Journal of Land Use Science
Vol. 1, Nos. 2–4, June–December 2006, 83–108

Research Article

Modelling urbanization patterns in two diverse regions of the world

B. C. PIJANOWSKI*†, K. T. ALEXANDRIDIS†‡ and D. MÜLLER§
†Human-Environment Modeling and Analysis Laboratory (HEMA), Department of Forestry and Natural Resources, Purdue University, 195 Marsteller Street, West Lafayette, Indiana, USA 47907-2033
‡CSIRO Sustainable Ecosystems, Davies Laboratory, University Drive, Townsville, QLD 4810, Australia
§Research Group on Postsocialist Land Relations, Humboldt University Berlin, Luisenstr. 56, 10099 Berlin, Germany

(*Received 13 January 2006; in final form 12 July 2006*)

We present work applying a similarly parameterized urbanization model to two diverse regions of the world, one in the USA and another in Albania. Eight calibration metrics are used to estimate model goodness of fit: four location-based measures (e.g. kappa), and four patch metrics based on patch size, shape and configuration. We conclude that if we use location goodness of fit estimates, the model fits observed data very well for most simulations. The model fit to data better in Albania than in the USA probably owing to top-down land ownership policies occurring in Albania and owing to the fact that commonly used land use change model drivers, such as distance to road, are not likely to capture individual behaviours that are important in the USA. Patch metrics provided additional information on model fit to observed data, and we suggest that, in some circumstances, patch metrics may be more useful than location metrics to calibrate a land use change model.

Keywords: Urbanization; Neural networks; Calibration metrics; Landscape patterns metrics

1. Introduction

Understanding how well a land use change model performs in diverse landscapes requires a characterization of the various spatial and temporal patterns of land use change occurring in a region. This is especially critical if model results are used for planning and management. For example, in many areas of the USA, urbanization is a rapid land use change process that produces different patterns depending on the proximity to large urban cities across the landscape (Wu 2004). Rates of change typically are greater at the urban–rural fringe, occurring in relatively large patches, but dispersed urbanization also occurs in the ex-urban environment (Daniels 1999). This latter pattern is sometimes referred to as ex-urban development or 'rural sprawl' (Carrion-Flores and Irwin 2004, Cavailhes *et al.* 2004) where housing development

*Corresponding author. Email: bpijanow@purdue.edu

Journal of Land Use Science
ISSN 1747-423X print/ISSN 1747-4248 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/17474230601058310

occurs at lower densities than at the urban fringe. It is hypothesized that the rural area is an attractive place to live compared with urban areas as land values are less, natural amenities are greater and transportation to the urban cores is affordable. In other areas of the world, however, the spatial patterns of urban development differ from that in the USA owing to differences in economics, political structure, and land ownership policies. For example, urbanization patterns in Albania after the fall of socialism are largely attributed to rural–urban migration movements. Urban growth was to a large extent unregulated which induced a construction boom after the privatization of land and buildings (Felstehausen 1999). The migration movements also led to large squatter settlements on empty urban spaces and in the peripheries following transportation routes and connection possibilities with electricity and water networks (Felstehausen 1999). The lack of adequate transport infrastructure does not allow for travelling large distances between the place of employment in urban areas and living in rural areas as occurs in the USA. As a result of these drivers of urbanization, urban expansion in Albania occurs in a more clumped spatial arrangement, similar to patterns in developing countries but different than in the USA.

The principal objective of the present study is to determine how well the same model applied to different areas of the world undergoing diverse patterns of urbanization fit to observed data. More specifically, we address the question: are fragmented urbanizing areas more challenging to model than clumped urbanization patterns? Few studies have compared the goodness-of-fit metrics of the same model applied to different areas of the world where land use change patterns differ. Such comparisons can help to address several important research questions. First, can the same drivers of land use change produce different urbanization patterns, such as those found in complex fragmented patterns of change as well as patterns generated from simple, radial growth from current urban centres? Second, in areas where land use change occurs in fragmented patterns with large and small patches of change and in varying amounts of change across the landscape, does the model satisfactorily predict these large and small patches? Similarly, does it predict change satisfactorily in areas where change is relatively rare? Third, do various model calibration metrics apply equally well in fragmented versus clumped urbanizing areas, and in particular, how important are common calibration metrics such as the kappa statistic (see below) in assessing the accuracy of models producing different patch size distributions and shapes?

The principal approach of the current study is to compare several calibration metrics for two similarly parameterized spatially explicit land use change models, one that is applied to a highly fragmented region composed of ex-urban and urban fringe areas and the other which is applied to an area where development occurs from city centres outwards. We apply a neural network-based model in two diverse regions of the world, one which is south-west Wisconsin, USA that includes the city of Milwaukee and large surrounding ex-urban areas, and the other located in south-eastern Albania, which is composed of several large towns surrounded by agricultural land. We compare a variety of calibration metrics, including mean square error of model fit, a percentage of correctly predicted urban change cells, location accuracy departing from chance using the kappa statistic, comparison of model goodness-of-fit given variable threshold values across probability distributions using the area under the receiver operating characteristic (AROC) curve statistic, and four different spatial pattern metrics. The model we employ, called the land transformation model (LTM), uses data from two time periods, spatially parameterized drivers of change using analysis tools within a geographic information system, and neural network software that applies pattern recognition algorithms to data. We also explore the behaviour of the neural network as it learns from patterns in input (e.g. drivers or independent variables) and output (e.g. urban change maps or dependent variable) data.

The outline of the paper is as follows. First, we present a brief overview of the LTM. Second, we present a description of two study sites chosen for our comparative analysis. Third, we describe our experimental design and describe the results of our modelling exercises. We conclude with a discussion on the relevance of our results in light of the differences in the spatial patterns and drivers of land use change in the two regions and discuss how the behaviour of neural networks differ from other land use modelling tools.

2. Description of the land transformation model

The LTM is a LUCC model that uses artificial neural networks and geographic information systems (Pijanowski *et al.* 2000, Pijanowski *et al.* 2002, Pijanowski *et al.* 2005). The independent variables for the model typically are elevation and distance to highways, streets, lakes, rivers and urban centres. Independent variables are presented to the neural network as inputs with the output as a Boolean map that reflects urban growth versus no urban growth between two points in time. Neural networks operate in the following manner. The neural net 'trains' on an input–output relationship throughout a certain subset of pixels within a dataset until it obtains a satisfactory fit between the independent variables and the data concerning urban change.

Neural networks (figure 1) are composed of input, hidden and output nodes and connection weights between the nodes, values for which are adjusted to fit data. The neural network compares fitted and observed values using a mean squared error (MSE) estimate such that smaller values of the MSE are sought over the course of the training exercise. Each pass through the network is called a 'cycle'. Pijanowski *et al.* (2005) found that the number of cycles used during the training exercise greatly influences the model goodness-of-fit as judged by accuracy of location using kappa statistic and spatial pattern of change. A testing procedure, which applies network weights achieved after any given cycle is applied to input values that are used to estimate the output, which are assigned values between 0.0 and 1.0; these values indicate relative propensity for urbanization. The relative propensity for change is analogous to a probability, although a neural network does not create 'probabilities' per se (Bishop 1999; for simplicity, we refer to these output values as 'probabilities'). This output map is then reclassified into a Boolean prediction map of urban growth (='1') or no urban growth (='0'), such that the number of pixels predicted as urban growth matches the number of pixels of observed urban growth.

3. Selection of study sites and historical background information

The two study regions selected for the modelling exercise are the south-eastern Wisconsin area (hereafter, SEWI), and four districts in south-eastern Albania in southern Europe (hereafter, ALB). The regional locations of these case studies are shown in figure 2. These two areas were selected because they both are undergoing significant urbanization but the patterns of change differ greatly.

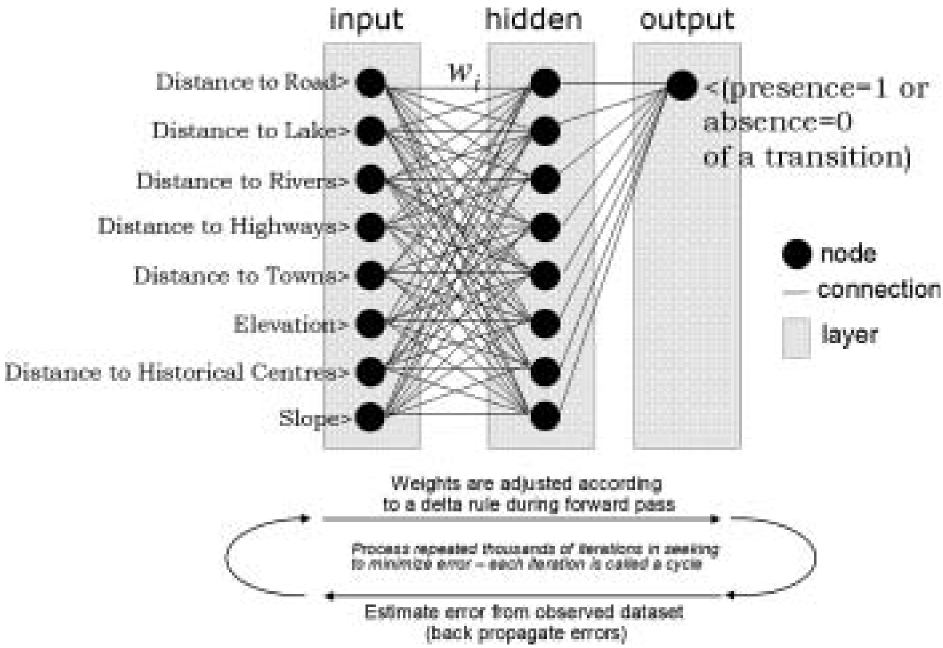

Figure 1. Neural network structure and terms used in the LTM.

