Environ Monit Assess (2017) 189: 300
DOI 10.1007/s10661-017-5986-3

Transition index maps for urban growth simulation: application of artificial neural networks, weight of evidence and fuzzy multi-criteria evaluation

Hossein Shafizadeh-Moghadam · Amin Tayyebi · Marco Helbich

Received: 15 August 2016 / Accepted: 9 May 2017 / Published online: 29 May 2017

Abstract Transition index maps (TIMs) are key products in urban growth simulation models. However, their operationalization is still conflicting. Our aim was to compare the prediction accuracy of three TIM-based spatially explicit land cover change (LCC) models in the mega city of Mumbai, India. These LCC models include two data-driven approaches, namely artificial neural networks (ANNs) and weight of evidence (WOE), and one knowledge-based approach which integrates an analytical hierarchical process with fuzzy membership functions (FAHP). Using the relative operating characteristics (ROC), the performance of these three LCC models were evaluated. The results showed 85%, 75%, and 73% accuracy for the ANN, FAHP, and WOE. The ANN was clearly superior compared to the other LCC models when simulating urban growth for the year 2010; hence, ANN was used to predict urban growth for 2020 and 2030. Projected urban growth maps were assessed using statistical measures, including figure of merit, average spatial distance deviation, producer accuracy, and overall accuracy. Based on our findings, we recomend ANNs as an and accurate method for simulating future patterns of urban growth.

Keywords Land cover change · Artificial neural networks · Weight of evidence · Fuzzy analytical hierarchical process · Relative operating characteristics

Electronic supplementary material The online version of this article (doi:10.1007/s10661-017-5986-3) contains supplementary material, which is available to authorized users.

H. Shafizadeh-Moghadam (✉)
Department of GIS & RS, Tarbiat Modares University, Tehran, Iran
e-mail: h.shafizadeh@modares.ac.ir
H. Shafizadeh-Moghadam
Institute of Geography, University of Heidelberg, Heidelberg, Germany

A. Tayyebi
Geospatial Big Data Engineer, Monsanto, MO, USA
e-mail: amintayyebi@gmail.com

M. Helbich
Department of Human Geography and Spatial Planning, Utrecht University, Utrecht, the Netherlands
e-mail: m.helbich@uu.nl

Introduction

Land cover change (LCC) modeling is highly complex so that non-linear relationships and inter-relations among driving forces require flexible geospatial models to precisely enumerate factors influencing urban growth processes (Almeida et al. 2008; Azari et al. 2016). Urbanization has been one of the most common forms of LCC and received considerable attention among scholars due to its various consequences on the environment such as biodiversity (Hansen et al. 2012), climate (Tayyebi and Jenerette 2016), and surface water (Tayyebi et al. 2015). Thus, a variety of data-driven and knowledge-based models have been developed to simulate urban growth (e.g., Sangermano et al. 2010; Zhu et al. 2010; Lin et al. 2011; Park et al. 2011; Tayyebi et al. 2014).

LCC models can be subdivided from different perspectives, for example, (a) aspatial and spatial statistical models, such as logistic regression and weight of evidence (WOE), which use historical data to quantify relationships between land transformations and their underlying driving forces (e.g., Tayyebi et al. 2010; Lin et al. 2011; Shafizadeh-Moghadam and Helbich 2015), (b) dynamic simulation models such as the rule-based cellular automata and agent-based models, which capture temporal dimensions and local conditions, and consider spatial interactions among land cover classes (e.g., Ligmann-Zielinska and Jankowski 2010; Munshi et al., 2014; Fathizad et al. 2015; Feng and Liu 2016; Feng et al. 2016, 2017), and finally (c) machine learning approaches, like artificial neural networks (ANNs), which have the appealing property of being able to represent non-linear relationships among input and output variables, without requiring a priori knowledge of the actual associations involved (e.g., Lin et al. 2011; Hagenauer and Helbich 2012; Grekousis et al. 2013; Azari et al. 2016; Tayyebi et al. 2016a).

Central to most LCC models are transition index maps (TIMs) (Almeida et al. 2008). These maps indicate the likelihood that the status of cells may change within a given area. If the model is well calibrated, the greater the ability of TIMs to identify the location of change, the more effective the LCC models will be in simulating future growth. LCC models produce TIMs with different levels of predictive power (Tayyebi et al. 2014). Thus, TIMs generated by various approaches might lead to different growth patterns, meaning that using a TIM which can robustly capture the influence of underlying drivers and precisely assign the most likely change locations is still a challenge. Therefore, it is important to evaluate the strengths and weaknesses of LCC modeling techniques using empirical comparative studies to determine the most suitable technique.

WOE is one of the well-known approaches to model LCC (Thapa and Murayama 2011). WOE is a statistical approach with a straightforward interpretation of the produced weights, as well as mapping the uncertainty of the posterior probability as a result of uncertainty in weights and missing data (Bonham-Carter et al. 1994). Almeida et al. (2003), for example, coupled CA and WOE to calculate the suitability of change occurring among different land cover classes. Their approach was to identify major drivers and enumerate their influence to predict the suitability of change among land cover types. Since each input variable should be presented to WOE in a binary format (Bonham-Carter et al. 1994; Agterberg and Cheng 2002), the conversion of continuous variables to binary maps is still challenging.

As an alternative to WOE, TIMs can be generated by coupling multi-criteria evaluation (MCE) and fuzzy membership functions (Mitsova et al. 2011; Shafizadeh-Moghadam and Helbich 2013). While MCE techniques such as the analytical hierarchical process (AHP) is used to determine the relative importance of driving forces and the precedence weighting, fuzzy membership functions can be utilized to transform explanatory variables from different scales into comparable units (FAHP) (Park et al. 2011). Inadequate experts' knowledge and subjectivity in the weighting process are the main limitations relevant to the FAHP. Another shortcoming of MCE is that the transition suitability changes in a linear fashion and, thus, assumes linear trends among the underlying spatio-temporal processes (Kamusoko 2012). In contrast, ANNs do not face this limitation and also have the following advantages: (a) the ability to deal with "noisy" and incomplete data (Tayyebi et al. 2014, 2016a, b), (b) the ability to detect potential inter-dependencies among the underlying driving forces, (c) provision of non-linearities (Pijanowski et al. 2002, 2014), and (d) less distributional assumptions regarding the input data than basic regression (e.g., Hagenauer and Helbich 2012; Tayyebi et al. 2014). This suggests ANN as a powerful technique when projecting urban growth. For instance, Tayyebi et al. (2011) developed an integrated framework using ANNs, GIS, and remote sensing to simulate the urban boundary in Tehran, Iran, to protect environmentally sensitive areas in rural landscapes. However, ANN gives fewer clues about how to interpret the interactions among variables compared to WOE and FAHP.

Despite the large body of studies in the field of LCC modeling, the application and evaluation of LCC models in different geographical regions will make the strengths and weaknesses of each model more evident. For example, Park et al. (2011) analyzed TIMs created by means of frequency ratios, AHP, logistic regression, and ANN models. Their results using relative operating characteristic (ROC) showed the highest and lowest values for logistic regression and ANN, respectively. On the contrary, the ANN achieved the highest overall accuracy among the four approaches. When compared

to MCE, ANNs do not provide straightforward measures able to reflect the relative importance of factors (Hagenauer and Helbich 2012). Tayyebi and Pijanowski (2014) contrasted the predictive abilities of ANNs, classification and regression trees, and multivariate adaptive regression splines. Confirming Park et al. (2011), ANNs provided the best predictive accuracy.

