its pattern. This is the adopted approach for development of the Land Transformation Model. Artificial neural networks, as described in this paper, are used to determine the *location* of land use change using landscape scale variables given a certain *amount* of change determined by regional and global scale variables. In this paper, local population and land use change data are used as a simplified approach to determining the amount of change.

Theobald and Hobbs (1998) described two basic types of spatially explicit land use change models: regression-type models and spatial transition-based models. Although their analysis of the former type focused on the logistic regression to implement the regression based models, our use of ANNs clearly fits within this category of models. The goal of these models is to establish functional relationships between a set of spatial predictor variables that are used to predict the locations of change on the landscape. The variable values and actual instances of land use change are typically observed from historical data and used to establish functional relationships that can be used to extrapolate land use change probabilities into the future. The spatial transition models are an extension of the aspatial Markov technique and a form of stochastic cellular automata (CA; Theobald & Hobbs, 1998, p. 66). Cellular automata incorporate simple rules about spatial adjacency effects that govern system dynamics and give rise to emergent behaviors and patterns that are usually more complex than those generated by simple equilibrium models (Batty & Longley, 1994; Clarke, Gaydos, & Hoppen, 1997), CA models are quite valuable for describing system dynamics and behavior (Theobald & Hobbs, 1998), but the focus of this paper is on an application of the computationally simpler regression type model.

The value of the regression-type models is that the relative contribution of different variables for predicting a given land use change is easily obtained. Because of the spatial nature of many of the input variables, integration with GIS is essential. GIS allows users to manage and analyze spatially explicit data associated with the models. For example, GIS can aid modelers in building input variables for the models, identifying spatial heterogeneity or pattern in data (Openshaw & Clarke, 1996), quantifying observed and/or predicted temporal changes in spatial pattern (de Koning et al., 1999), and assessing factors that operate across a variety of scales (Qi & Wu, 1996). Most GIS-based models of land use change use data stored in the raster data structure (Clarke et al., 1997; Landis, 1994; Veldkamp & Fresco, 1996) because the structure simplifies the representation of space by breaking it into many units of equal size and shape. Further, remotely sensed data, which is inherently grid-based, is often used for model validation and calibration.

2.3. LTM

The LTM follows four sequential steps: (1) processing/coding of data to create spatial layers of predictor variables; (2) applying spatial rules that relate predictor variables to land use transitions for each location in an area; the resultant layers contain input variable values in grid format; (3) integrating all input grids using one of three techniques; and (4) temporally scaling the amount of transitions in the study area in order to create a time series of possible future land uses. The GIS portion of the LTM is encoded in ESRI's (2000) ArcView GIS 3.2 Avenue scripting language. A collection of routines written in C is used to process and analyze data.

In Step 1, *processing of spatial data*, inputs are generated from a series of base layers that are stored and managed within a GIS. These base layers represent land uses (such as agriculture parcels and urban areas) or features in the landscape (e.g. roads, rivers, lakeshores). Grid cells are coded to represent predictors as either binary (presence = 1 or absence = 0) or continuous variables depending on the type of attribute.

For Step 2, *applying spatial transition rules*, inputs are developed using a set of spatial transition rules that quantify the spatial effects that predictor cells have on land use transitions (Pijanowski et al., 2000, for details). We use four classes of transition rules: (1) neighborhoods or densities; (2) patch size; (3) site specific characteristics; and (4) distance from the location of a predictor cell. Neighborhood effects are based on the premise that the composition of surrounding cells has an effect on the tendency of a central cell to transition to another use. Patch sizes relate the variable values of all cells within a defined patch (e.g. parcel) to likelihood of land use transition. Site-specific characteristics are values assigned to a cell based on biophysical or social characteristics that are specific to each grid cell. An example of a site-specific characteristic is the location of quality views. The distance spatial transition rule relates the effect of the Euclidean distance between each cell and the closest predictor variable.

Certain locations are coded so that they do not undergo transitions. This is necessary for areas within which development is prohibited, such as public lands. We code cells with a “0” if a transition cannot occur; all other locations are assigned a “1”. All such layers are then multiplied together to generate one single layer of “exclusionary zones.”

Step 3, *integration of predictor variables*, one of three different integration methods are used: multi-criteria evaluation (MCE), ANNs, and logistic regression (LR). Each integration procedure requires a different type of data normalization. Although we only present information relevant to the ANN integration method here, Pijanowski et al. (2000) described the approach used with the MCE method. With all of the integration methods, the cell size (100×100 m in the present analysis) and analysis window are set to a fixed base layer. The output from this step is a map of “change likelihood values”, which specifies the relative likelihood of change for each cell based on the ANN result given for the cell’s aggregate value for change derived from the total of predictor variable values.

In Step 4, *temporal indexing*, the amount of land that is expected to transition to urban over a given time period is determined using a “principle index driver” or PID (Pijanowski et al., 2000). This paper describes two methods for calculating the PID. The first method involves simply calculating the amount of area that underwent transition to urban use based on analysis of historical land use data. Future projections can be made for each 10-year time step by assuming that the same number of cells will transition to urban in each 10-year period as in the observed 10-year period. However, we recommend against using this method for projection unless future demographic projection for the study area are not available. Our use of this approach is used here merely to test the LTM’s ability to predict the spatial patterns of change.

The second approach to calculating the PID is to base its value on the population growth (i.e. number of people) over a time interval (i.e. number of years) for a region. This might include: (1) applying population growth statistics to subregions and forecasting subregions separately, output of which is integrated within a GIS; or (2) adding all population projection estimates for subregions (such as municipalities, counties or states) and using this information to forecast the entire region at the same time. By using population and land use data for an initial time period, per capita requirements for land (i.e. the number of hectares of developed land per person) are calculated such that the total amount of new urban land at a later time is:

$$U(t) = \left(\frac{dP}{dt}\right) \times A(t) \tag{1}$$

where U is the amount of new urban land required in the time interval t, dP/dt is the number of new people in any given area in a given time interval and A is the per capita requirements for urban land. In this paper, we illustrate the first method of using Eq. (1); namely, using county population projections to scale our land use changes into the future.

The PID is used to determine the number of cells that need to transition to urban. Projections are made by selecting the appropriate number of cells in priority order, which is based on the change likelihood value for that cell (determined in Step 3). The lowest change likelihood value required to transition the appropriate number of cells is referred to as the critical threshold value (CTV).