3.1 *Albania*

3.1.1 Rationale and historical information. The conversion of land devoted for crop production into built-up areas is one of the most significant land changes in Albania (Agrotec Spa Consortium 2004). The expansion of urban areas mainly occurs on productive agricultural land, close to main market centers. The difference in land use and socio-economic characteristics between urban and rural areas in Albania is very significant. Following Albania's radical land reform after the collapse of socialism in 1991 virtually all agricultural land was distributed among former employees of the agricultural collectives (Cungu and Swinnen 1999). According to the Albanian Ministry of Agriculture and Food (2002) the land privatization process resulted in 420 000 farms with a mean farm size of 3.2 acres (1.3 ha), fragmented into an average of 3.6 non-contiguous parcels. Permanent land transactions are limited and Albanian agriculture is to a large extent subsistence-based. Besides having the highest emigration rates of all transition economies, Albania also witnessed large internal migration movements and a rapid process of urbanization with the share of the urban population increasing from 35% to 42% between 1989 and 2001 (Zezza *et al*. 2005). One quarter of the population are classified as poor in 2001 and two thirds of the poor live in rural areas (Zezza *et al*. 2005).

3.1.2 Study site characteristics. The study site in Albania includes the districts of Elbasan, Gramsh, Librazhd, and Pogradec in south-eastern Albania and covers an area of 3651 km^2. Figure 3 displays the major land use categories and their spatial distribution across the area under analysis.

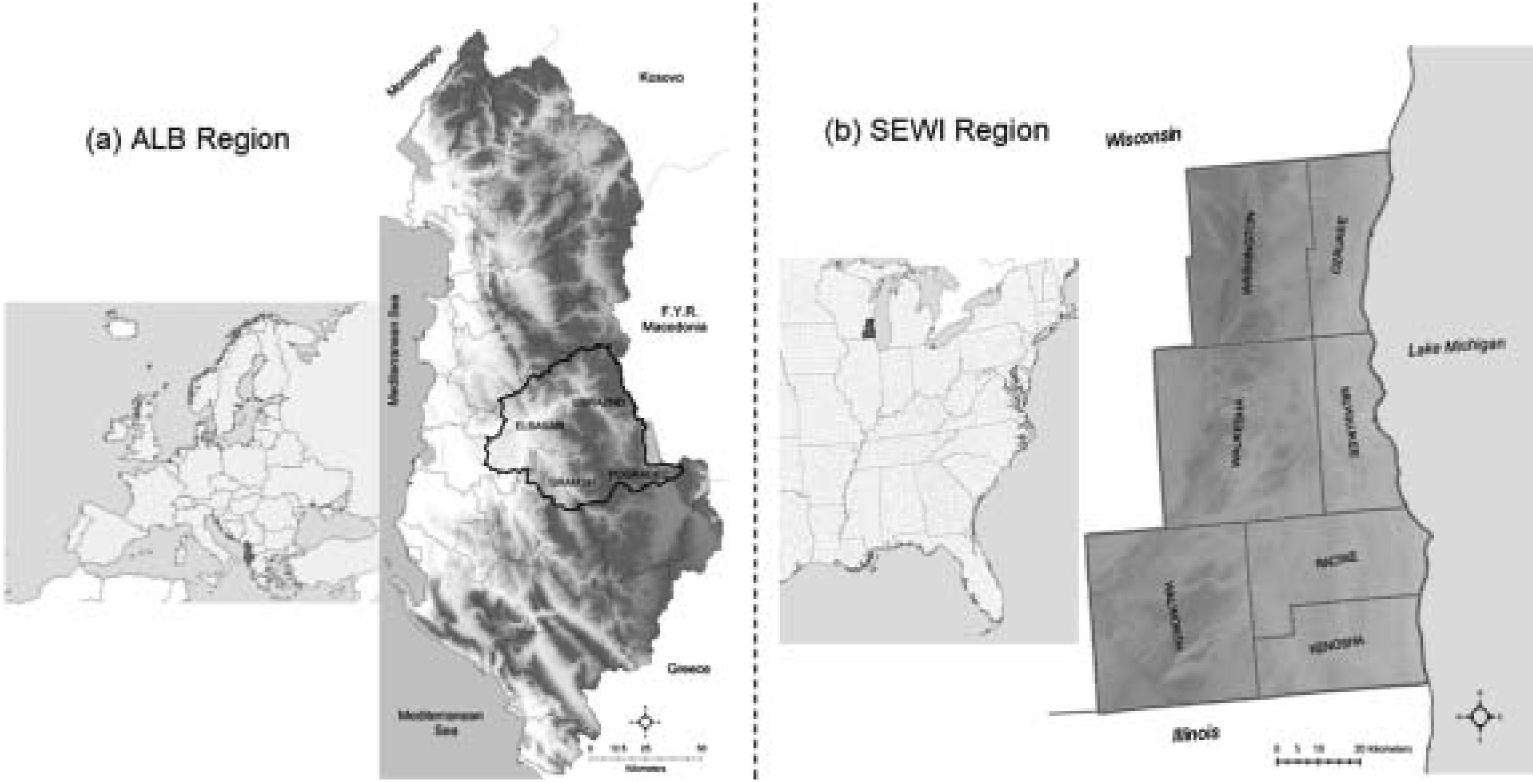

Figure 2. Location of study sites with respect to (a) Europe and (b) the USA. [The colour version of this figure is included in the online version of the journal].

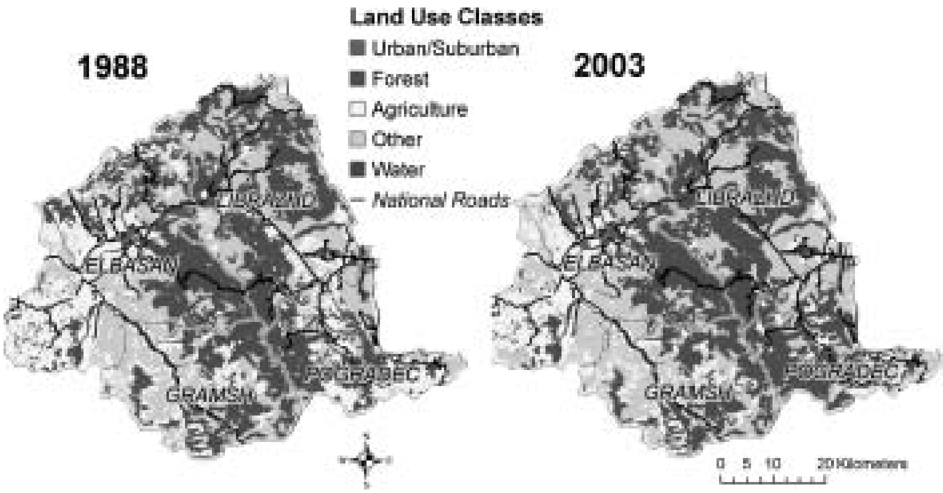

Figure 3. Major land uses in the ALB case study basin. Also shown are roads and the boundaries of the four districts in ALB. (Figures will appear in colour online).

The area is ecologically diverse with plains in the east that are intensively used for agricultural production. The southern and northern parts are mountainous areas with disadvantageous market access and adverse geophysical conditions. Livelihood strategies are dominated by remittances from internal and external emigration and smallholder subsistence-based production. Especially in rural mountainous areas, the percentage of subsistence agricultural production exceeds 30% of total production (World Bank 2003).

The district of Pogradec in the east, situated along Lake Ohrid, is one of the major tourist attractions in Albania. Elbasan is the third largest city in Albania with approximately 100000 inhabitants and high growth rates. Both Librazhd and Gramsh districts are among the poorest in Albania.

3.1.3 Geographic information system (GIS) data for model. All spatial data were resampled to a spatial resolution of 30×30 m and projected to the *Albers Equal Area* projection system, European datum 1950. Data were stored in ArcGIS 9.0 GRID format as integer values.

The land cover data are derived from visual on-screen interpretation of Landsat and Aster satellite images for the pre-socialist state in 1988 (thematic mapper, TM) and for 2003 (TM and Aster). Interpretation was conducted at an on-screen resolution of no less than 1:40 000. Land-cover categories were resampled to a binary grid comprising urban and non-urban classes. Urban changes were calculated using the GIS and locations of change were presented to the neural networks.

Geophysical driving variables included raster data for elevation and slope degrees. 10 m contour lines from 1:25 000 topographic maps were used to interpolate a digital elevation model (DEM). Elevation and slope were calculated from this DEM.

Euclidean distances were calculated to urban areas derived from the satellite image in 1988 and to the four district capitals. The road and river network were digitized by the Albanian National Forest Inventory (ANFI) from 1:25 000 topographic maps. We use the Euclidean distance to the entire road network, to the major, surfaced roads, and to rivers. We calculated the distance from each cell to the nearest lake as an additional driver.

Non-development zones comprised of already existing urban areas in 1988, water bodies (lakes and rivers), and the road network, were combined to define a unified exclusionary zone for the LTM simulation experiments. Cells in the exclusionary zone are never presented to the neural network.

3.2 *South-eastern Wisconsin*

3.2.1 Rationale and historical information. The south-eastern Wisconsin (SEWI) area contains a typical mix of historical land use development traced back to the nineteenth century, as well as relatively recent and dynamic patterns of urbanization and suburbanization change. The city of Milwaukee, one of the major cities in the Great Lakes region, is located within the study site. The SEWI region covers a total area of 6 965 km^2, and consists of seven counties: Kenosha, Milwaukee, Ozaukee, Racine, Walworth, Washington, and Waukesha Counties. Figure 4 displays the major land use classes of the SEWI area and their spatial distributions.

The region can be considered as unified in terms of its land use patterns, since land use changes occur under the planning jurisdiction of the South-eastern Wisconsin Regional Planning Commission (SEWRPC), since its establishment, in 1960. The SEWRPC is the official regional planning agency for the case study area.