To conclude, despite these first attempts at conducting LCC model assessments, studies regarding the comparison of the models from the knowledge-based and data-driven perspectives for creating TIMs are still limited and not conclusive. Therefore, this paper aims to compare the performance of three approaches to derive TIMs, including WOE, MCE coupled with FAHP as well as ANNs using the rapidly growing mega city of Mumbai, India, as a case study. Mumbai is an excellent study area as the city experienced extensive urban expansion in the last decade (Shafizadeh-Moghadam and Helbich 2013).

Materials and methods

Study area

Mumbai is the capital of the state of Maharashtra in western India that is located between the latitudes 18° 53′ to 19° 16′ N and longitudes 72° to 72° 59′ E (Fig. 1). With a population of approximately 11.8 million people, Mumbai is the largest city in India (Patel et al. 2014). Along with natural population growth and high rates of immigration, the city is characterized by the relocation of industries to urban peripheries and the subsequent migration of employees into those areas, as well as the growth of a significant slum population. As a result, Mumbai's administrative boundary has been extended twice, first in 1950 and again in 1957 (Pacione 2006). A retrospective LCC analysis of Mumbai conducted by Shafizadeh-Moghadam and Helbich (2013) confirmed that the city underwent fast-paced physical growth between 1973 and 2010. Monitoring how Mumbai is physically expanding and simulating how it will continue to spread, therefore, is crucially important for local urban planners and the authorities there (Tayyebi et al. 2016a).

Data sources and pre-processing

The database was constructed from a range of sources (Table 1), including Landsat images from 1990 (TM), 2001 (ETM), and 2010 (ETM^+) with a spatial resolution of 30 m. The remote sensing images were provided by the US Geological Survey. Moreover, a digital elevation model (DEM) with a cell size of 30 m was acquired from the US Shuttle Radar Topography Mission (SRTM) website and a transportation network layer extracted from OpenStreetMap (Helbich et al. 2012).

Utilizing Landsat data, land cover classes were classified (Fig. 2) using a supervised maximum likelihood algorithm (Tayyebi et al. 2016a). Accuracy assessments of the classified maps were carried out using reference data including topographic maps as well as Google Earth and OpenStreetMap data. Statistical assessments were conducted using the Kappa index based on 250 randomly selected points for each land cover category (van Genderen and Lock 1977; Azari et al. 2016). Different approaches have been proposed for the determination of sample size (Foody 2009) that most often it depends on the purpose and the intended accuracy of the produced maps. Van Genderen and Lock (1977) suggested to take at least 20 sample points per class for 85% accuracy.

In phase of prediction, urban growth during 2001 and 2010 was considered as dependent variable and several driving forces as explanatory variables with the initial timestamp of 2001. However, the determination of potential drivers is still inconclusive and varies between studies (Pijanowski et al. 2014). The complexity of drivers involved, local characteristics within geographic regions and a lack of access to data, particularly in developing countries, give rise to the issue of how to select variables. As guided by previous studies (e.g., Thapa and Murayama 2011) and by expert knowledge, the variables shown in Table 2 were prepared in a GIS environment.

The distance-based variables were created by computing the Euclidean distance and the density-based variables by using a neighborhood function as frequently done (e.g., Hu and Lo 2007). Distance-based variables include the distance to built-up areas, water bodies, railway lines, main roads, wetlands, forested areas, and central business districts (CBD). Density-based factors were obtained using a 7×7 neighborhood filter (e.g., Shafizadeh-Moghadam and Helbich 2015; Tayyebi et al. 2016a), accounting for the density of built-up areas and the availability of both open land and crop land. This filter provides the total number of cells, which represent the investigated variable within a specified radius of the central cell, and yields a value between 0 and 49. To match the spatial resolution of the Landsat data and to avoid inconsistency among the

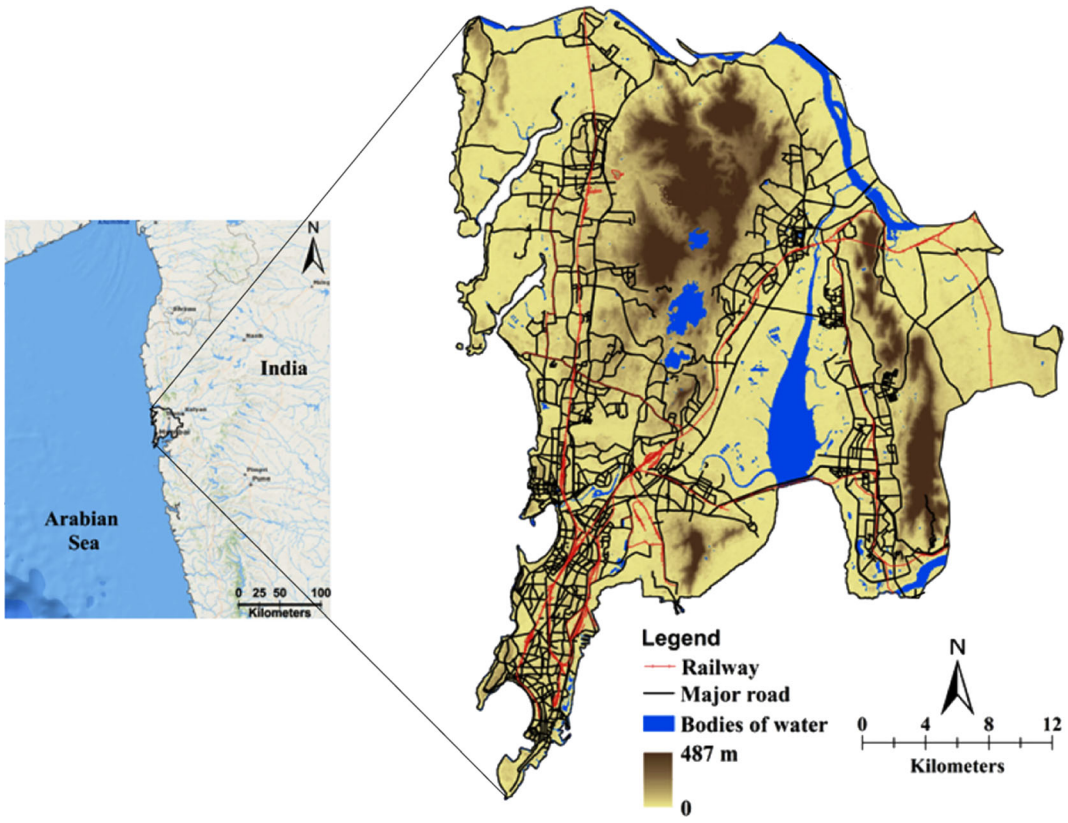

Fig. 1 The study area

variables, explanatory factors were prepared with a 30m cell resolution (see Figure S1 in the supplemental information). Slope and height were computed on the basis of the DEM. Finally, water bodies, wetlands, national parks, and built-up areas in 2001 were not further considered in subsequent analyses and were labeled as "excluded areas," indicating physical or legal restrictions to further growth.