3. Methods

3.1. Study area and data sources

Michigan's Grand Traverse Bay Watershed (GTBW) was selected as the test site for this project. The GTBW, located in the northwestern portion of Michigan's Lower Peninsula (Fig. 2), is one of the most rapid population growth and land use change regions in the USA (Vesterby & Heimlich, 1991). The modeling extent is the six Michigan counties that comprise the watershed: Grand Traverse, Kalkaska, Antrim, Leelanau, Charlevoix and Benzie. From 1970 to 1997, resident population in the watershed nearly doubled (US Census Bureau, 1998). Traverse City, with a resident population of approximately 18,000 (oftentimes having a seasonal tourist population exceeding 500,000) is the largest city in the watershed. Land use in the GTBW is predominantly forest (49%) and agriculture (20%). Much of the forested portions of this watershed are managed within the Pere Marquette State Forest, which encompasses most of the upland reaches of the Boardman River tributaries (Fig. 2). Urban land use comprises about 6% of the total area of the watershed. The other main land covers are open herbaceous/shrub/grasslands (15%), water (9%), and wetlands (1%).

The MiRIS (Michigan Resource Information System) database (developed around 1980 from 1:24,000 aerial photography) was used as the source of land use data in this project. For the present application of the LTM, all land use codes were reclassed to Anderson Level I land uses (Anderson, Hardy, Roach, & Witmer, 1976). All land use files were then rasterized at a resolution of 100×100 m. MiRIS line files, digitized from 1:24,000 scale topographic maps, were integrated with our database to represent the transportation network and locations of rivers, lakes, as well as the Grand Traverse Bay shoreline to provide the appropriate inputs to the GIS-based LTM. USGS digital elevation models (DEM), corresponding to the 7.5-min topographic quandrangles covering the watershed, were then resampled to 100×100 m cell sizes. Boundaries for public lands were digitized from county digital DRG databases. Locations of recreational sites, such as golf courses, casinos, ski lodges, and marinas were obtained from published county road maps and stored as point coverages.

3.2. Implementing the LTM

For this project, the LTM was applied to two groups of runs. The first run (referred to as the control run) was used to project the pattern of urban land development in

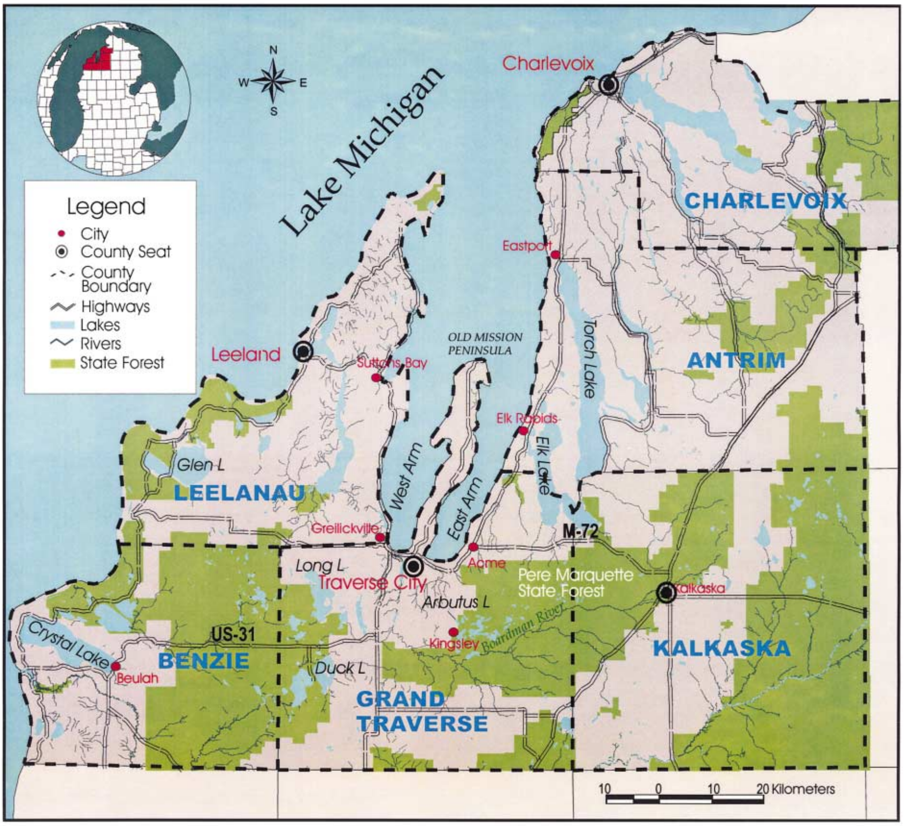

Fig. 2. Map of Michigan's Grand Traverse Bay Watershed counties and important locations within the watershed.

1990 using an ANN trained on the actual changes between 1980 and 1990 solely for Grand Traverse County. The second run is an extension of the same ANN in order to project 1990 urban land development across all six counties in the watershed. The first three steps of LTM implementation (described earlier) were identical for the two runs; it is the fourth step that distinguishes them. The land use data-based approach to indexing change was used for the control run. Since 1990 land use data were not available for the entire watershed, the population-based approach for implementing projections was used in the second run.

3.3. GIS-based predictor variables

Ten predictor variables and the exclusion zones were compiled in Arc/Info Grid format (Table 1; Fig. 3) using the LTM GIS Avenue interface. The agricultural density variable represents the amount of agriculture, from the 1980 land use data-base, within a 1 km radius surrounding each cell. This variable describes the degree