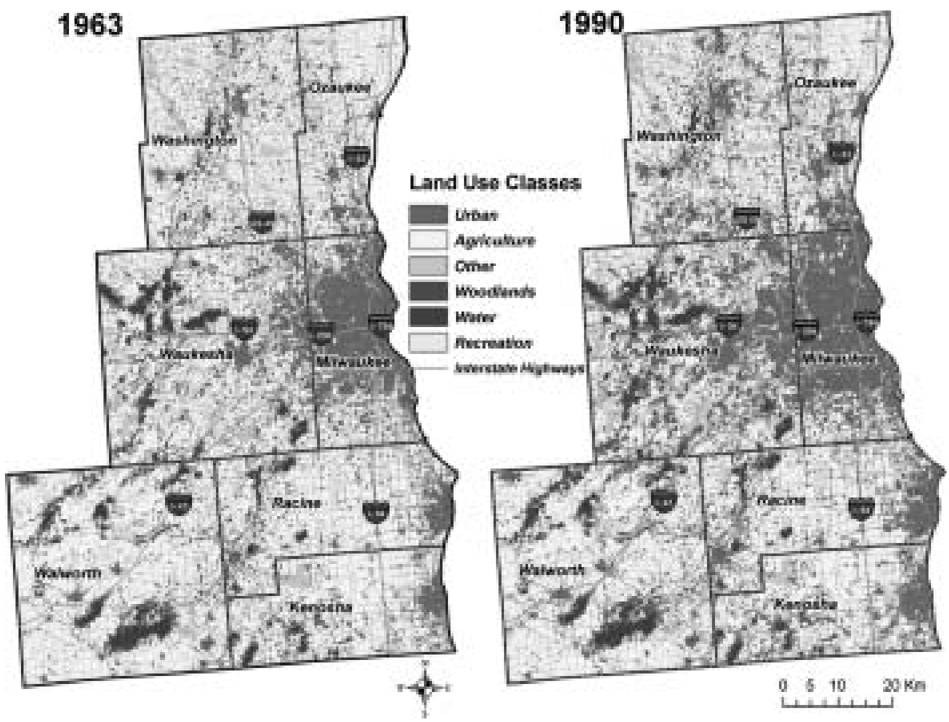

Figure 4. Major land uses in the SEWI case study area. Also shown are roads, interstate highways and the seven counties in SEWI. (Figures will appear in colour online).

3.2.2 Study site characteristics. The study area has a wider regional planning importance to the SEWRPC, and is considered to include the broader character of changes for the Milwaukee area along the eastern shores of Lake Michigan. Extensive suburban development has occurred during the study period under analysis (1963 to 1990) with greater socio-economic importance. A significant amount of growth has historically occurred in local communities in relative close proximity to the edges of the city of Milwaukee, especially in local communities and townships with favourable physical and natural features (e.g. small lakes, rivers, etc.). Lake shore suburban development accounts for a very significant amount of suburban growth (figure 4). The same is true for road-side development, especially along highway and state road corridors. At least four recognizable development corridors can be identified: the south corridor alongside Interstate highway I-94; the north corridor alongside state highway 145/45; the west Interstate-94 highway corridor; and the south-west corridor alongside state highway 43.

3.2.3 GIS data for model. All spatial data were resampled to a spatial resolution of 30×30 m and referenced to the *Albers Equal Area* projection system, North American datum 1983. Data were stored in ArcGIS 9.0 GRID format as integer values.

The land cover data were derived from polygon features of land use for the SEWRPC seven-county land use inventory from board-digitized 1″ = 400′ scale aerial photographs, in five-year intervals between 1963 and 1990 (SEWRPC 2000). Elevation and slope for the SEWI region were obtained from the USGS's Shuttle Radar Topography Mission (SRTM) 30 m data (http://seamless.usgs.gov).

Euclidean distances were calculated to urban areas derived from the SEWRPC historic urban growth inventory for historical urban areas in 1900 and point-locations of major cities and towns. The road and river network were digitized by line conversion from the land-use polygon maps. We used the Euclidean distance to two road categories (major interstate highways and state and subdivision roads), and to rivers. We used the GIS to calculate the distance from each cell to the nearest lake as an additional driver.

Non-development zones comprised of already existing urban areas in 1963, water bodies (lakes and rivers), and the road network, were combined to create an exclusionary zone for the simulations.

4. Experimental design

4.1 *Selection of simulation boxes*

We selected 2.5×2.5 km subsets of the seven county SEWI region and the four province ALB region to develop separate simulation areas; a total of 60 simulation areas were created, 30 for each region. Each subset represented a grid composed of 83 columns and 83 rows with 30×30 m cells. We used ArcGIS 9.0 (ESRI 2004) to create a polygon shape file with 2.5×2.5 km sampling boxes for both regions and then, using the tabulate area function, calculated the amount of urban change and amount of exclusionary area in each box. Sampling boxes falling partially outside the region's geographical extent were excluded from the simulations (i.e. boxes around the edges). For SEWI, 1409 2.5×2.5 km boxes were created; all of these boxes experienced urban change over the course of the study. For Albania, 508 boxes were created but only 49 boxes contained urban change.

Initial simulations indicated that a significant amount of variation in the area covered by the exclusionary zones existed within the sampled boxes in the SEWI region (but not in the ALB region), thus rendering comparative assessment problematic because of the substantial variability in valid cells for modelling. To reduce potential bias in our simulations owing to the presence of a large number of exclusionary cells, we selected only those simulation boxes where the amount of urban change exceeded the amount of exclusionary zone cells for all SEWI simulations. We sorted the remaining candidate simulation boxes according to percent urban change in each sampling bin, in order to eventually select 30 boxes for each study area. A random seed generator was used to select one box from each 30-tile bin. We then grouped the simulation boxes into ten groups such that the first three simulations in the 30-tile group represented group 1, the second three simulations, group 2, and so on. This sampling procedure resulted in the 30 simulations boxes for each study area compressed into ten simulation groups. Table 1 summarizes the results of the selection procedure providing the amount of urban change in each simulation box and the mean amount of urban change by simulation group. Note that the least amount of urban change was 3.16% and 0.01% in SEWI and ALB respectively, and the most amount of urban change was 38.86% and 36.52% in SEWI and ALB respectively. We report the means for calibration metrics for the three simulations in each of the 10 simulation groups for the two study areas.

4.2 *Modelling*

We followed the procedures of Pijanowski *et al.* (2002) and Pijanowski *et al.* (2005) to process data by the GIS for input to the Stuttgart Neural Network Simulator. First, all data were exported from the GIS and converted to a neural network pattern file with eight inputs, eight hidden nodes and one output node. Output was coded as binary with '0' = no predicted change to urban and '1' = change to urban. We selected every other cell to present to the neural network during the training exercise. We used a back propagation, feed forward neural network and a logistic activation function. During each training cycle, the neural net was set to shuffle, so that the order of presentation of input data to the neural network randomly changed during each cycle of the training session. We created separate pattern files for each of the simulation boxes in the two study areas. We saved to a neural network configuration file the neural network connection weights, activation and bias values for input, hidden and output nodes, as well as the MSE for every 100 cycles of the simulation runs. We then created pattern files representing all possible transition locations (areas outside the exclusionary area) with only input values and the neural network configuration files and had the neural network estimate output values. This constituted the testing exercise. We followed Pijanowski *et al.* (2002, 2005) to calibrate the model by thresholding the number of cells that transitioned in each of the simulation boxes. Because of the large amount of output results (50 000 networks and output transition maps), we saved 44 different training cycle simulations for analysis of the results and the performance analysis that follows. We selected more training cycles to save and analyze during the early training session due to the general pattern of fast learning initially by the neural nets with diminishing learning rate with more training cycles.

Table 1. Summary of the amount of urban change in SEWI and ALB simulation boxes and how the 30-tile sampling boxes were aggregated into simulation groups.

Ranked Area Box	Simulation Group	% Urban Changes		Ranked Area Box	Simulation Group	% Urban Change	
		SEWI	ALB			SEWI	ALB
1	1	3.16	0.01	16	6	14.44	4.53
2	1	3.57	0.40	17	6	18.00	5.12
3	1	5.65	0.42	18	6	18.78	5.18
	mean	*4.13*	*0.28*		*mean*	*17.07*	*4.94*
4	2	6.60	0.56	19	7	19.55	5.63
5	2	7.54	0.76	20	7	19.60	7.46
6	2	7.99	1.34	21	7	20.83	7.50
	mean	*7.38*	*0.89*		*mean*	*19.99*	*6.86*
7	3	8.67	1.62	22	8	23.24	7.95
8	3	8.80	1.71	23	8	23.71	8.63
9	3	9.16	2.03	24	8	23.77	8.83
	mean	*8.87*	*1.79*		*mean*	*23.57*	*8.47*
10	4	9.91	2.19	25	9	24.52	13.22
11	4	10.73	2.60	26	9	24.93	13.35
12	4	11.44	2.69	27	9	29.63	14.01
	mean	*10.69*	*2.49*		*mean*	*26.36*	*13.53*
13	5	11.79	2.80	28	10	30.12	18.41
14	5	13.17	2.91	29	10	32.56	22.92
15	5	14.29	4.04	30	10	38.90	36.52

4.3 *Model calibration*

We employed several model calibration metrics to determine how well the neural net performed between the two study areas. First, we saved the mean square error calculated for each training cycle. Second, we calculated three metrics that quantify agreement between observed and simulated data, a percent correct metric (PCM), kappa agreement statistic (K) and the AROC curve. All three are derived from a standard two-by-two contingency table (cf. Pontius 2002) created when simulated and observed maps are compared. PCM is the percentage of true positives divided by the total number of urban cells changing in each simulation box. Kappa statistic quantifies the level of agreement between two maps (e.g. simulated and observed) compared against the null hypothesis that the maps do not differ by chance from a random map. We follow Pontius (2002) and Sousa *et al.* (2002) who suggest that K values of less than 0.4 reflect poor performing models, 0.4 to 0.6 are fair, 0.6 to 0.8 are good and K values greater than 0.8 represent excellent agreements between model and observed datasets.