Land cover change models

Weight of evidence

Weight of evidence (WOE) is a Bayesian approach that examines the probability of an event occurring, here, urban growth (U) in relation with specific factors (F), based on prior and posterior probabilities. The prior probability, $P\{U\}$, is the probability of a built-up area occurring across the entire region, which is obtained by dividing the ratio of the area or total number of built-up cells $N\{U\}$, by the area or number of all cells (built-up and non-built-up) $T\{U\}$ (Eq. 1).

$$P\{U\} = \frac{N\{U\}}{T\{U\}} \tag{1}$$

The posterior probability is the probability that urban change will take place in the presence or absence of a particular factor. As described in Lee and Choi (2004), utilizing an existing factor, the

Table 1 Data description

Data	Source	Date	Spatial resolution
Landsat (TM, ETM, ETM+)	US Geological Survey, German Aerospace Centre	1990, 2001, 2010	30 m
Digital elevation model	ASTER (NASA)	2009	30 m
Transportation network	OpenStreetMap	2011	–

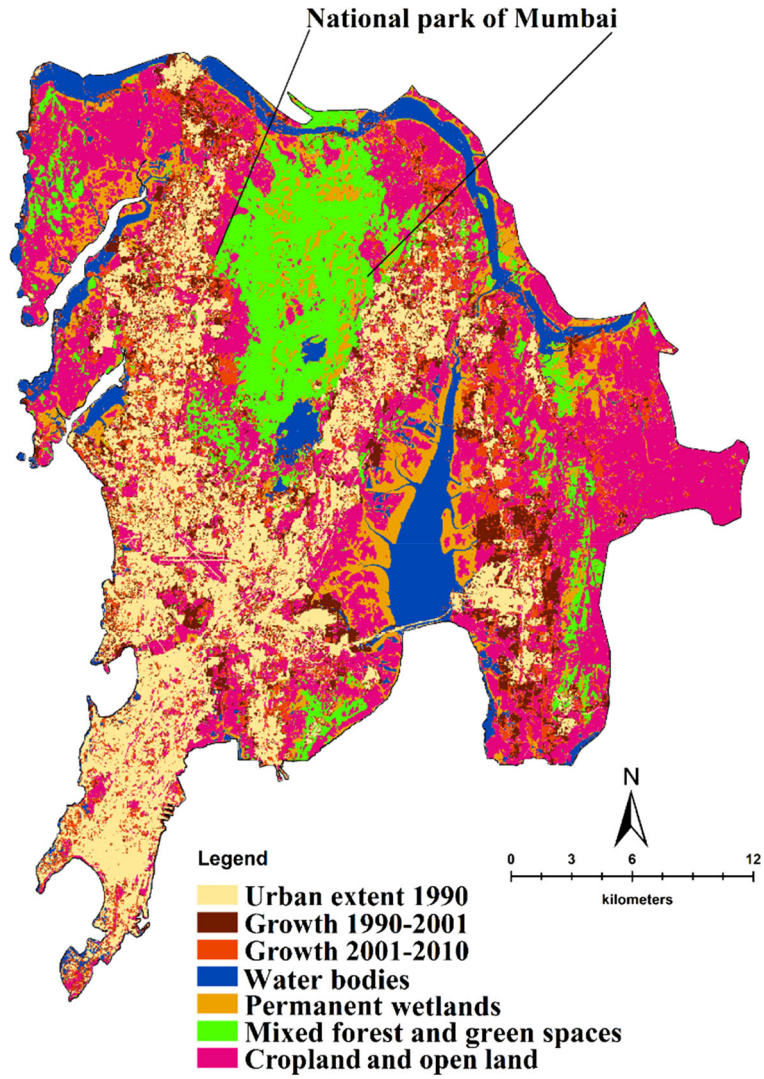

Fig. 2 Urban growth from 1990 to 2010

Table 2 Explanatory variables of urban growth in 2001

Covariates	Range	Mean	SD
Distance to the CBD (m)	0–33,000	18,069	7482
Distance to main roads (m)	0–4330	597	748
Distance to built-up areas (m)	0–4300	536	764
Distance to water bodies (m)	0–23,000	2636	2150
Distance to wetlands (m)	0–22,000	4188	3775
Distance to forests (m)	0–16,000	2793	2698
Density of built-up areas (cells per area)	0–49	29	15
Density of avail. Open/arable land (cells per area)	0–49	24.8	16.8
Height (m)	0–487	39	60
Slope (degree)	0–60	5.86	6.01

posterior probability can be calculated as follows (Eq. 2):

$$P\{U|F\} = \frac{P\{U \cap F\}}{P\{F\}} = P\{U\}\frac{P\{F|U\}}{P\{F\}} \quad (2)$$

and when the factor is absent as (Eq. 3)

$$P\left\{U|\overline{F}\right\} = \frac{P\left\{U \cap \overline{F}\right\}}{P\left\{\overline{F}\right\}} = P\{U\}\frac{P\left\{\overline{F}|U\right\}}{P\left\{\overline{F}\right\}} \quad (3)$$

$P\{U|F\}$ and $P\{U|\overline{F}\}$ are respectively the posterior probability of U based on the presence (F) and the absence of a factor (F). Similarly, $P\{F\}$ and $P\{\overline{F}\}$ are the prior probabilities of being inside or outside the predictor domain. Figure 3, for example, illustrates the spatial relationship between the occurrence of built-up areas in the presence of height variable. Each predictor variable yields both a positive and negative weighting. The positive weighting in the presence of a factor (Eq. 4) and negative weighting (Eq. 5) in the absence of a factor can be calculated as follows:

$$W^{+} = \log_e \frac{P\{F|U\}}{P\left\{F|\overline{U}\right\}} \quad (4)$$

$$W^{-} = \log_e \frac{P\left\{\overline{F}|U\right\}}{P\left\{\overline{F}|\overline{U}\right\}} \quad (5)$$

The positive weight of a factor is the logarithmic ratio of the probability of F (presence of the factor) if urban change occurs (U) to the probability of F if no urban change occurs ($\overline{U}$). The negative weight of evidence is the logarithmic ratio of the probability of not F (absence of the factor) in the presence of urban change to the probability of not F when the event (urban change) is absent (U).

Since WOE was developed for binary assessments, each factor is classified based on the standard deviation (SD) criterion, which is based on the data distribution and prevent arbitrary classification schemes. The number of SDs here is assigned based on the range of factors. The correlation between sub-classes (number of SDs) for each factor and an urban occurrence is measured using the contrast index ($C = W^{+} - W^{-}$). The maximum contrast is used as a cut-off point. Finally, all the weighted factors are combined as follows (Eq. 6):

$$PC = \sum Fw \quad (6)$$

where Fw is W^{+} and W^{-} for each binary map. For WOE, binary maps were created after conducting a conditional independence test (Bonham-Carter et al. 1994; Lee and Choi 2004) (see Figure S2 in the supplemental information). Once the variables fulfilled the condition of being independent in a pairwise comparison, they were entered into the WOE model and considered for inclusion in the final combination to create the TIM. All the WOE calculations were performed using Microsoft Excel.

Fuzzy multi-criteria evaluation

An alternative to compute a TIM is to couple MCE and fuzzy membership functions (Eastman 2006). Fuzzy membership functions were utilized to make the variables comparable and dimensionless, while MCE was employed to assess the suitability of locations based on the predefined criteria. Appropriate fuzzy membership functions, including monotonically increasing and decreasing S-shaped, J-shaped, and linear functions, were employed (Mitsova et al. 2011). For example, in the model using a J-shaped membership function, areas within 50-m buffer of roads are assumed to be the most appropriate for urban growth, and beyond this buffer zone, suitability decreases until 1 km but never reaches zero (Araya and Cabral 2010). After that, relative importance of each criterion was determined by expert knowledge and experimental LCC analysis.