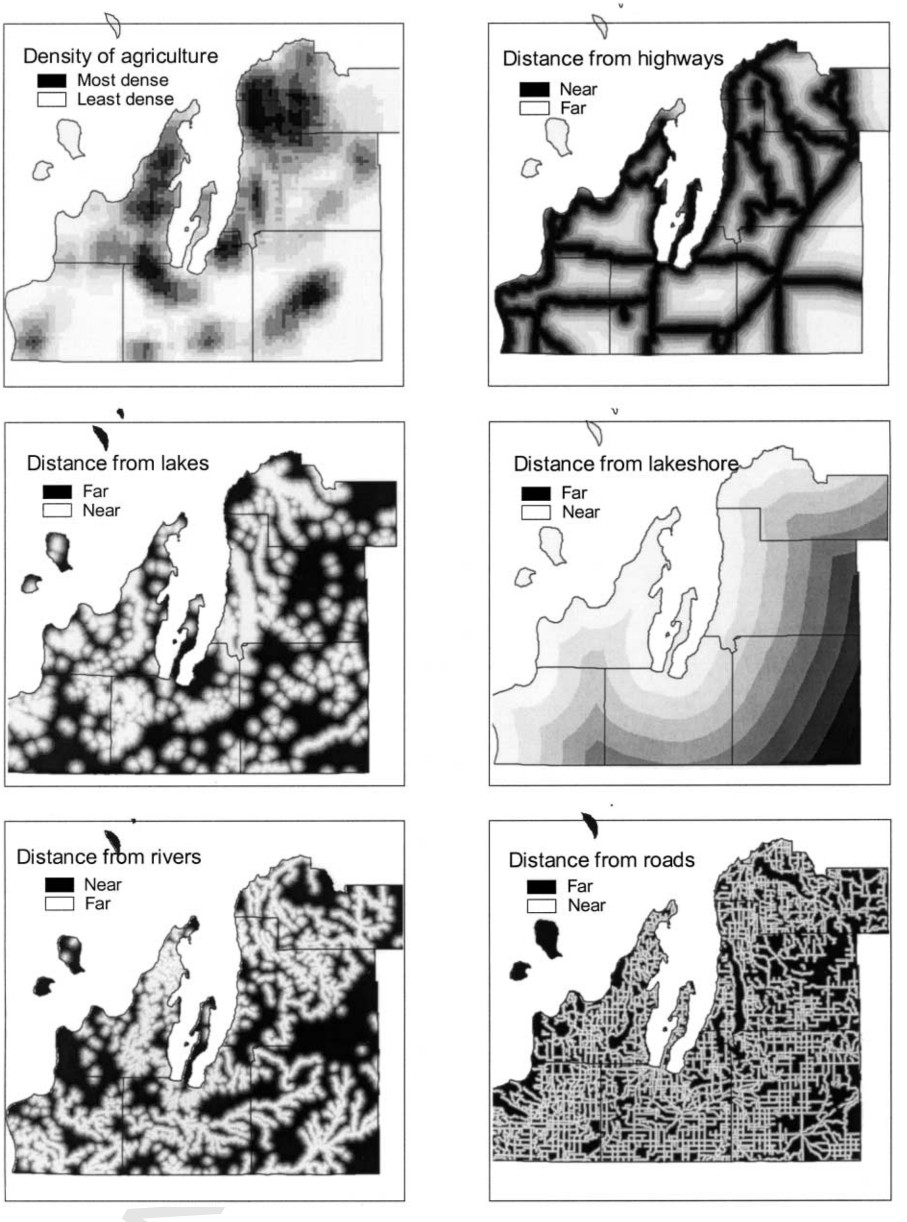

Fig. 3. Maps of the 10-predictor variables used for the training exercise. Also shown is the exclusion zone.

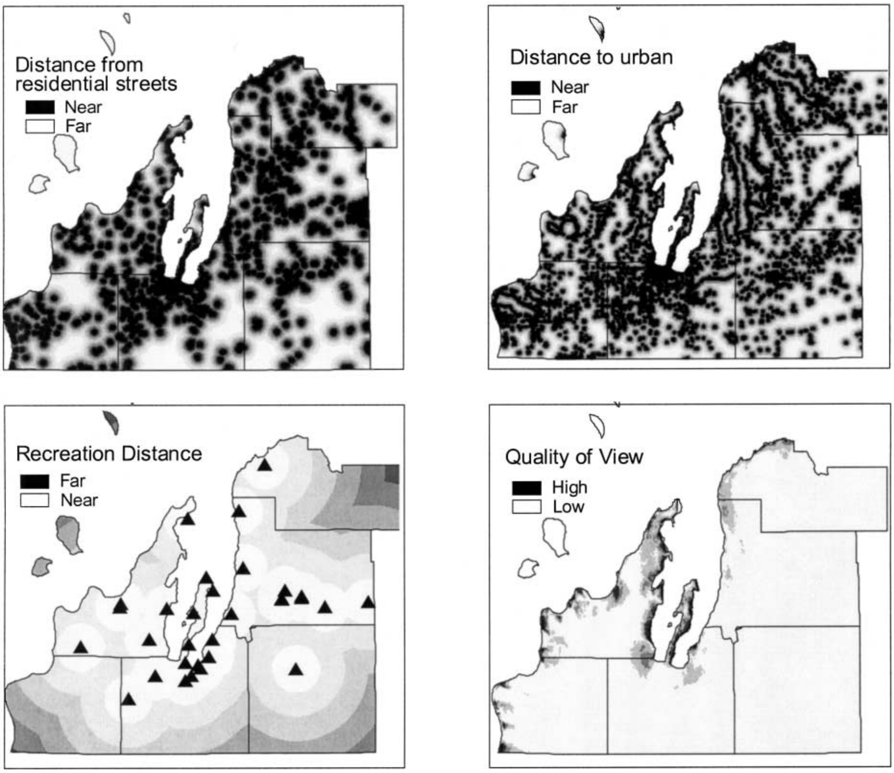

Fig. 3. (*Continued*).

to which the landscape is dominated by agricultural land uses. Agriculture can be seen as an amenity on the landscape that attracts development. However, it is possible that agricultural land use can serve as an impediment to development, especially in this area where agricultural activities are specialized (vineyards and other fruit production) and profitable.

For the variables of county roads distance, highway distance, shoreline distance, inland lake distance, and river distance, the minimum Euclidean distance to each feature was calculated. These features serve to either improve the access of the site to larger urban areas (i.e. county roads and highways) or to increase the amenity value of a site (i.e. shoreline, inland lakes, rivers). The urban distance variable was the minimum distance of each cell to an urban cell, from the 1980 Grand Traverse County land use database. Since access to urban services affects development patterns, it is expected that sites nearer to existing urban land uses would be more likely to develop. Distance from recreation sites, which were coded as point coverages, was also used as a predictor variable. For view quality, the height above lake level and the distance from the lakeshore were calculated for every location in the watershed. The sine of the angle of incidence of a line-of-sight from each location in the watershed to the lake was then calculated and used as a surrogate for quality of

Table 1
Summary of the LTM predictor variables used for the Grand Traverse Bay Watershed application

Variable name	Description
Agricultural density	Amount of agricultural land that surrounds each location within a 1 km radius
Highway distance	Distance to nearest highway
Inland lake distance	Distance to nearest inland lake (not including Lake Michigan or Grand Traverse Bay)
Lakeshore distance	Distance from the lakeshore along Lake Michigan and Grand Traverse Bay
River distance	Distance from rivers and streams; not including intermittent streams
County road distance	Distance to nearest county road
Residential street distance	Distance to nearest residential street
Urban distance	Distance to 1980 urban use
Recreation distance	Distance from a major recreational site, such as golf courses, casinos, ski slopes and marinas
Quality views	A quality of view estimate derived from a combination of proximity to the bay or Lake Michigan and elevation above lake level

view. Larger angles represent locations that are highly elevated above the bay and are close. These locations are hypothesized to be in great demand for residential use.