We calculated the AROC for each of the simulations and plotted AROC as a function of simulation training cycle and percentage of urban change. Briefly, the receiver operating characteristic curve plots the rate of true positive to positive classifications (termed *sensitivity*) against the rate of false positive to negative classifications (*1-specificity*) as a discrimination or threshold value is varied between 0.0 and 1.0. We use a nonparameteric approximation using SPSS (SPSS Inc. 2003) to estimate the area under the curve that is produced by varying the threshold and plotting sensitivity against 1-specificity.

Our final set of calibration metrics focused on the number, size and shape of urban patches. We used the landscape ecology program FRAGSTATS to calculate the urban patch metrics contained in each simulation box for the two study areas. Details on these statistics can be found in McGarigal and Marks (1994) and Pijanowski *et al.* (2005). We used four different patch metrics calculated for each simulation box: (a) number of urban patches, (b) the average fractal dimension of urban patches, (c) the average contiguity index of all urban patches and (d) the average perimeter-area ratio of urban patches. We computed these for: (a) the observed urban use change between initial and final; and, (b) new urban predicted by the model. These maps were coded such that '1'=location changed to urban and '0'=location did not undergo change to urban.

5. Results

Figure 5 summarizes the amount of change in each of the major land use categories for SEWI and ALB during our study periods. Between 1963 and 1990, the amount of urban increased by 50% in SEWI; the amount of urban doubled (i.e. increased by 100%) in ALB during a much shorter period (1988–2003). Agriculture was the loser land use for SEWI during the 27 year time period shrub/grass was the major loser (Figure 5a2) during the 12 year study period in ALB. Urban in SEWI was highly fragmented, with over 20 000 patches of urban located in this area in both time periods; far fewer patches occurred in ALB in 1988 and 2000. In SEWI, urban was arranged in more patches than any other land use; in ALB, urban contained the fewest number of patches of land use compared to other land uses.

Figure 6 summarizes the results of urban patches predicted and observed for all 30 simulation boxes in each study area for 500 000 training cycles. Note that for the

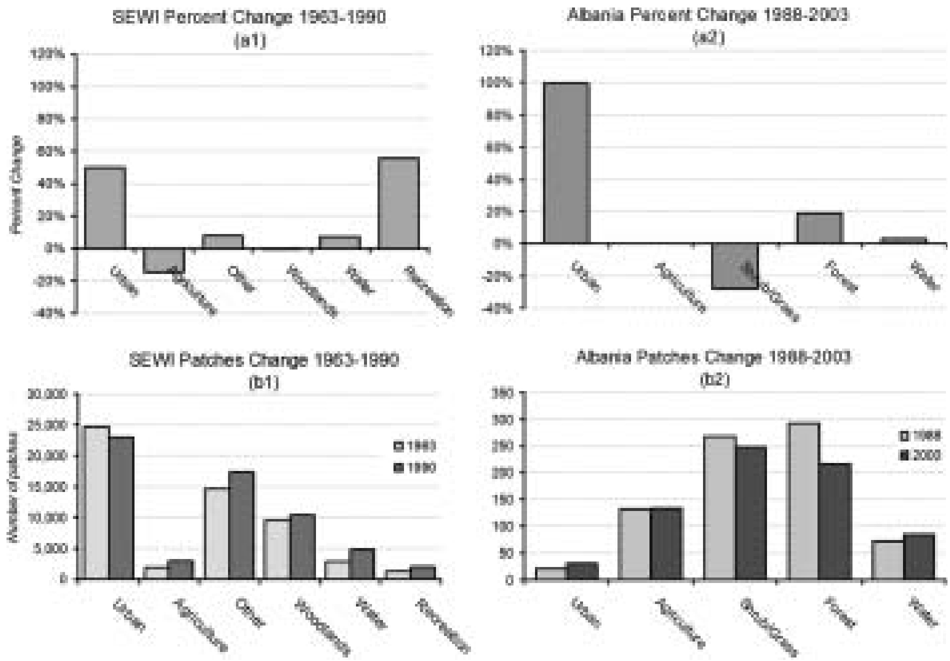

Figure 5. Summary of changes in land use classes in (a1) SEWI and (a2) ALB and number of patches in each land use class for (b1) SEWI and (b2) ALB.

SEWI simulations, the model predicted 449 new urban patches but 787 urban change patches actually occurred within these simulation boxes. Thus, the model created fewer number of patches in all SEWI simulations than observed, producing slightly more than half (57.1%) of the number of patches than observed. In particular, the model produced fewer patches of new urban within the smaller sized categories. For new urban patches composed of six to ten cells in size, the model under-estimated the number of patches; the model produced less than a third (31.5%) of the number of patches compared to observed. For next two larger patch size categories (i.e. 51–100 and 101–250), the model also produced fewer urban patches than observed but once patches exceed 251 cells, the model either produced more of the larger patches than observed (patch sizes 251–500) or produced the same as observed (501 or more). We compared the distribution of the number of urban patches across the range of patch size categories using a two-tailed, two sample Kolmogorov-Smirnov test and found that the simulated and observed distributions did not differ significantly ($p = 0.185$) from each other. Interestingly, in Albania, the model produced almost twice as many new urban patches as were observed. For small patches (between 1 and 50 cells) the model produced a greater number of smaller patches than observed. For urban patch size categories greater than 51 cells, the model produced fewer patches than observed. The differences in the number of urban patches for ALB across the range of patch size categories for simulated and observed were not significantly different when compared using a two-tailed, two-sample Kolmogorov-Smirnov test ($p = 1.000$). These results suggest that, in general, the model clumped new urban cells in SEWI compared with observed urban change and fragmented new urban cells in ALB compared with observed urban change.

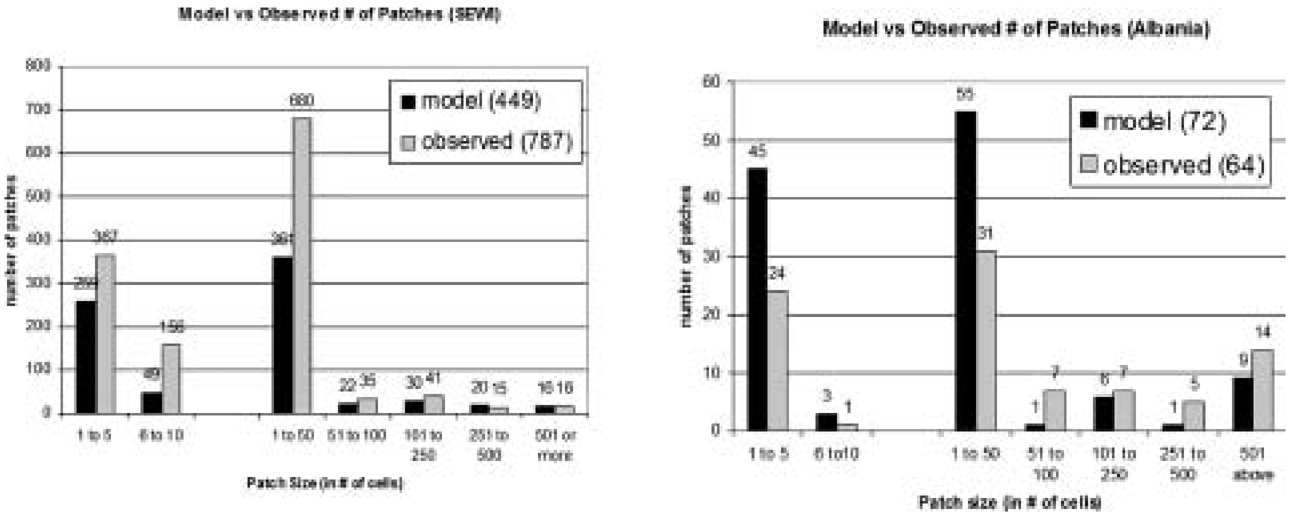

Figure 6. New urban patch size distributions of simulated and observed for (a) SEWI and (b) ALB study areas.

Our four non-shape performance metrics (MSE, PCM, K and AROC), grouped by the number of training cycles and by area box are summarized as the mean of the three simulations in each group, and are shown in figures 7 and 8. MSE [figure 7(a)] varies greatly by area box for the first 100 000 training cycles in SEWI compared to the MSE values across training cycles in ALB. In fact, for only a few area boxes, the MSE remains very low and close to zero for nearly all training cycles between 1000 and 500 000. In both study areas, simulations containing larger amounts of change (shown in orange and red) have larger MSE values.

The PCM [figure 7(b)] shows similar trends between SEWI and ALB. Simulation boxes in SEWI that contain little urban change produce the largest PCM values. Only a few area boxes and the earliest training cycles (<1000) produce PCM values smaller than 75. In ALB, nearly all simulations produced PCM values greater than 95.

Kappa values [figure 8(a)] show a reverse trend from MSE and PCM when comparing how the amount of change impacts model goodness-of-fit. Note that in SEWI, the best K are for simulations boxes with the most amount of urban change.