Belonging to the MCE methods, using analytic hierarchy process (AHP), the relative importance of factors were assigned and weighted in a pairwise comparison manner. In doing so, all the factors were compared using a pairwise matrix, then using a numerical scale ranging from 0 to 9, the factors were assessed with respect to each other. Next, using the pairwise matrix, the weight of each factor is calculated. Finally, consistency ratio between the factors and the obtained weights were calculated (Chowdhury and Maithani 2014). By overlying the weighted factors, a suitability map was generated showing the potential of each cell for future urban growth. For the implementation of FAHP, we used the IDRISI Taiga software.

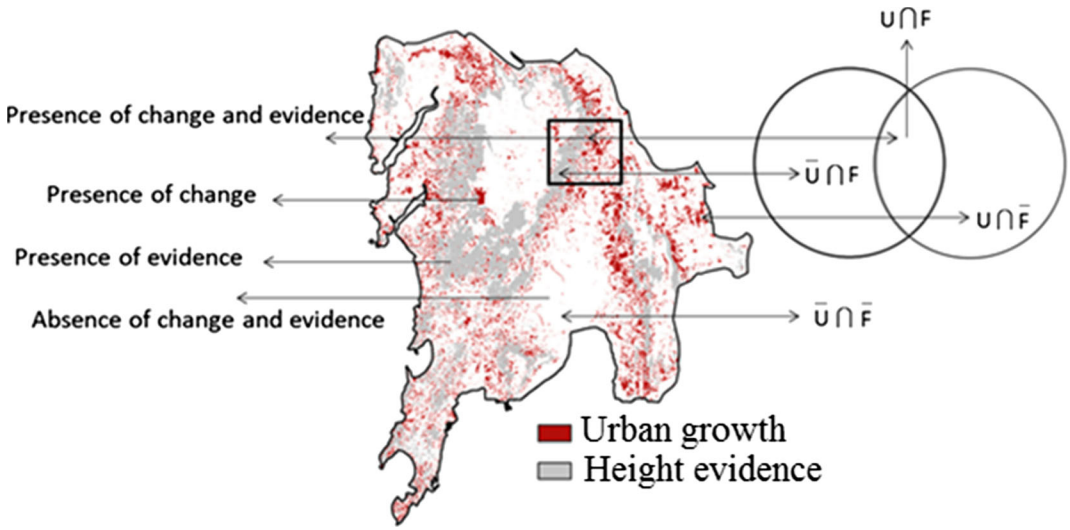

Fig. 3 Spatial relationship between the occurrence of built-up areas in the presence of height variable in a WOE model

Artificial neural networks

ANNs are machine learning techniques capable to capture non-linear associations between urban change and the underlying drivers (Bishop 1995; Pijanowski et al. 2002, 2014; Shafizadeh-Moghadam et al. 2015). This is done using non-linear functions and weightings that are applied to units within the network (Pijanowski et al. 2009; Tayyebi and Pijanowski 2014). The multi-layer perceptron, as used in this paper, typically includes three different layers, the input, hidden, and output layer. These three layers are connected to each other in a feed-forward manner (Hagenauer and Helbich 2012; Azari et al. 2016). A suitability map is then produced using an activation function that applies a neuron which receives the weighted output from connected neurons of the preceding layers. Back propagation is a commonly applied training algorithm (Bishop 1995). ANNs start by randomly assigning the weightings and calculating the model error, after which the cycle continues until a terminating criterion is met (Bishop 1995). If the level of performance is not attained, the error is distributed back to the neurons in the hidden layer, thus allowing them to update the weightings and mitigate the error. Once the training step is successfully completed through validation runs (e.g., using a dataset exclusively set aside), the model can be generalized for other data sets.

To set up the ANN model, the urban growth map was coded as 1 if the cells from any other land cover classes had become built-up, or otherwise as 0. The explanatory variables were standardized between 0 and 1 using maximum values. The data set was split into a training set (70%) and a testing set (30%). The former was used to train the ANN model, represented as a feed-forward multi-layer perceptron. The latter was used to evaluate the model performance. We applied a sigmoid function for the hidden layer and a linear function for the output layer as an activation function (Pijanowski et al. 2000; Lin et al. 2011; Bagheri et al. 2015). The best transfer function for the hidden layer is hyperbolic tangent sigmoid function while the best transfer function for the output layer is a linear one. These functions are well established in the literature (e.g., Bagheri et al. 2015).

Model evaluation

Assessing model performance is a critical task in LCC modeling (Tayyebi et al. 2014). Observed maps of 1990 and 2001 were used for model calibration, while the year 2010 was used for validation. One of the most common assessment metrics used to evaluate LCC models is ROC (Pontius and Schneider 2001). ROC compares transition index maps with the observed binary map, accounting for coincidences among the observed changes and their locations in the index maps (Mas et al. 2013). Furthermore, we used figure of merit (FOM), producer accuracy (PA), overall accuracy (OA), as well as the average spatial deviation distance (ASDD) index for accuracy assessment (Pontius et al. 2008). Formulas and description of these metrics are furnished in Table 3.

Results

Model evaluation of transition index maps

The classification of the Landsat images resulted in five classes including built-up areas, water bodies, wetlands, forest and green spaces, and open land and crop lands. Accuracy assessment of the classified maps using the Kappa index showed 84% and 86% agreement. As seen from Table 4, net change exploration of the land cover classes shows a gain for the built-up class and a loss for the remaining classes. From 1990 to 2010, built-up areas doubled while the least affected land cover class was forest and green spaces with a loss of 9.5%. Water bodies, wetlands, and open land and croplands decreased by 14.5%, 15.9%, and 36.8%. The classified maps were aggregated to a binary map (i.e., 1 for built-up areas and 0 for other classes). This binary map together with the explanatory variables (Table 2) was considered for the training of ANN, WOE, and FAHP models.

The outcome of ANN, WOE, and FAHP was a TIM, indicating the suitability of future urban growth. Figure 4 shows the TIMs generated by the three models; the brighter colors refer to a higher suitability of change to built-up areas. To allocate the amount of cells that were going to change, we used a top-down procedure (e.g., Pijanowski et al. 2002; Tayyebi et al. 2014; Shafizadeh-Moghadam et al. 2017a, b). The top-down approach starts by searching for the highest values in the transition probability map and stops when the aggregate number of cells that should change had been allocated (Pijanowski et al. 2002). Cells that converted to urban between 2000 and 2010 were allocated to the cells with the highest value from the TIM. As a result, three simulated maps were generated and to assess how accurate each map could correctly allocate the spatial location of changes, all the maps were overlaid with the observed maps. Figure 5 provides a visual assessment of how well the selected models performed in simulating urban growth. Visual interpretation indicates that the ANN model was able to generate more accurate and spatially explicit simulations while the other two models resulted in less accurate simulated maps.