The exclusion zone for this execution of the model was composed of the following GIS layers: areas that were urban in 1980; locations of open water; locations of wetlands; locations of public land (e.g. local, state and federal parks) and locations of current transportation corridors.

3.3.1. ANN-based integration

In order to develop a network with adequate predictive capacity, it was necessary to train and test the ANN with different input data (Skapura, 1996). Training involves presenting input values and adjusting the weights applied at each node according to the learning algorithm (e.g. back-propagation). Testing presents a separate data set to the trained network independently to calculate the error rate. For these reasons, ANNs were applied to the prediction of land use change in four phases: (1) design of the network and of inputs from historical data; (2) network training using a subset of inputs; (3) testing of the neural network using the full data set of the inputs; and (4) using the information from the neural network to forecast changes.

The ANN in this project is a feed forward network with one input layer, one hidden layer and one output layer. The Simple Backpropagation Algorithm was used as the learning process. SNNS (Stuttgart's Neural Network Simulator) version 4.2 was used for the design, training and prediction of the ANN (Zell et al., 1996). The neural network was designed to have a flexible number of inputs depending on the number of predictor variables presented to it, an equal number of hidden units as input units and a single output. All input grids, which existed in Arc/Info Grid format, were then normalized to a range from 0.0 to 1.0 and converted into ASCII representations (called a pattern file), which is the required format for

SNNS. The pattern file contained information from the 10 final input grids and one output file so that each line in the pattern file corresponded to one location. The output of the ANN represents the likelihood of a non-urban cell changing to urban. A change likelihood of zero indicates "no readiness" to change whereas a change likelihood of one indicates the "highest readiness" to change to urban.

To avoid over-training of the network (Skapura, 1996), the neural network was trained with a partial set of data by providing it with data from every other cell in the county. To further reduce over-fitting, the cells in a cycle were first presented to the network in random order (Zell et al., 1996). A cycle is defined as one complete presentation of all training cells to the network. The network was trained with the training data and the overall mean squared error generated by SNNS and each cycle was stored in a file for analysis. Based on these results, it was concluded that about 4000 cycles were adequate to stabilize the error level to a minimum value. Thus, each model run was set for 5000 cycles, which meant that the entire pattern file would be presented to the network 5000 times.

The ANN was tested as follows: First, the network files generated from the training exercise were applied to a pattern file that contained all of the cells in the county (except those within the exclusion zone). SNNS used the pattern file and the network file to generate an output file of activation values, which is called a "result file". The result file contains values ranged from 0.0 (no likelihood of changing to urban) to 1.0 (highest likelihood of changing to urban). Next, the GIS was used to determine that 2073 cells transitioned into urban in Grand Traverse County during the 10-year period 1980–1990 (i.e. the first approach to PID calculation). Thus, 2073 cells were selected from the result file that had the greatest change likelihood values; these cells were then classification as new urban. Since there were many cells that contained the same change likelihood values, selecting the change likelihood values at the 2073 cell threshold required a random selection from among the remaining cells to transition from the population of cells that had the same change likelihood value.

Testing was completed by comparing those cells that were observed to transition, based on the data, with those cells with the highest likelihood of transition, based on the model. The following metric was then used to assess the performance of the model:

$$\frac{\text{No. of cells predicted to change}}{\text{No. of cells that transitioned (i.e. 2073)}} \tag{2}$$

3.3.2. *Evaluating the model*

Two questions were raised in regard to the predictive ability of the ANN version of the LTM: (1) does the model predict the locations of urban development accurately? and (2) what predictor variables were found most influential in the model's ability to identify urban land use change?

To address the first question, model results were overlaid from the control run at the county level with observed changes in land use from 1980 to 1990. A layer was created with the following codes:

0 = no observed change and no predicted change;
1 = observed change but not predicted by the model;
2 = no observed change but change predicted by the model;
3 = observed change and predicted change.

The resulting display (Fig. 4) was used throughout the model building process in order to visualize the spatial distribution of the two types of error (codes 1 and 2) for Grand Traverse County.

Since errors tended to be spatially autocorrelated (i.e. errors of the same type tend to occur in multi-pixel patches), a modified version of Costanza's (1989) scalable window goodness of fit algorithm was applied to assess predictability across spatial scales. Our approach entailed creating window sizes of 3×3, 5×5, 7×7 and so on, across the entire county grid, and searching for cells coded as "1" and "2" within the window. For each pair of "1" and a "2" that were found in the window, the number of correct predictions (starting with the number of 3s for the entire region) were incremented by one. Once they were added to the counter, each "1" and "2" were removed from the database so that they would not be double counted as the window frame shifted. The window was shifted by one cell at a time so that the entire study area (the county) was covered. The proportion of correct predictions (No. predictions/No. cells transitioning in the observed database) was then plotted against window size.

The influence of each predictor variable on model performance (question 2) was quantified by creating 10 alternative versions of the model, each version created by dropping one variable out to create nine-predictor variable models, and repeating the training exercise described earlier. New network files were created for each model and then testing was repeated by presenting a pattern file with nine predictor variables and the network file created from the training to generate result files used to determine transition locations. The scalable window index was then calculated for each nine-variable model. The difference in goodness of fit between the 10-variable model and the 10 versions of the nine-variable model were then plotted in order to assess the relative power of each predictor variable has on model performance.

3.3.3. *Applying the model to the watershed*

The ANN generated from the county control run was used to scale-up our forecasts to the entire watershed. In this run, 1980 MiRIS land use and line files (e.g. roads and rivers) were used to create predictor variable inputs for the entire watershed region. Eq. (1) was used to determine the number of cells that would transition across the entire watershed for four decadal time steps, to the years 1990, 2000, 2010, and 2020. Projected population estimates for these time periods were obtained from the US Census Bureau. Table 2 contains a summary of the number of urban cells per person for each county and the number of future urban cells required given the projected population growth for each county. Note that Grand Traverse County has a population density that is twice that of the other counties.