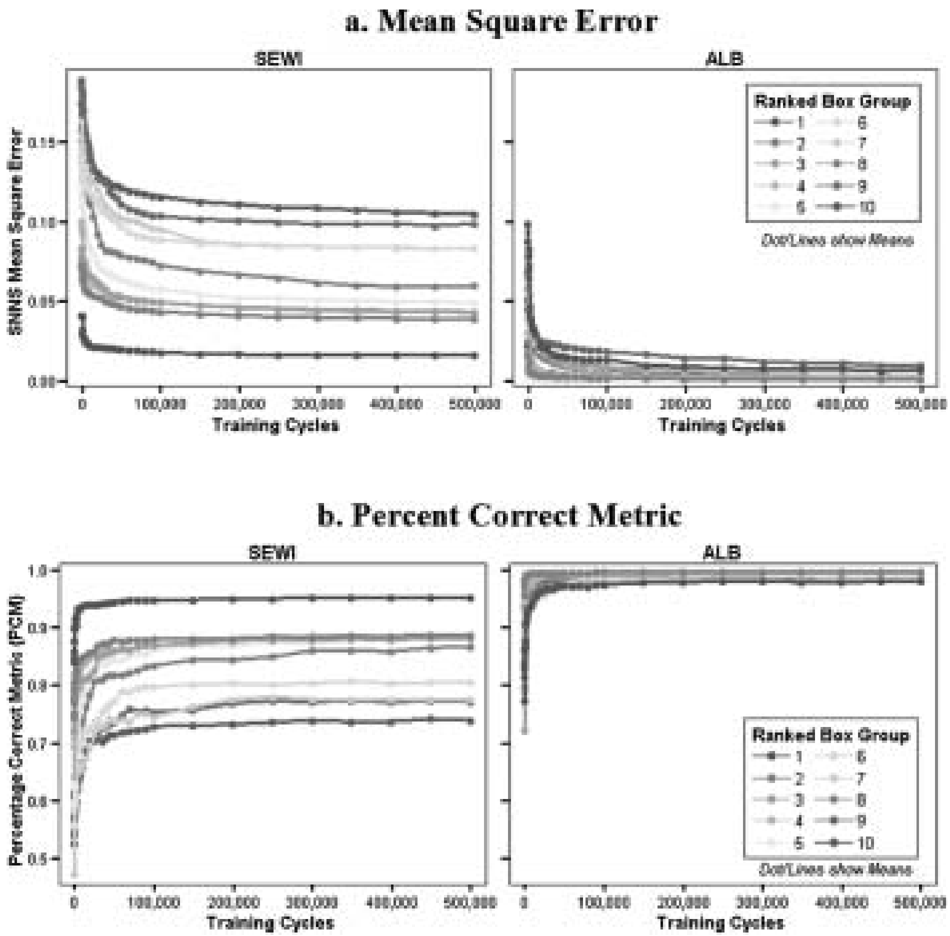

Figure 7. Accuracy assessment in the study areas across training cycles: (a) MSE and (b) PCM. (Figures will appear in colour online).

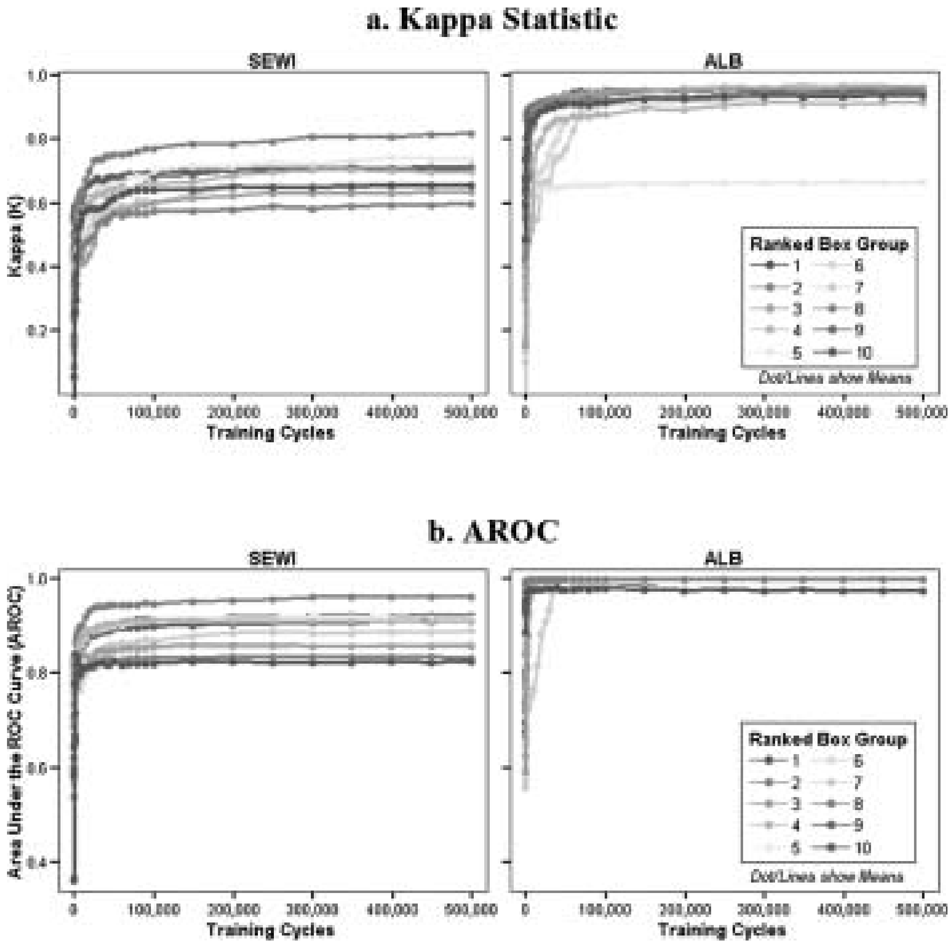

Figure 8. Accuracy assessment in the study areas across training cycles: (a) kappa and (b) AROC. [The colour version of this figure is included in the online version of the journal].

Most simulations in SEWI give ‘good’ K values between 0.6 and 0.8. Note that in ALB, nearly all simulations (except group 5) produced ‘very good’ K (>0.8) values and most of these were achieved very early in the training session (i.e. with 50 000 cycles or more).

The AROC comparisons for the same four study groups of simulations are shown in figure 8(b). AROC shows similar patterns as K with a few minor differences. First, the amount of variability between area boxes within each SEWI is less than K. Second AROC produces sharper rises in model goodness-of-fit in the earlier training cycles. Finally, most models in ALB appear to be nearly perfect, even for early training cycles. Of the 30 ALB area boxes, three simulations resulted in values of 1.0 (perfect model) after 10 000 training cycles.

A comparison of class shape metrics across training cycles for the two study areas are given in figure 9. The percent change represents the amount of change between what the model predicted divided by observed. Note that the model over predicts

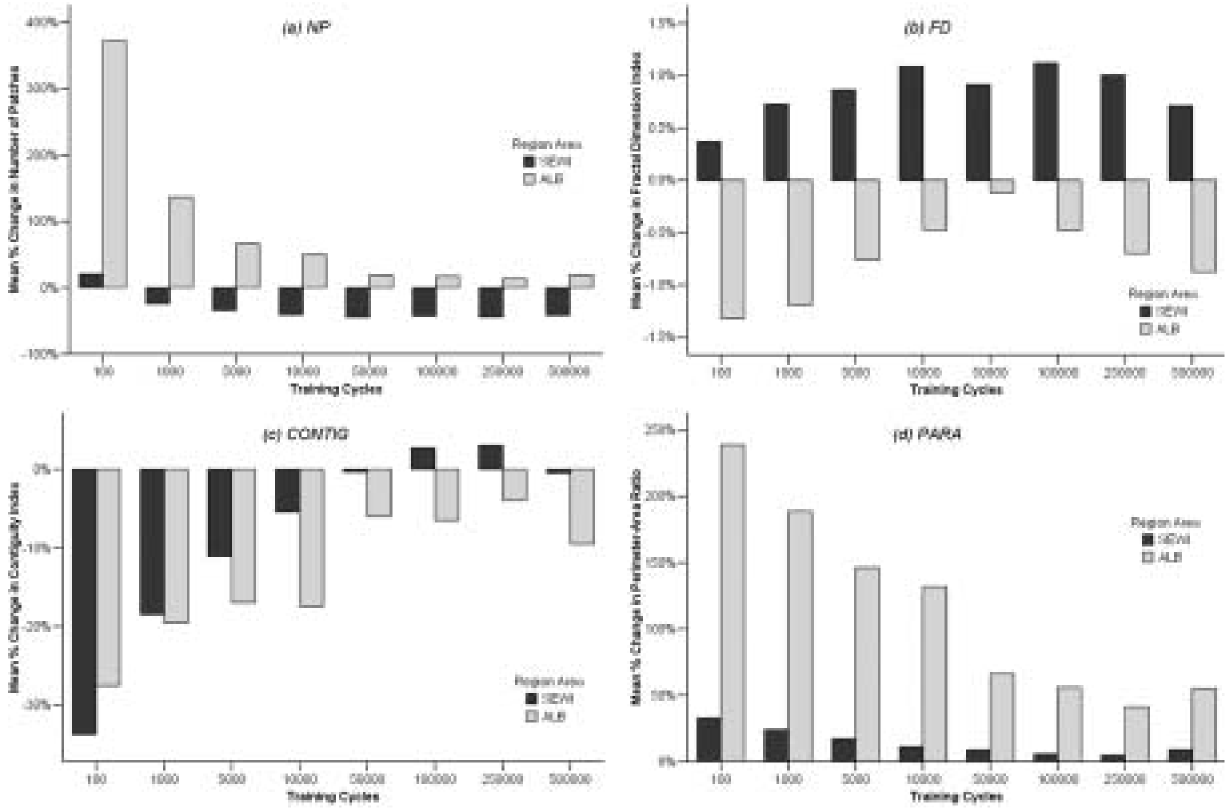

Figure 9. Patch metrics for the study areas expressed as a percentage deviation from observed organized across training cycles: (a) NP, (b) FD, (c) CONTIG and (d) PARA.

the number of patches (NP) in ALB for all training cycles and under predicts the number of urban patches for all training cycles in SEWI except for 100 training cycles [figure 9(a)]. The model improves greatly during the training exercise in the prediction for number of patches in ALB but appears to stabilize for SEWI with increasing number of training cycles. The difference in fractal dimension (FD) of predicted and observed urban patches is relatively small with most simulations producing differences in percent FD less than 1.0% [figure 9(b)]. Trends across training cycles appear to vary, though with the intermediate number of training cycles performing best in ALB and the fewest training cycles (e.g. 100) performing best in SEWI. The contiguity index measures the amount of clumpiness [figure 9(c)]. Note that this metric shows significant improvement as training progresses, producing values starting around 30% early in all training cycles for the two study areas and progressing to near 0.0% (a near perfect match between simulated and observed) at 500 000 cycles. The perimeter to area ratio (PARA) provides some of the greatest variability of trends between the two study areas [figure 9(d)]. In ALB, nearly all training cycles produced urban maps that had greater PARA than the observed urban change map. The differences were smaller in SEWI with more training cycles producing better fits between simulated and observed urban change maps.