Model performance was carried using (i) the ROC statistic, (ii) a set of statistical measures (Table 3), and (iii) by overlaying the simulated and reference maps to quantify the amount of correctness and erroneous predicted cells. The area under the ROC curve was 85% for ANN, 75% for FAHP, and 73% for WOE. From 2001 to 2010, a total amount of 8323 ha converted to the built-up class. The implemented ANN model correctly predicted the largest amount of spatial location of these changed cells (60.62%), while FAHP had the lowest accuracy (41.93%). From the remaining 35,744 ha, the number of correctly predicted unchanged cells was 90.84% for the ANN while FAHP had the lowest accuracy with 84.05%. Figure 5 shows the locations where the simulated map matches observed changes during 2001–2010, and also areas in which the model fails to classify changes correctly. True positive refers to cells classified as "changed" in both the reference and simulated maps. False positives indicate cells that wrongly predicted as changed by the model. Similarly, false negatives indicate locations classified as unchanged when they were found in the reference map to have actually changed. True negatives refer to locations where both the simulated and observed maps classify as places that remained unchanged. These changes were obtained by overlaying the simulated map of 2010 with a reference map of 2001 and 2010.

Following Pontius et al. (2008), we considered referenced maps in 2001 and 2010 as well as a predicated map for 2010, to characterize the dynamics of change, model behavior, and spatial prediction accuracy. The FOM, which ranges from 0% (no overlap) to 100% (full match between observed and predicted changes), indicates the amount of spatial overlap between the predicted model and the observed change. The FOM was 43.5%, 25.8%, and 32.5% for the ANN, FAHP, and WOE. The PA was 60.6%, 41.9%, and 48.9% for the ANN, FAHP, and WOE. The OA was 85.1%, 75.9%, and 78.1% for the ANN, FAHP, and WOE. The ASDD was 20.1%, 29.2%, and 24.2% for the ANN, FAHP, and WOE, respectively. Figure 6 depicts the statistics for the selected measures. In conclusion, the model assessment indicated that the ANN performed best and thus we used the ANN to simulate future growth patterns for the years during 2010–2020 and 2020–2030.

Simulating urban growth using the ANN

As displayed in Fig. 7, most urban growth will occur within the vicinity of existing areas. Between 2020 and 2030, Mumbai is expected to experience an increase in built-up density towards its eastern and northern regions, where less developed, open, and agricultural land will be more available than in other parts. Moreover, developments in the middle and northern half of

Table 3 Spatial metrics for model evaluation (Pontius et al. 2008)

Metrics	Formula	Description
Figure of merit (FOM)	$\mathrm{FOM} = \frac{B}{A+B+C+D}$	FOM is the intersection of the observed and predicted change divided by the union of the observed and predicted change.
Producer accuracy (PA)	$\mathrm{PA} = \frac{B}{A+B+C}$	PA indicates the proportion of pixels that the model predicts accurately as change, given that the reference maps indicate observed change.
Overall accuracy (OA)	$\mathrm{OA} = \frac{B+E}{A+B+C+D+E}$	OA provides the overall agreement between the reference and predicted maps.
Average spatial deviation distance (ASDD)	$\mathrm{ASDD}(A,P) = \frac{1}{K}\sum_{k=1}^{K} D(A_k, A_k.\mathrm{nearest}(P))$	ASDD measures Euclidean distance between the locations of observed and predicted changes.

For the FoM, PA, and OA, A is the proportion of error cells due to the observed change—predicted as persistence, B is the proportion of correct cells due to observed change—predicted as change, C is the proportion of error cells due to observed change modeling—predicted as a wrong gaining category, D is the proportion of error cells due to observed persistence—predicted as change, and E denotes the area of correct cells due to observed persistence—predicted as persistence. Also for ASDD, A_k is the observed built-up cell, P is the map layer of predicted built-up areas, and A_k.nearest(P) denotes the nearest predicted location to A_k, and K is the number of actual built-up cells.

Mumbai will occur around environmentally sensitive areas including the national park which will be increasingly surrounded by buildings. Considering the relationship between physical growth and environmental issues (e.g., Mumbai is among the world's most polluted cities), continuous urban growth processes can be seen as a serious threat potentially having disastrous consequences for the entire region.

Discussion

Accuracy assessment of LCC models has great importance (Pontius et al. 2008). It is critical for environmental modelers that couple their models with LCC models to assess the future consequence of LCC on other environmental dimensions. For example, Tayyebi et al. (2015) assessed the future consequence of agriculture expansion and urbanization on water quality. However, LCC models can be clustered and compared from different perspectives. For example, from the viewpoint of the role of expert knowledge, the LCC models can be considered as data-driven and knowledge-based approaches, while from the view point of statistical assumptions, they could take either the class of parametric or non-parametric methods (Tayyebi et al. 2014). Furthermore, the simulated maps created by each model are different. Thus, the models should be compared and evaluated from the side of spatial accuracy of the simulated maps. Different pattern of simulated maps could be attributed to the integrated functions of each model, their associated assumptions, and the type of cell allocation strategies (Shafizadeh-Moghadam et al. 2017b).

This paper showed that the ANN was the most accurate, and the WOE was the least accurate model for simulating urban growth. WOE has a major drawback that the model requires each input layer to be presented as a binary map to the model. As described in data

Table 4 LCC in Mumbai between 1990 and 2010 (in ha and %)

Land cover class	1990 (ha)	2001 (ha)	2010 (ha)	1990–2001 (%)	2000–2010 (%)
Built-up areas	18,455	25,498	35,607	38.2	39.7
Water bodies	8271	7805	7100	−5.6	−9.0
Permanent wetlands	9734	9086	8187	−6.7	−9.9
Mixed forest and green spaces	11,418	11,057	10,329	−3.2	−6.6
Open land and cropland land	39,682	34,124	26,353	−14.0	−22.8

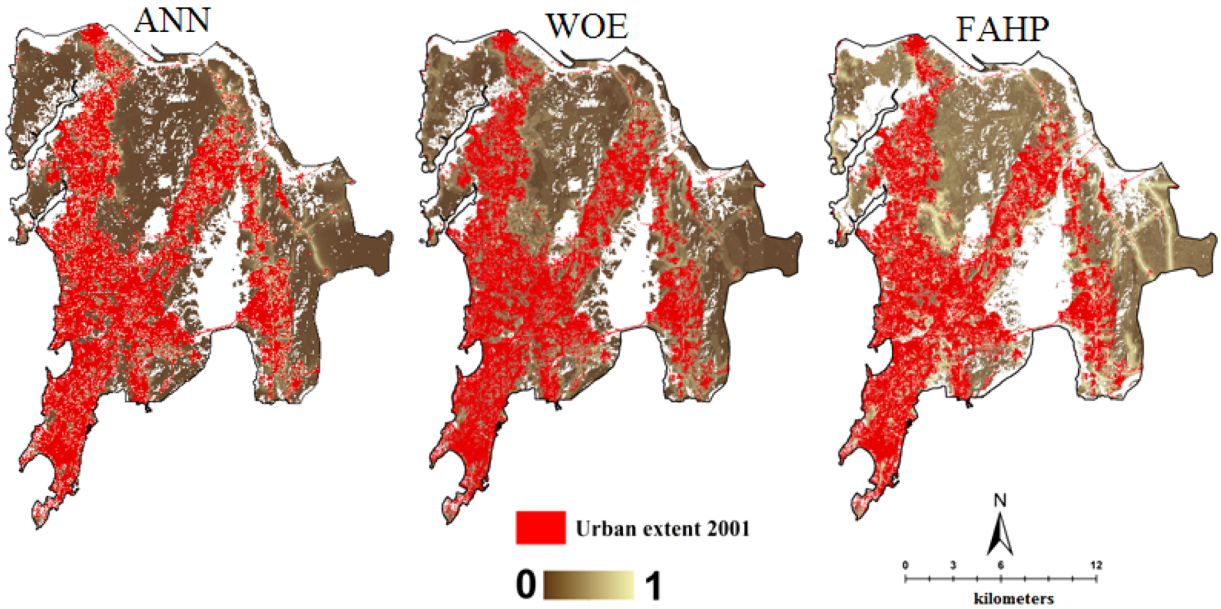

Fig. 4 TIMs produced by ANN, WOE, and FAHP

section, some of our driving forces (input layers) were continuous that was required to be discretized beforehand. Optimum discretization of input layers is not an easy task. In this paper, we simply considered the statistical distribution of the layers and used standard deviation for layer discretization. Integration of the optimum discretization algorithms might heavily increase the quality of the TIM produced by the WOE and finally could boost the positional accuracy of the simulated maps.