Between 1980 and 1990, 9374 people were added to Grand Traverse County. The number of cells that transitioned were 2073, thus each person added between 1980

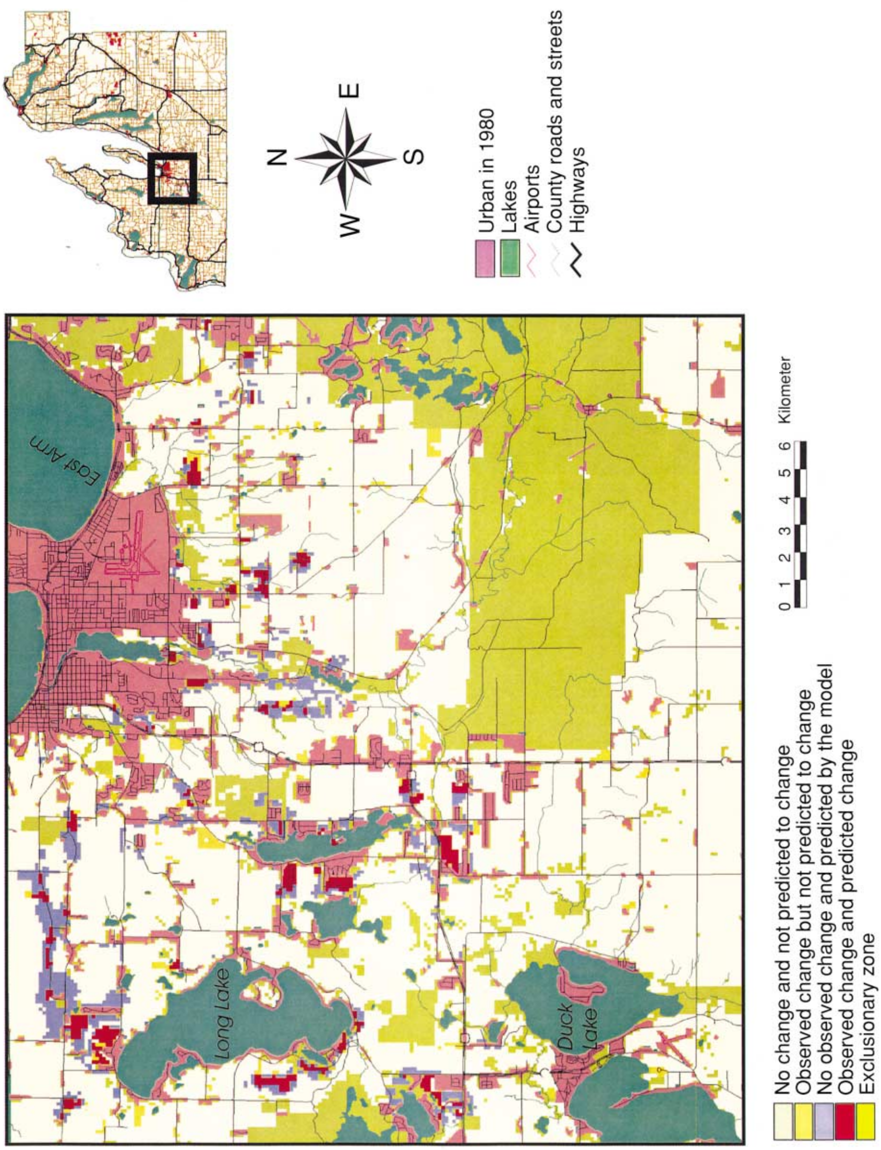

Fig. 4. An overlay of model predictions and observed changes in an area southwest of Traverse City in Grand Traverse County.

and 1990 required 4.52 cells (or 0.452 ha) of urban land. Over 92,000 people will be added to the six county area between 1980 and 2020, (Table 2) representing an increase of 73% in total population. Percentage decadal growth for each county is expected to average around 13%, with many counties potentially increasing total population by nearly 20% in some 10-year periods (Table 2).

4. Results and discussion

4.1. Model performance

Of the 63,744 cells that could have undergone transition in Grand Traverse County the ANN estimated that, 56,762 (or 89.0%) of these had change likelihood values of 0.0 (or no likelihood that these cells could transition to urban) while 208 had likelihood values of 1.0. Of the 2073 cells with the highest change likelihood values, 581 (28.0%) had values that were greater than or equal to 0.90. The critical threshold value, which is the lowest change likelihood value selected for 2073 cells to transition during the 10-year period, was 0.28. These results suggest that the model does well at separating change and no change.

Only 3.25% of all non-exclusionary areas changed to urban in the observed databases; thus a great majority of the area did not undergo urbanization. However, our results show that the ANN has a more difficult time learning the characteristics that lead to change than those that lead to no change, given that a critical threshold value of 0.28 was used to transition enough cells to urban during this 10-year time step. Increasing the critical threshold may require: (1) more information about change, and/or (2) knowledge about specific non-spatial aspects of change that were not provided as input to the model. More information about change could mean that it may be necessary to model a large enough area so that an adequate number of cells that undergo transition are provided to the neural net. Alternatively, longer time intervals between the land use changes contained in databases (e.g. larger than 5–10 years) may be needed to supply the neural network with enough locations that changed. Learning about change will never be 100% complete given that urbanization has elements of both unpredictability and likely has aspects that are non-spatial. For example, demographic factors, such as the age of a farmer, could be very important to what locations transition from agriculture to urban. In the USA, many retiring farmers are not passing along their farms to their children. More research is needed to assess the importance of both of these elements to the predictability of this and other regression type models that use GIS inputs.

For the execution of the model version presented here, the proportion of transitioned cells (using Eq. 2) that were predicted correctly was 0.46 (i.e. 941/2073). This represents the accuracy of the model at a 100×100 m cell size. The proportion of correct predictions using the goodness of fit metric versus window size is shown in Fig. 5. Note that within a window size of 1 km (i.e. 10 cells), the improvement of the model increases to a 0.65 goodness of fit. As expected, the goodness of fit approaches

Fig. 5. Goodness of fit metric plotted across window sizes. Note that because the cell size is 100 m, a window size of 10 corresponds to a 1 km window.