The differences in shape metrics organized by simulation groups are shown in figure 10. Note that the NP across different amounts of urban change [figure 10(a)], as reflected in the area boxes organized from 1–10, show a large amount of variation across the groups. The greatest difference in NP between simulated and observed change maps are in ALB where very little urban change occurred, especially in boxes 2 and 3 [figure 10(a)]. In nearly all cases, the simulations produced more patches of urban change in ALB than observed; in SEWI, nearly all simulation boxes produced fewer patches than observed, although in group 8 the model produced more urban patches than observed. The FD metric [figure 10(b)] varies considerably across simulations in SEWI and ALB with most simulations producing greater FD than observed; no discernible pattern across the area boxes from small to large amounts of change are evident [figure 10(b)]. Most simulations produced CONTIG metrics that were smaller than observed for SEWI and ALB [figure 10(c)]. The PARA metric shows a slight improvement in matching patterns of urban change between simulated and observed as the amount of urban change increases in simulation groups [figure 10(d)]. When comparing all urban patch metrics across the simulation groups, it appears that urban change occurring in the mid-range of the values produced the best goodness-of-fit values (i.e. values closest to zero) and the worse are for simulation groups that experienced the least amount of urban change.

We selected two simulation boxes, one in SEWI and another in ALB, to follow how new urban patches are arranged by the neural net over the course of the training cycles. For our SEWI simulation box 11 (in group 4), the projected urban class after 100 training cycles is concentrated in the right and lower portions of the simulation area (figure 11). As the number of training cycles increases, the neural net begins to allocate urban change to areas in the middle-left portion of the simulation box and then progresses to the correct distribution and shape of urban change patches by 500 000 cycles. Small, but potentially important differences, if patch shape is critical, exist in models produced from 100 000 through 500 000 cycles. We also selected a simulation box (also in group 4) for ALB to follow maps of new urban change predicted by the model compared to the observed (figure 12). Note that in Albania, the model produces many small patches of new urban at the beginning of training, with four

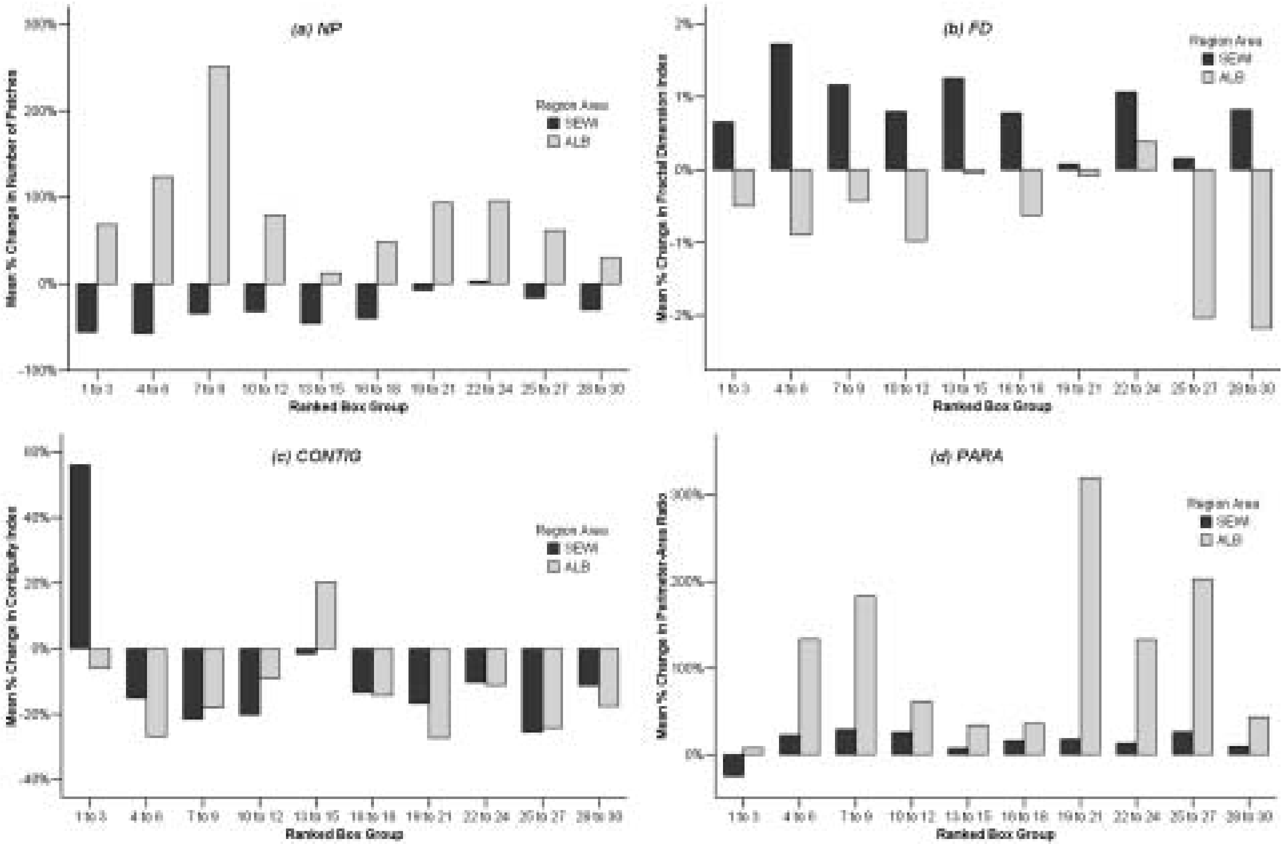

Figure 10. Patch metrics for the study areas expressed as a percentage deviation from observed across simulation boxes ranked from low to high amount of urban change (a) NP, (b) FD, (c) CONTIG and (d) PARA.

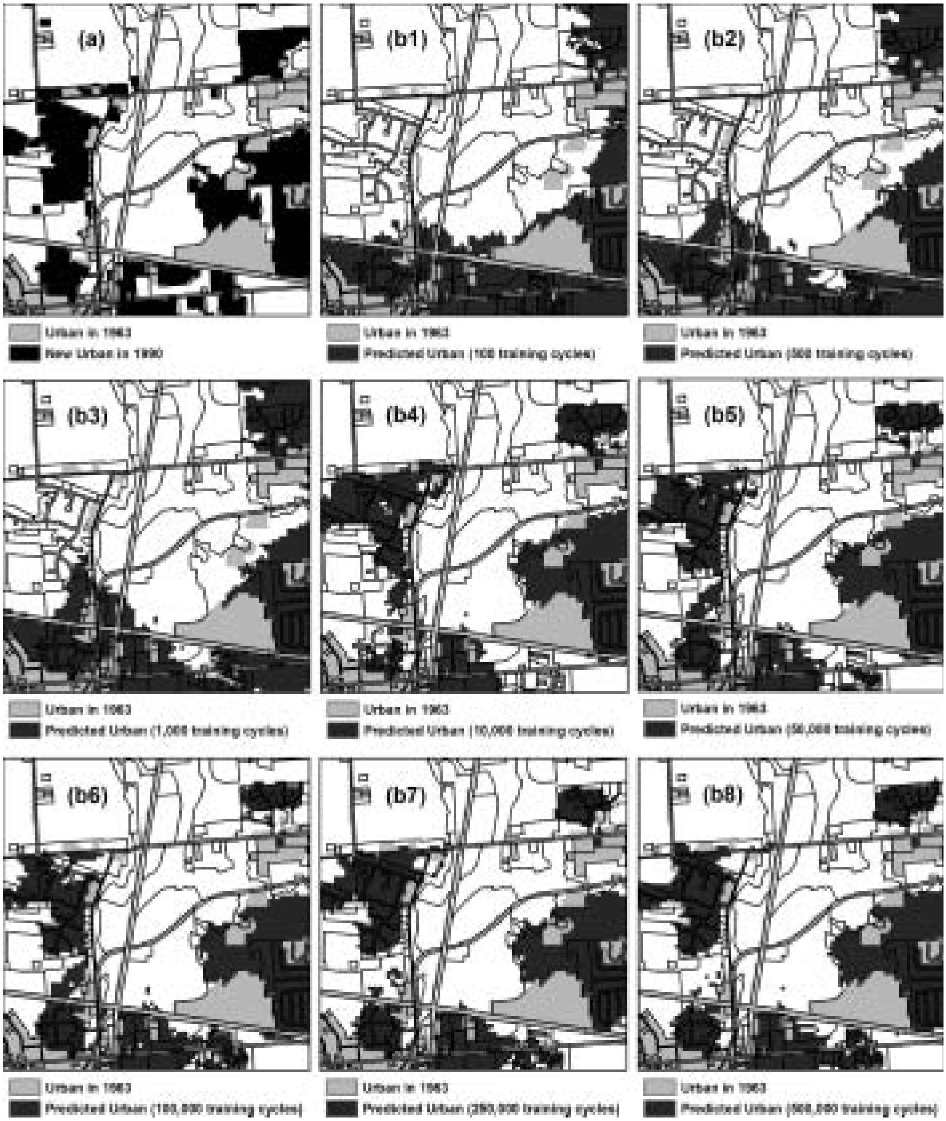

Figure 11. Illustration of spatial arrangement of new urban in the SEWI region: (a) observed change in urban cells between 1963 and 1990; (b1–8) sequence of new urban cells predicted by the model (for 1990).