In a comparative view, each of the utilized models exhibited specific strengths and limitations. ANN is an

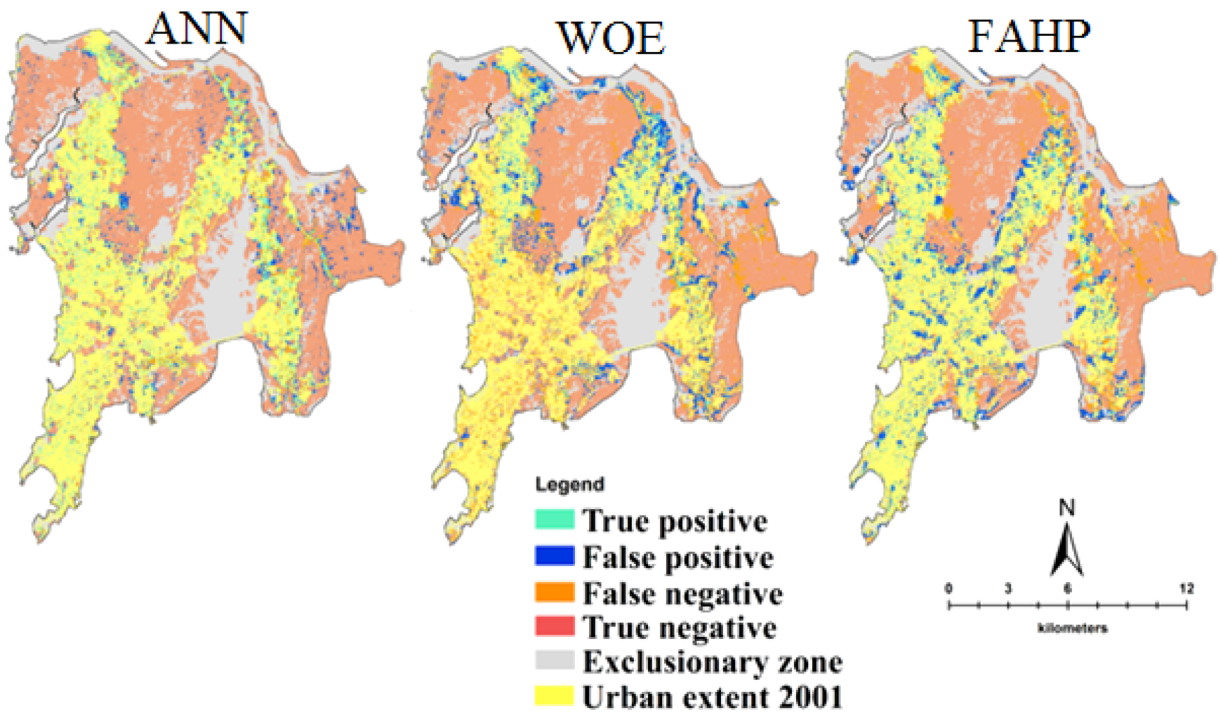

Fig. 5 Spatial overlay of the ANN, WOE, and FAHP model predictions and observed changes between 2001 and 2010

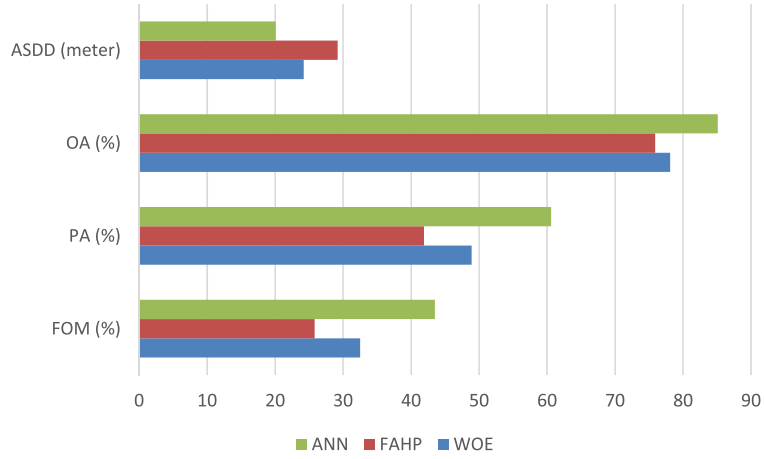

Fig. 6 Comparing the accuracy of ANN, FAHP, and WOE

appropriate machine learning model for simulating urban growth, in which various spatial drivers stimulate land transformation in a non-linear way (Tayyebi et al. 2014). ANN relaxes the statistical assumptions and