1.0 with very large window sizes. These values suggest that the ANN is learning well the conditions that give rise to urban development.

4.2. Relative effect of predictor variables

We compared the predictive ability of the 10 versions of the reduced-variable model with that with the full complement of predictor variables. We subtracted the goodness of fit for the full model from that for each of the nine-variable models across all of the window sizes. The resultant value, which could be positive or negative, represents the relative effect, or contribution, of each predictor variable on the model performance (Fig. 6). Note that all predictor variables have a positive influence on the goodness of fit for small window sizes. The predictor variable with the most influence at small window sizes (<10) is the location of quality of views. For small window sizes (<5) the rank order of predictor variables according their influence on the model performance, after quality of views, (Fig. 6) is: distance to previous urban, distance to residential streets, distance to highways, distance to recreational facilities, distance to rivers, density of agriculture, distance to inland lakes, distance to county roads, and distance from lakeshore.

The relative influence of the predictor variables changes over multiple scales (Fig. 6), and these changes can be used to interpret the scale over which the variables influence land use change. The differential scale of influence of the variables was evident in the variables describing road access. The influences of both county roads and highways increased with distance to level off at about 2.5 km, whereas the influence of residential streets was highest at the shortest distances and dropped to about zero influence at a scale of about 3.5 km. County roads and highways relate

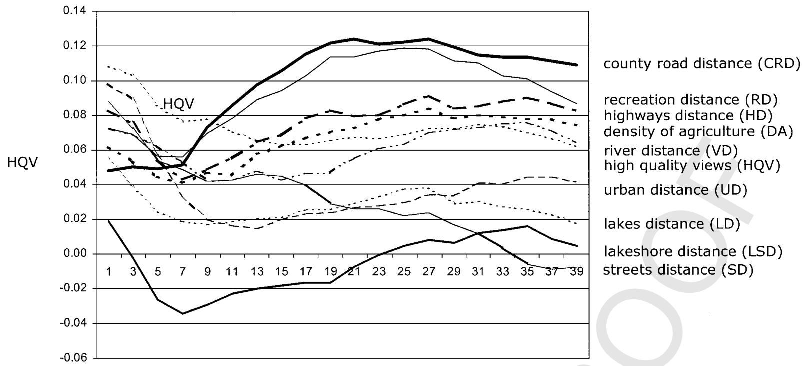

Fig. 6. Goodness of fit metric difference between each of the nine-variable models and the model containing the full compliment of predictor variables.

more to the regional accessibility of a location (i.e. its situation) than to its local character (i.e. site). Access to residential streets describes the local character of the place. Similarly, the influence of the view variable decreases in importance with increasing window size. Because the value of a striking view declines rapidly as one leaves a given property, the variable tended to have a more localized influence on likelihood of urban development.,

Interestingly, the curves in Fig. 6 appear to adhere to two major inflection points. At the first, a window size of about seven (i.e. 700 m), nearly all of the predictor variables change direction from a negative slope (decreasing in their relative affect on model performance) to a positive slope. Also at this window size, ranking of relative influence of variables is very volatile. This scale appears to reflect a natural break in the scales of influence, from local scale influence at smaller window sizes to more regional scale influences at larger window sizes. The second inflection point occurs at a window size of about 19 (i.e. 1.9 km) and reflects a stabilization of the relative effects of each variable on model performance at large window sizes. The volatility of influence is diminished at larger scales, probably because the changes in overall fit of the model (Fig. 5) are much smaller at larger scales.

4.3. Watershed-scale land use projections

Four projections for the six county area were produced for the years 1990, 2000, 2010, and 2020 (Fig. 7) using the PID values derived from US Census Projections (Table 2). Fig. 7 shows the results of this regional forecast of land use changes. A comparison of 1980 and 2020 land uses shows that most of the new urban development is predicted to occur along Crystal Lake (Benzie County), along the US-31 east-west thoroughfare between Beulah (Benzie County) and Duck Lake (Grand Traverse County), around Glen Lake (Leelanau County), along the Five Mile Road corridor especially around Lake Arbutus which connects the cities of Acme and Kingsley (Grand Traverse County), areas around the city of Kalkaska (Kalkaska County), the cities of Elk Rapids and Eastport in Antrim County, and around the