isolated one cell patches of urban located in this simulation box. As the training progresses, the patches become more clumped and with fewer single-celled urban patches. The new urban patches are also increasingly aggregated toward the lower-right corner of the simulation box as training progresses. The final training cycle, 500 000 cycles, shows only one cell that is not correctly matched with the observed data. Both of these simulation sequences are consistent with other results, namely, that the model fits data better in ALB but creates more patches of new urban than observed, and that in SEWI, the fit to observed data is not as great, that the model produced

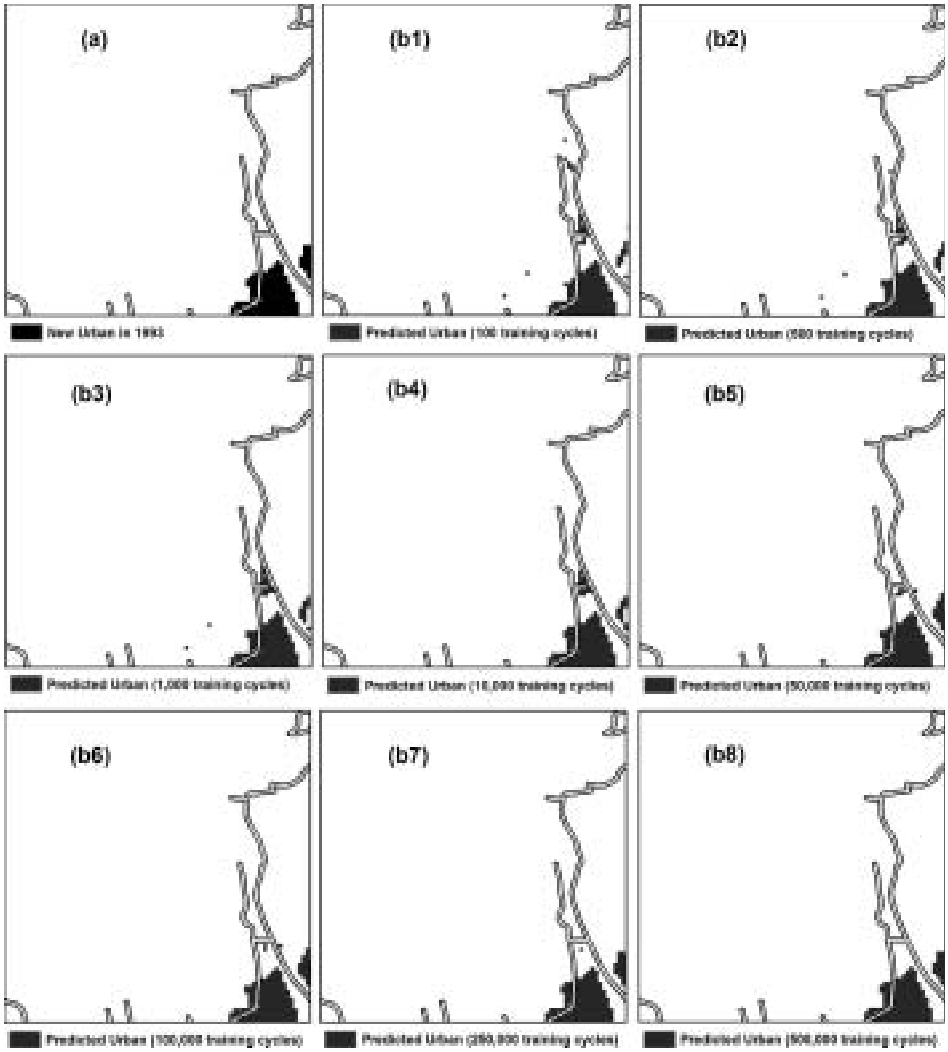

Figure 12. Illustration of spatial arrangement of new urban in the ALB region: (a) observed change in urban cells between 1988 and 2003; (b1–8) sequence of new urban cells predicted by the model (for 2003).

more clumped patches, especially early during the training session, and that the shapes of the new urban patches do not match precisely with observed data.

6. Discussion

6.1 *Comparing calibration metrics*

It is interesting that models using the same eight drivers of land use change were able to fit observed data relatively well, in most cases, in two different urbanizing areas despite drastically different economic, policy and human behaviours occurring in these two regions. We found that the model performed well in a majority of simulation boxes in both locations, especially after 100 000 training cycles, when location metrics, such as MSE, PCM, K and AROC are examined. For most simulations in both areas, MSE values were less than 0.10, PCM values were greater than 80%, K values were greater than 0.6 and AROC measures were greater than 0.8; all of these are considered to measure adequate model fits to observed data. The model clearly performed better in ALB, where new urban patches are more clumped than in SEWI. The model also produced better results earlier during the training session in ALB than in SEWI; most simulations reached a plateau after only a few thousand training cycles in ALB.

The analysis of new urban patches produced by the model (after 500 000 training cycles) compared to observed new urban patches yielded interesting results. In SEWI, which contained an enormous number of new urban patches between 1963 and 1990 ($N > 20\,000$ for the entire area including area outside the simulation boxes), the model produced more clumped new urban patches than observed; the model especially under-estimated the number of very small urban patches by more than 50% compared to observed. In ALB, however, where urban change was more prevalent but highly clumped, the model produced twice as many urban patches than observed. Statistical tests comparing frequency distributions of new urban patch sizes between simulation and observed were, however, not significant. We also saw that the number of urban patches departed greatly from those observed in early training cycles for ALB but did not change greatly across the training cycles for SEWI. New urban patch shapes, as reflected in the fractal dimension and perimeter-area ratio, varied across training cycles and between the study areas with the best fit model (e.g. FD for ALB) sometimes occurring during the early portion of our training exercise (e.g. 50 000 cycles). In other words, increasing the number of training cycles did not necessarily produce the best new urban patch shapes. The contiguity index, which measures the dispersion of urban patches across the simulation box, also produced the best fit to observed for both areas during the middle of the training exercise (e.g. also after 50 000 cycles but for SEWI). These results are important because if only location statistics (e.g. K, AROC) are used, then one might assume that the model does not improve after a certain number of training cycles.

The results of the simulations in SEWI and ALB are consistent with our understanding of the main drivers of land use change and is redundant information that would be required to improve the model in SEWI. In ALB, drivers of land use change are more 'top down', driven by market developments that pull many rural residents to urban and peri-urban areas. Distance to previous urban and roads are likely to be strong predictors of new urban in this system. In SEWI, on the other hand, as with most of the USA, land ownership and the ability to buy and sell land is made at the level of the individual. Information on individual decisions, such as socio-economic variables like age, wealth, attitudes and beliefs, would likely help capture more individualistic choices that are reflected in this landscape. It is unlikely that factors such as distance to roads, rivers, and previous urban are detailed enough to provide the necessary information to increase the model fit. Indeed, it is surprising that the model in SEWI fit observed data as well as it did in so many of the simulations. We also found that when we compared different calibration metrics such as location and new urban patch size, shape and configuration, that location metrics do not provide as much information as patch metrics, especially across training cycles. Location metrics tended to stabilize over training cycles giving the false impression that the model fit was not improving during training.

Predicting correct patch distributions and shapes may, in some cases, be more important than developing a model with high (viz. $K \approx 0.8$) accuracy of location. For example, it has been shown that wildlife movement and distributions depend upon the spatial distribution of habitats in a landscape (Lambeck 1997, Conroy *et al*. 2002, Johnson *et al*. 2004). Corridors, distances between patches, the amount of edge or core area of a patch, are especially important for wildlife management. The degree of land use/cover fragmentation is also important to the spread of invasive species with highly fragmented uses/covers being more susceptible to greater incidence of invasive species. The impact of land use change on water quality is also influenced by how many patches of natural areas exist in an area; at large scales, highly fragmented land uses are thought to impact surface roughness properties that influence land-atmosphere interactions (Houghton 1999, IPCC 1997).

Also of interest was the finding that in both areas, the calibration metrics were sensitive to the amount of urban change occurring in the simulated area. For example, SEWI, K and AROC showed that the best fit models were in the simulations areas with the most urban change, MSE and PCM best in areas with little urban change. However, when we compare the two study areas, there was more urban change across all SEWI simulation boxes (table 1) than in ALB and SEWI still yielded poorer fit models compared with ALB. In addition, there was greater variation in model calibration metrics between simulations in SEWI compared to ALB when compared across different amounts of urban change. This is interesting because the range of the amount of urban change was greater in the ALB simulation boxes (range = 0.3% – 26%) compared with that in SEWI (range = 4.1% – 33.9%). Pontius *et al*. (in review) have found a similar pattern of model performance when several different models are compared using persistence and null models. Models applied to areas containing a large amount of change (e.g. deforestation in the tropics) did better than models that attempted to simulate relatively 'rare' and 'salt and pepper' arranged land use change.

These results suggest that model goodness-of-fit is affected by the amount of land use change occurring in the study site and the amount of dispersion of new urban patches occurring in the study area. More urban change produces better model fits to observed data and more clumped urban change produces better model fits as well. When one considers the results of the new urban patch analysis, location metrics, such as K, may be poor estimates of model goodness-of-fit in areas that undergo small amounts of change and where changes are distributed in a patchy arrangement throughout the landscape. Indeed, if one could cautiously generalize these findings to consider the matrix of amount of change versus pattern of change (figure 13), certain model calibration metrics may be better than others depending on the rate and pattern of land use change. Study sites with small amounts of change distributed in many patches are likely never to yield significant K values, and modelling in these environments might be more difficult than in more clumped urbanizing areas. In these situations, the number of patches and patch shapes might provide a better estimate of model goodness-of-fit. Likewise, studies of areas with lots of urban change arranged in a clumped arrangement might use location and pattern metrics together. Overall however, more work is needed to determine what calibration metrics are best suited to particular areas undergoing distinct patterns of land use change.