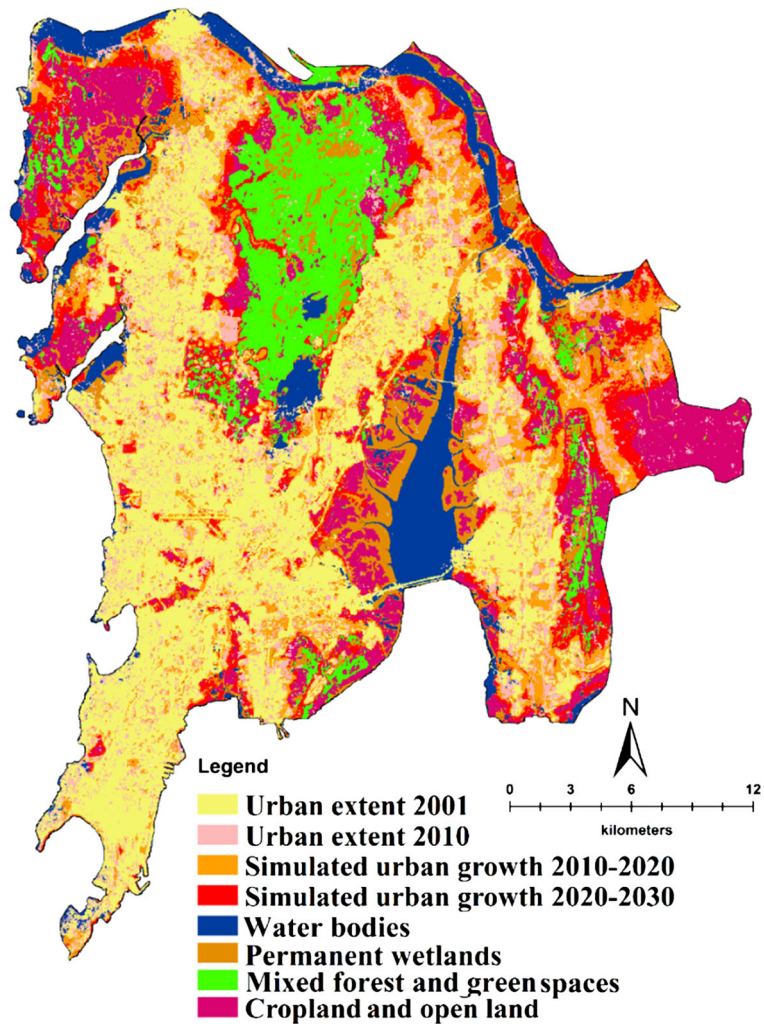

Fig. 7 Observed and predicted urban growth using the ANN model from 2001 to 2030

properly handles data with any distribution (Xie et al. 2008). Paliwal and Kumar (2009) highlighted the advantages of ANN over the statistical models which automatically approximated any non-linear function. FAHP, on the other hand, relies heavily on expert knowledge regarding the nature and importance of the drivers. Due to the complexity of land transformations taking place in cities, it seems doubtful that experts would have a detailed enough understanding of the underlying spatial processes to effectively apply FAHP (i.e., the selection of the membership function and factor weighting), which further supports the application of ANN. In addition, data-driven approaches reduce the likelihood of biased or incorrect decisions being made by experts (Hosseinali and Alesheikh 2008), while the knowledge-based systems such as MCE can explicitly reveal the magnitude and relationship between each driver causing change in urban areas, by allowing experts to contribute.

Findings of this paper correspond to those reported by Lin et al. (2011), who observed a higher ROC for the ANN, when compared to logistic and auto-logistic regression models. Similar conclusions were drawn by Tayyebi and Pijanowski (2014) and Tayyebi et al. (2014), who also reported more effective prediction capabilities for ANN than for tree-based models and multivariate adaptive regression splines.

This study can be improved in the future by using additional socio-economic data, including population density and other indicators on the microscale. It is expected that such drivers would increase the models' performances. However, such a lack of data is common in most developing countries, where fine resolution census data are often unavailable. Future research can also develop hybrid models on the basis of knowledge-based and data-driven approaches, which benefit from the complementing advantages of different models (Azari et al. 2016). Such integrations would allow capturing various drivers and dynamics behind urban transformations, as well as to set up more precise urban growth simulation models. Finally, the incorporation of the spatio-temporal variability of land cover drivers might also improve the models' spatial accuracy.

Conclusion

This study compared three TIMs produced from data-driven models (i.e., ANN and WOE) and knowledge-based approaches (i.e., FAHP). On the one hand, the ROC value indicated that the ANN has the ability to generate a more accurate TIM for urban growth than WOE and FAHP. On the other hand, the statistical measures quantifying the spatial accuracy of the simulations clearly confirmed the suitability of ANN. Using the most accurate approach, ANN was then calibrated to simulate urban growth for the years 2020 and 2030 in Mumbai. The projected maps generated by the ANN were indicative of continuous urban growth through both infilling of the already built-up areas and rapid expansion to the north and east, into new areas. In this case, the simulated maps are beneficial for urban planners and decision makers by knowing how the extent and pattern of urban development in Mumbai would be expected in the next two decades for directing future urban development, environment-oriented plans. In conclusion, due to their high accuracy and their robustness, ANNs are recommended for future studies to generate TIMs instead of WOE and FAHP.

References

Agterberg, F. P., & Cheng, Q. (2002). Conditional independence test for weights-of-evidence modeling. *Natural Resources Research, 11*(4), 249–255.

Almeida, C. M., Batty, M., Vieira-Monteiro, A. M., Câmara, G., Soares-Filho, B. S., Cerqueira, G. C., & Pennachin, C. L. (2003). Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation. *Computers, Environment and Urban Systems, 27*(5), 481–509.

Almeida, C. M., Gleriani, J. M., Castejon, E. F., & Soares-Filho, B. S. (2008). Using neural networks and cellular automata for modelling intra-urban land-use dynamics. *International Journal of Geographical Information Science, 22*(9), 943–963.

Araya, Y. H., & Cabral, P. (2010). Analysis and modelling of urban land cover change in Setúbal and Sesimbra, Portugal. *Remote Sensing, 2*, 1549–1563.

Azari, M., Tayyebi, A., Helbich, M., & Reveshty, M. A. (2016). Integrating cellular automata, artificial neural network, and fuzzy set theory to simulate threatened orchards: application to Maragheh, Iran. *GIScience & Remote Sensing, 53*(2), 183–205.

Bagheri, M., Mirbagheri, S. A., Ehteshami, M., & Bagheri, Z. (2015). Modeling of a sequencing batch reactor treating municipal wastewater using multi-layer perceptron and radial basis function artificial neural networks. *Process Safety and Environmental Protection, 93*, 111–123.

Bishop, C. M. (1995). *Neural networks for pattern recognition*. New York: Oxford University Press.

Bonham-Carter, G. F. (1994). *Geographic Information Systems for Geoscientists, modelling with GIS*. Oxford: Pergamon Press.

Chowdhury, P. R., & Maithani, S. (2014). Modelling urban growth in the Indo-Gangetic plain using nighttime OLS data and cellular automata. *International Journal of Applied Earth Observation and Geoinformation, 33*, 155–165.

Eastman, J. R. (2006). *IDRISI Andes*. Worcester: Clark University.

Fathizad, H., Rostami, N., & Faramarzi, M. (2015). Detection and prediction of land cover changes using Markov chain model in semi-arid rangeland in western Iran. *Environmental Monitoring and Assessment, 187*(10), 1–12.

Feng, Y. (2017). Modeling dynamic urban land-use change with geographical cellular automata and generalized pattern search-optimized rules. *International Journal of Geographical Information Science*, 1–22.

Feng, Y., & Liu, Y. (2016). Scenario prediction of emerging coastal city using CA modeling under different environmental conditions: a case study of Lingang New City, China. *Environmental Monitoring and Assessment, 188*(9), 540.

Feng, Y., Yang, Q., Hong, Z., & Cui, L. (2016). Modelling coastal land use change by incorporating spatial autocorrelation into cellular automata models. *Geocarto International*, 1–19.

Foody, G. M. (2009). Sample size determination for image classification accuracy assessment and comparison. *International Journal of Remote Sensing, 30*(20), 5273–5291.

Grekousis, G., Manetos, P., & Photis, Y. N. (2013). Modelling urban evolution using neural networks, fuzzy logic and GIS: the case of the Athens metropolitan area. *Cities, 30*, 193–203.

Hagenauer, J., & Helbich, M. (2012). Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks. *International Journal of Geographical Information Science, 26*(6), 963–982.

Hansen, A. J., DeFries, R. S., & Turner, W. (2012). Land use change and biodiversity. In Land Change Science (pp. 277–299). Springer Netherlands.

Helbich, M., Amelunxen, C., Neis, P., & Zipf, A. (2012). Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary geodata. In T. A. Jekel, A. Car, J. Strobl, & G. Griesebner (Eds.), *Geospatial Crossroads @ GI_Forum 2012. Proceedings of the Geoinformatics Forum Salzburg* (pp. 24–33). Heidelberg: Wichmann.

Hosseinali, F., & Alesheikh, A. A. (2008). Weighting spatial information in GIS for copper mining exploration. *American Journal of Applied Sciences, 5*(9), 1187–1198.