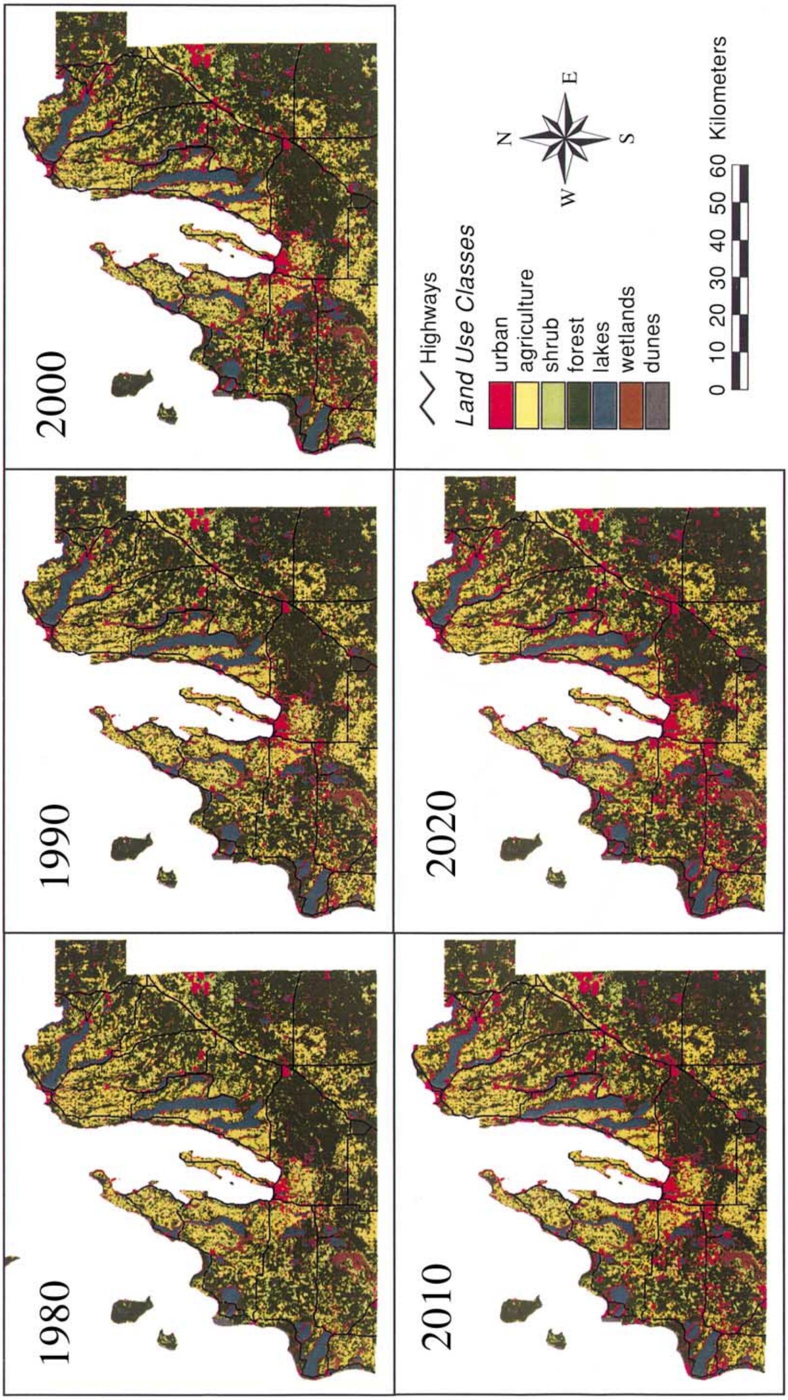

Fig. 7. An illustration of scaling up the results to the other five counties using the PID and the neural network file generated from the training of Grand Traverse County.

city of Charlevoix in Charlevoix County. These model projections are, anecdotally, consistent with current development patterns observed in the watershed. Development is also projected to occur along the western shore of the West Arm of Grand Traverse Bay from Greilickville to just south of Suttons Bay (Leelanau County). This area also contained the largest area of quality views in the watershed. These projections illustrate how the ANN could be trained on relationships between urbanization and all of the predictor variables that occurred in Grand Traverse County and, through our approach, applied to the same predictor variables scaled to a larger region to provide reasonable results for these five other counties in the watershed.

5. Conclusions

The Land Transformation Model presented in this paper examines the relationship between 10 predictor variables and urbanization. The model performs with a relatively high predictive ability (46%) at a resolution of 100×100 m. Using a scalable window approach, we show that model prediction increased substantially when results are aggregated over larger window sizes such that, at a resolution of 1 km, the model predicted changes to urban which were correct 65% of the time. By developing 10 versions of the LTM, each with one of the variables removed, we assessed the relative contributions of each variable on model performance. The location of high quality views was the best predictor of new urban during the 1980–1990 period in Grand Traverse County.

Using the ANN pattern file generated for the watershed, we applied the network file created from the control run to create a file with changing likelihood values for each location in the entire six county area. The forecasts were indexed to population projections, using each county's population projection and per capita urban use requirements. The forecasts of new urban growth for the counties not used to train the model appeared reasonable and were concentrated in tourist towns near inland lakes or along the lakeshore.

We made several assumptions in order to keep the model simple. First, we assumed that the pattern of each predictor variable remained constant beyond 1990. For example, the location of roads and highways are likely to change (e.g. new roads will be built) and they may respond to changes in land use. The model can reflect changes in transportation if updated layers are available and are added between time steps and then applied to the same neural network file. Second, spatial rules used to build the interactions between the predictor cells and potential locations for transition are assumed to be correct and remain constant over time. It is possible for people to respond to the density of inland lakes (one type of spatial rule) rather than distance to any inland lake (an alternative spatial rule). Third, the neural network itself was assumed to remain constant over time. Thus, the relative affect of each predictor variable is assumed to be stable. Finally, the amount of urban per capita undergoing a transition is assumed to be fixed over time. Given the availability of data (e.g. new roads, more temporal information about land use change and population estimates), it is possible to relax many of these assumptions in order to examine the potential effect each of these assumptions have on the performance of model forecasts.

6. Uncited references

Environmental Systems Research Institute, 2000; Turner, 1987; Turner, 1988; Turner, 1989; Yin and Xu, 1991

Acknowledgements

This work was partially supported through grants from NASA's Land-Cover and Land-Use Change Program (NAG5-6042), from NASA's grant NAG13-99002 to the University of Minnesota, University of Wisconsin and Michigan State University entitled "An Upper Great Lakes Regional Environmental Science Applications Center (RESAC)" and from the Michigan Great Lakes Protection Fund administered by the Office of the Great Lakes, Michigan Department of Environmental Quality. All responsibility for errors in the execution of the research lies with the authors.

References

Alig, R. J. (1986). Econometric analysis of the factors influencing forest acreage trends in the southeast. *Forest Science*, *32*, 119–134.

Alig, R. J., & Healy, R. G. (1987). Urban and built-up land area changes in the United States: an empirical investigation of determinants. *Land Economics*, *63*(3), 216–226.

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensor data*. US Geological Survey, Professional Paper 964, p. 28, Reston, VA.

Atkinson, P., & Tatnall, A. (1997). Neural networks in remote sensing. *International Journal of Remote Sensing*, *18*(4), 699–709.

Babaian, R., Miyashita, H., Evans, R., Eshenbach, A., & Ramimrez, E. (1997). Early detection program for prostate cancer: results and identification of high-risk patient population. *Urology*, *37*(3), 193–197.

Batty, M., & Longley, (1994). Urban modeling in computer-graphic and geographic information system environments. *Environment and Planning B*, *19*, 663–688.

Bingham, G., Bishop, R., Brody, M., Bromley, D., Clark, E., Cooper, W., Costanza, R., Hale, T., Hayden, G., Kellert, S., Norgaard, R., Norton, B., Payne, J., Russell, C., & Suter II, G. (1995). Issues in ecosystem valuation: improving information for decision making. *Ecological Economics*, *14*, 73–90.

Bockstael, N., Costanza, R., Strand, I., Boyton, W., Bell, K., & Wagner, L. (1995). Ecological economic modeling and valuation of ecosystems. *Ecological Economics*, *14*, 143–159.

Boutt, D. F., Hyndman, D. W., Pijanowski, B. C., & Long, D. T. (2001). Identifying potential land use-derived solute sources to stream baseflow using ground water models and GIS. *Groundwater*, *39*(1), 24–34.