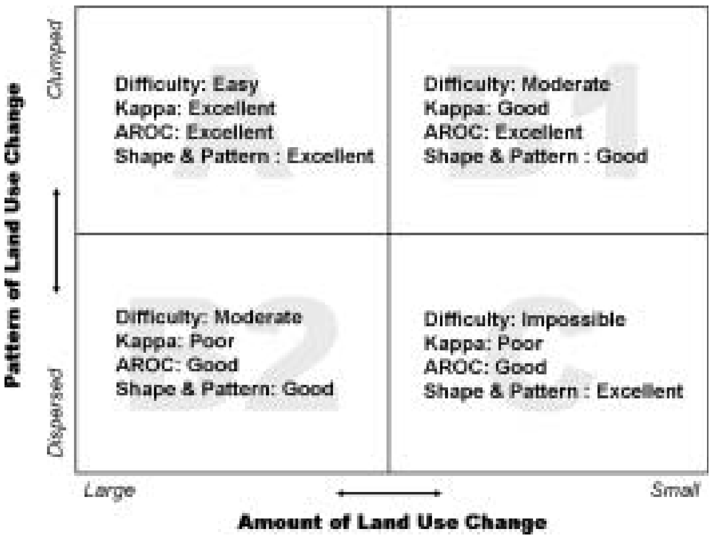

Figure 13. A proposed general relationship between amount and pattern of land use change in a study area and usefulness of metrics organized by location and pattern calibration metrics.

6.2 *Neural-network-based LUCC models*

Neural networks and statistically-based models, particularly logistic regression models, share important similarities while retaining critical differences. First, in order to fit the model to empirical data, neural networks employ a learning algorithm, rather than fitting data to a statistical model such as a least squares or maximum likelihood model. Over the course of the training exercise we have found here, and elsewhere (Pijanowski *et al.* 2005), too few training cycles result in very poor calibration metrics. Second, logistic regression models (e.g. Serneels and Lambin 2001) provide the relative weight of each independent variable. With neural nets, the relative contribution of a driver is not directly obtainable as internal weights cannot be interpreted as coefficients of independent variables of the model (Bishop 1999). In a previous paper (Pijanowski *et al.* 2002), we presented a 'remove one variable at a time' method that allowed us to determine the relative strength of each variable. Third, neural networks eventually produce many different models, some of which might be 'equally valid' because they fit data nearly identically when one compares a variety of calibration metrics. Each of these neural network models then represents a different 'view' of the data which could be useful either individually or collectively. As a group, these models could represent an ensemble of valid models that could be used for planning and management purposes each representing a plausible outcome. For example, we could select the 100 best fit models generated from thousands to millions of training cycles and then use each model to develop forecasts of land use patterns into the future (following the procedures of Pijanowski *et al.* 2002). These 100 land use maps could then be input into hydrologic impact models (e.g. Tang *et al.* 2005) to produce 100 impact assessments, the mean of which could be used to improve robustness of forecasts. This would be helpful from a policy standpoint as land use forecasts are fraught with problems associated with error inherent in maps (Pontius and Spencer 2005).

6.3 *Conclusions*

We found that a neural network based land use change model applied to two very diverse areas, each differing in the pattern of urbanization resulting from policy, economics and behaviour, and parameterized using the same drivers, could produce models that satisfactorily fit to observed data. We employed eight different calibration metrics and found that each metric provides different information on how well the model compares to observed urban change. However, we found that the model for Albania, where urban change is very clumped, performed consistently better than in south-eastern Wisconsin which contains patchy urbanization patterns. The number of urban patches, urban patch shape and arrangement of urban patches across the simulation areas differed between study areas, across training cycles and between simulation boxes grouped by the amount of urban change. We found that more training did not necessarily produce the best goodness-of-fit patch metrics. In general, we suggest that patch metrics are useful for calibration of land use change models and that they may be more useful than location metrics, such as kappa, in areas where urbanization is fragmented and diffuse, and especially for applications where patches of land use/cover are important for policy and management.

Acknowledgements

We would like to acknowledge Amelie Davis, Snehal Pithadia and Deepak Ray for technical assistance on this project. This work was partially funded by NSF (Grant WCR:0233648), EPA STAR Biological Classification Program, the Great Lakes Fisheries Trust, the Purdue University Department of Forestry and Natural Resources, and the Purdue University Research Foundation. Dr Muller's visit to Dr. Pijanowski's lab was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft) under the Emmy Noether Program. Suggestions from two anonymous reviewers helped to improve the manuscript. Comments and editorial improvements by Richard Aspinall are greatly appreciated.

References

AGROTEC SPA CONSORTIUM, 2004, *Albanian National Forest Inventory (ANFI): Analysis of the Spatio-Temporal and Semantic Aspects of Land Cover/Use Dynamics* (Rome: Agrotec SpA Consortium).

ALBANIAN MINISTRY OF AGRICULTURE AND FOOD, 2002, *Annual Report 2002* (Tirana: Ministry of Agriculture and Food).

BISHOP, C. M., 1999, *Neural Networks for Pattern Recognition* (Oxford: Oxford University Press).

CARRION-FLORES, C. and IRWIN, E. G., 2004, Determinants of residential land-use conversion and sprawl at the rural-urban fringe. *American Journal of Agricultural Economics*, **86**, pp. 889–904.

CAVAILHES, J., PEETERS, D., SEKERIS, E. and THISSE, J. F., 2004, The periurban city: why to live between the suburbs and the countryside. *Regional Science and Urban Economics*, **34**, pp. 681–703.

CONROY, M. J., MOORE, C. T., SCOTT, J. M., HEGLUND, P. J., MORRISON, M. L., HAUFLER, J. B., RAPHAEL, M. G., WALL, W. A. and SAMSON, F. B., 2002, Wildlife habitat modeling in an adaptive framework: the role of alternative models. In *Predicting Species Occurences: Issues of Accuracy and Scale*, pp. 205–218 (Washington D.C.: Island Press).

CUNGU, A. and SWINNEN, J. F. M., 1999, Albania's radical reform. *Economic Development and Cultural Change*, **47**, pp. 605–619.

DANIELS, T., 1999, *When City and Country Collide: Managing Growth in the Metropolitan Fringe* (Washington: Island Press).

ESRI, 2004, ArcGIS Desktop Help. Available online at: http://webhelp.esri.com/arcgisdesktop/9.0/.

FELSTEHAUSEN, H., 1999, *Urban Growth and Land Use Changes in Tirana, Albania: With cases describing urban land claims* (Albania Series, Working Paper No. 31) (Madison: Land Tenure Center).

HOUGHTON, R. A., 1999, The annual net flux of carbon to the atmosphere from changes in land use 1850–1990. *Tellus*, **50B**, pp. 298–312.

IPCC, 1997, *The Regional Impacts of Climate Change: An Assessment of Vulnerability* (Report of IPCC Working Group II): United Nations Environmental Programme - Intergovernmental Panel on Climate Change.

JOHNSON, C. J., BOYCE, M. S., MULDERS, R., GUNN, A., GAU, R. J., CLUFF, H. D. and CASE, R. L., 2004, Quantifying patch distribution at multiple spatial scales: applications to wildlife-habitat models. *Landscape Ecology*, **19**, pp. 869–882.

LAMBECK, R. J., 1997, Focal species: a multi-species umbrella for nature conservation. *Conservation Biology*, **11**, pp. 849–856.

MCGARIGAL, K. and MARKS, B. J., 1994, *FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure* (No. Gen. Tech. Rep. PNW-GTR-351) (Portland: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station).

PIJANOWSKI, B. C., BROWN, D. G., SHELLITO, B. A. and MANIK, G. A., 2002, Using neural networks and GIS to forecast land use changes: A land transformation model. *Computers, Environment and Urban Systems*, **26**, pp. 553–575.

PIJANOWSKI, B. C., GAGE, S. H. and LONG, D. T., 2000, A land transformation model: Integrating Policy, socioeconomics and environmental drivers using a geographic information system. In *Landscape Ecology: A Top Down Approach*, J. Sanderson and L. Harris (Eds) (Boca-Raton: CRC Press, Lewis Publishers).

PIJANOWSKI, B. C., PITHADIA, S., SHELLITO, B. A. and ALEXANDRIDIS, K., 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 Sciences*, **19**, pp. 197–215.

PONTIUS, G. R., Jr., 2002, Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. *Photogrammetric Engineering and Remote Sensing*, **68**, pp. 1041–1049.

PONTIUS, G. R., Jr. and SPENCER, J., 2005, Uncertainty in extrapolations of predictive land-change models. *Environment & Planning B: Planning and Design*, **32**, pp. 211–230.

SERNEELS, S. and LAMBIN, E., 2001, Proximate causes of land-use change in Narok, District: Kenya: a spatial statistical model. *Agriculture, Ecosystems and Environment*, **85,** pp. 65–81.

SOUTHEASTERN WISCONSIN REGIONAL PLANNING COMMISSION, 2000, *Aerial Photography and Orthophotography Inventory (1963–2000)* (Waukesha: SEWRPC). Available online at: http://www.sewrpc.org/.

SOUSA, S., CAEIRO, S. and PAINHO, M., 2002, *Assessment of Map Similarity of Categorical Maps Using Kappa Statistics: The Case of Sado Estuary*. Paper presented at the ESIG 2002, Tagus Park, Oeiras.

SPSS INC., 2003, SPSS for Windows (Version 12.0) (Chicago: SPSS Inc).

TANG, Z., ENGEL, B., LIM, K., PIJANOWSKI B. and HARBOR, J., 2005, Minimizing the impact of urbanization on long-term runoff. *Journal of the Water Resources Association*, **41**, pp. 1347–1359.

WORLD BANK, 2003, *Albania – Poverty Assessment* (Economic Report No. 26213-AL). Washington, D.C.: The World Bank, Human Development Sector Unit, Europe and Central Asia Region.

WU, J., 2004, Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology*, **19**, pp. 125–138.

ZEZZA, A., CARLETTO, G. and DAVIS, B., 2005, Moving away from poverty: a spatial analysis of poverty and migration in Albania. *Journal of Southern Europe and the Balkans*, **7**, pp. 175–193.