Hu, Z., & Lo, C. P. (2007). Modelling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems, 31*, 667–688.

Kamusoko, C. (2012). Markov–cellular automata in geospatial analysis. In progress in Geospatial Analysis (pp. 107–124). Springer Japan.

Lee, S., & Choi, J. (2004). Landslide susceptibility mapping using GIS and the weight-of-evidence model. *International Journal of Geographical Information Science, 18*(8), 789–814.

Ligmann-Zielinska, A., & Jankowski, P. (2010). Exploring normative scenarios of land use development decisions with an agent-based simulation laboratory. *Computers, Environment and Urban Systems, 34*, 409–423.

Lin, Y. P., Chu, H. J., Wu, C. F., & Verburg, P. H. (2011). Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land use change modelling–a case study. *International Journal of Geographical Information Science, 25*(1), 65–87.

Mas, J. F., Soares Filho, B., Pontius, R. G., Farfán Gutiérrez, M., & Rodrigues, H. (2013). A suite of tools for ROC analysis of spatial models. *ISPRS International Journal of Geo-Information, 2*(3), 869–887.

Mitsova, D., Shuster, W., & Wang, X. (2011). A cellular automata model of land cover change to integrate urban growth with open space conservation. *Landscape and Urban Planning, 99*(2), 141–153.

Munshi, T., Zuidgeest, M., Brussel, M., & van Maarseveen, M. (2014). Logistic regression and cellular automata-based modelling of retail, commercial and residential development in the city of Ahmedabad, India. *Cities, 39*, 68–86.

Pacione, M. (2006). City profile–Mumbai. *Cities, 23*(3), 229–238.

Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: a review of applications. *Expert Systems with Applications, 36*(1), 2–17.

Park, S., Jeon, S., Kim, S., & Choi, C. (2011). Prediction and comparison of urban growth by land suitability index mapping using GIS and RS in South Korea. *Landscape and Urban Planning, 99*(2), 104–114.

Patel, A., Koizumi, N., & Crooks, A. (2014). Measuring slum severity in Mumbai and Kolkata: a household-based approach. *Habitat International, 4*, 300–306.

Pijanowski, B.C., Gage, S.H., & 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*, Larry Harris and James Sanderson eds.

Pijanowski, B. C., Brown, D. G., Shellito, B. A., & Manik, G. A. (2002). Using neural networks and GIS to forecast land use changes: a land transformation model. *Computers, Environment and Urban Systems, 26*(6), 553–575.

Pijanowski, B. C., A. Tayyebi, M. R. Delavar, & M. J. Yazdanpanah. (2009). Urban expansion simulation using geospatial information system and artificial neural networks. *International Journal of Environmental Research*, 3(4).

Pijanowski, B. C., Tayyebi, A., Doucette, J., Pekin, B. K., Braun, D., & Plourde, J. (2014). A big data urban growth simulation at a national scale: Configuring the GIS and neural network based land transformation model to run in a high performance computing (HPC) environment. *Environmental Modelling and Software, 5*, 250–268.

Pontius, R. G., & Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment, 85*(1), 239–248.

Pontius, R. G., Boersma, W., Castella, J. C., Clarke, K., de Nijs, T., Dietzel, C., & Verburg, P. H. (2008). Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science, 42*(1), 11–37.

Sangermano, F., Eastman, J. R., & Zhu, H. (2010). Similarity weighted instance-based learning for the generation of transition potentials in land use change modeling. *Transactions in GIS, 14*(5), 569–580.

Shafizadeh-Moghadam, H., & Helbich, M. (2013). Spatiotemporal urbanization processes in the megacity of

Mumbai, India: a Markov chains-cellular automata urban growth model. *Applied Geography, 40*, 140–149.

Shafizadeh-Moghadam, H., & Helbich, M. (2015). Spatiotemporal variability of urban growth factors: a global and local perspective on the megacity of Mumbai. *International Journal of Applied Earth Observation and Geoinformation, 35*, 187–198.

Shafizadeh-Moghadam, H., Hagenauer, J., Farajzadeh, M., & Helbich, M. (2015). Performance analysis of radial basis function networks and multi-layer perceptron networks in modeling urban change: a case study. *International Journal of Geographical Information Science, 29*(4), 606–623.

Shafizadeh-Moghadam, H., A. Asghari, M. Taleai, M. Helbich, & A. Tayyebi. (2017a). Sensitivity analysis and accuracy assessment of the land transformation model using cellular automata. *GIScience & Remote Sensing*, 1–18.

Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., & Taleai, M. (2017b). Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. *Computers, Environment and Urban Systems, 64*, 297–308.

Tayyebi, A., & Jenerette, G. D. (2016). Increases in the climate change adaption effectiveness and availability of vegetation across a coastal to desert climate gradient in metropolitan Los Angeles, CA, USA. *Science of the Total Environment, 548*, 60–71.

Tayyebi, A., & Pijanowski, B. C. (2014). Modeling multiple land use changes using ANN, CART and MARS: Comparing tradeoffs in goodness of fit and explanatory power of data mining tools. *International Journal of Applied Earth Observation and Geoinformation, 28*, 102–116.

Tayyebi, A., Delavar, M. R., Yazdanpanah, M. J., Pijanowski, B. C., Saeedi, S., & Tayyebi, A. H. (2010). A spatial logistic regression model for simulating land use patterns: a case study of the Shiraz Metropolitan area of Iran. In *Advances in earth observation of global change* (pp. 27–42). Springer Netherlands.

Tayyebi, A., Pijanowski, B. C., & Tayyebi, A. H. (2011). An urban growth boundary model using neural networks, GIS and radial parameterization: an application to Tehran, Iran. *Landscape and Urban Planning, 100*(1), 35–44.

Tayyebi, A., Pijanowski, B. C., Linderman, M., & Gratton, C. (2014). Comparing three global parametric and local non-parametric models to simulate land use change in diverse areas of the world. *Environmental Modelling and Software, 59*, 202–221.

Tayyebi, A., Pijanowski, B. C., & Pekin, B. K. (2015). Land use legacies of the Ohio River basin: using a spatially explicit land use change model to assess past and future impacts on aquatic resources. *Applied Geography, 57*, 100–111.

Tayyebi, A., Tayyebi, A. H., Arsanjani, J. J., Moghadam, H. S., & Omrani, H. (2016a). FSAUA: A framework for sensitivity analysis and uncertainty assessment in historical and forecasted land use maps. *Environmental Modelling & Software, 84*, 70–84.

Tayyebi, A., Meehan, T. D., Dischler, J., Radloff, G., Ferris, M., & Gratton, C. (2016b). SmartScape™: a web-based decision support system for assessing the tradeoffs among multiple ecosystem services under crop-change scenarios. *Computers and Electronics in Agriculture, 121*, 108–121.

Thapa, R. B., & Murayama, Y. (2011). Urban growth modelling of Kathmandu metropolitan region, Nepal. *Computers, Environment and Urban Systems, 35*(1), 25–34.

Van Genderen, J. L., & Lock, B. F. (1977). Testing land-use map accuracy. *Photogrammetric Engineering and Remote Sensing, 43*(9).

Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology, 1*(1), 9–23.

Zhu, Z., Liu, L., Chen, Z., Zhang, J., & Verburg, P. H. (2010). Land-use change simulation and assessment of driving factors in the loess hilly region—a case study as Pengyang County. *Environmental Monitoring and Assessment, 164*(1–4), 133–142.