Brown, D. G., Duh, J. D., & Drzyzga, S. (2000). Estimating error in an analysis of forest fragmentation change using North American Landscape Characterization (NALC) Data. *Remote Sensing of Environment*, *71*, 106–117.

Brown, D. G., Lusch, D. P., & Duda, K. A. (1998). Supervised classification of glaciated landscape types using digital elevation data. *Geomorphology*, *21*(3–4), 233–250.

Brown, D. G., Pijanowski, B. C., & Duh, J. D. (2001). Modeling the relationships between land-use and land-cover on private lands in the Upper Midwest, USA. *Journal of Environmental Management*, *59*, 247–263.

Burchell, R. W. (1996). Economic and fiscal impacts of alternative land-use patterns. In *The land use decision making process: its role in a sustainable future for Michigan. January 9–10, 1996*. East Lansing, Michigan.

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

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

Costanza, R., Kemp, W. M., & Boynton, W. R. (1993). Predictability, scale and biodiversity in coastal and estuarine ecosystems: implications for management. *Ambio*, *22*, 88–96.

Dale, V., O'Neill, R., Pedlowski, M., & Southworth, F. (1993). Causes and effects of land use change in Central Rondonia, Brazil. *Photogrammetric Engineering and Remote Sensing*, *59*(6), 97–1005.

deKoning, G. H. J., Verburg, P. H., Veldkamp, A., & Fresco, L. O. (1999). Multi-scale modelling of land use change dynamics in Ecuador. *Agricultural Systems*, *61*, 77–93.

Drummond, S., Joshi, A., & Sudduth, K. (1998). Application of neural networks: precision farming. *IEEE Transactions on Neural Networks*, 211–215.

Environmental Systems Research Institute. (1999). *Cell-based modeling with GRID*. Atlanta, Georgia: Environmental System Research Institute.

Fishman, M., Barr, Dean S., & Loick, W. J. (1991). Using neural nets in market analysis. *Technical Analysis of Stocks & Commodities*, *4*, 18–21.

Flintrop, C., Hohlmann, B., Jasper, T., Korte, C., Podlaha, O., Scheele, S., & Veizer, J. (1996). Anatomy of pollution: streams of North Rhine-Wesphalia, Germany. *American Journal of Science*, *296*(296), 58–98.

Fukushima, K., Miyake, S., & Takayuki (1983). Neocognitron: a neural network model for a mechanism of visual pattern recognition. *IEEE Transactions on Systems, Man, and Cybernetics, SMC*, *13*(5), 826–834.

Houghton, R. A. (1994). The world-wide extent of land-use change. *Bioscience*, *44*(5), 305–313.

Landis, J. D. (1994). The California urban futures model: a new-generation of metropolitan simulation-models. *Environment and Planning B*, *21*, 399–420.

Medley, K., Okey, B. W., Barrett, G. W., Lucas, M. F., & Renwick, W. H. (1995). Landscape change with agricultural intensification in a rural watershed, southwestern Ohio. *USA Landscape Ecology*, *10*(3), 161–176.

Openshaw, S., & Clarke, G. (1996). Developing spatial analysis functions relevant to GIS environments. In M. Fischer, H. J. Scholten, & D. Unwin (Eds.), *Spatial analytical perspectives on GIS* (pp. 21–37). London: Tayler and Francis.

Pijanowski, B. C., Gage, S. H., Long, D. T., & Cooper, W. C. (2000). A land transformation model: integrating policy, socioeconomics and environmental drivers using a geographic information system; In L. Harris, & J. Sanderson (Eds.), *Landscape ecology: a top down approach*.

Pijanowski, B. C., Machemer, T., Gage, S., Long, D., Cooper, W., & Edens, T. (1995). *A land transformation model: integration of policy, socioeconomics and ecological succession to examine pollution patterns in watershed* (Report to the Environmental Protection Agency, Research Triangle Park, North Carolina, pp. 72).

Qi, Y., & Wu, J. (1996). Effects of changing spatial resolution on the results of landscape pattern analysis using autocorrelation indices. *Landscape Ecology*, *11*(1), 39–49.

Richards, J. (1990). Land transformation. In: B. L. Turner (Ed.), *Earth as transformed by human action: global and regional changes in the biosphere over the past 300 years* (pp. 713). Cambridge University Press with Clark University Press.

Ritter, N., Logan, T., & Bryant, N. (1988). Integration of neural network technologies with geographic information systems. *Proceedings of the GIS symposium: integrating technology and geoscience applications* (pp. 102–103). Denver, Colorado.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*, 386–408.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning internal representations by error propagation. In D. E. Rumelhart, & J. L. McClelland (Eds.), *Parallel distributed processing: explorations in the microstructures of cognition* (Vol. 1; pp. 318–362). Cambridge: MIT Press.

Skapura, D. (1996). *Building neural networks*. New York: ACM Press.

Steiner, F. R., & Osterman, D. A. (1988). Landscape planning: a working method applied to a case study of soil conservation. *Landscape Ecology*, *1*(4), 213–226.

Theobald, D. M., & Hobbs, N. T. (1998). Forecasting rural land-use change: a comparison of regression- and spatial transition-based models. *Geographical and Environmental Modelling*, *2*(1), 65–82.

Theobald, D. M., Miller, J. M., & Hobbs, N. T. (1997). Estimating the cumulative effects of development on wildlife habitat. *Landscape and Urban Planning*, *39*, 25–36.

Turner, M. G. (1987). Spatial simulation of landscape changes in Georgia: a comparison of 3 transition models. *Landscape Ecology*, *1*, 29–36.

Turner, M. G. (1988). Changes in landscape patterns in Georgia, USA. *Landscape Ecology*, *1*, 241–251.

Turner, M. G. (1989). Landscape changes in nine rural counties in Georgia. *Photogrammetric Engineering and Remote Sensing*, *56*(3), 379–386.

Veldkamp, A., & Fresco, L. O. (1996). CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. *J. D.Ecological Modeling*, *91*, 231–248.

Vesterby, M., & Heimlich, R. (1991). Land use and demographic change: results from fast-growing counties. *Land Economics*, *67*(3), 279–291.

Wilder, M. G. (1985). Site and situation determinations of land use change: an empirical example. *Economic Geography*, *61*, 332–344.

Yin, Y., & Xu, X. (1991). Applying neural net technology for multi-objective land use planning. *Journal of Environmental Management*, *32*, 349–356.

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