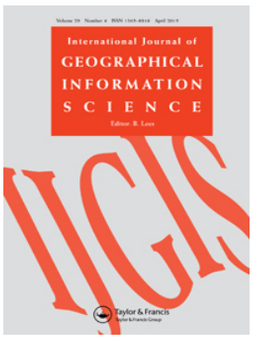

International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: www.tandfonline.com/journals/tgis20

Performance analysis of radial basis function networks and multi-layer perceptron networks in modeling urban change: a case study

Hossein Shafizadeh-Moghadam, Julian Hagenauer, Manuchehr Farajzadeh & Marco Helbich

To cite this article: Hossein Shafizadeh-Moghadam, Julian Hagenauer, Manuchehr Farajzadeh & Marco Helbich (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, DOI: 10.1080/13658816.2014.993989

To link to this article: https://doi.org/10.1080/13658816.2014.993989

International Journal of Geographical Information Science, 2015
Vol. 29, No. 4, 606–623, http://dx.doi.org/10.1080/13658816.2014.993989

Taylor & Francis
Taylor & Francis Group

Performance analysis of radial basis function networks and multi-layer perceptron networks in modeling urban change: a case study

Hossein Shafizadeh-Moghadam[a,b]*, Julian Hagenauer[a], Manuchehr Farajzadeh[b] and Marco Helbich[c]

[a]*Institute of Geography, Heidelberg University, Heidelberg, Baden-Württemberg, Germany;* [b]*Department of GIS and Remote Sensing, Tarbiat Modares University, Tehran, Iran;* [c]*Department of Human Geography and Spatial Planning, Utrecht University, Utrecht, The Netherlands*

(*Received 2 July 2014; final version received 25 November 2014*)

The majority of cities are rapidly growing. This makes the monitoring and modeling of urban change's spatial patterns critical to urban planners, decision makers, and environment protection activists. Although a wide range of methods exists for modeling and simulating urban growth, machine learning (ML) techniques have received less attention despite their potential for producing highly accurate predictions of future urban extents. The aim of this study is to investigate two ML techniques, namely radial basis function network (RBFN) and multi-layer perceptron (MLP) networks, for modeling urban change. By predicting urban change for 2010, the models' performance is evaluated by comparing results with a reference map and by using a set of pertinent statistical measures, such as average spatial distance deviation and figure of merit. The application of these techniques employs the case study area of Mumbai, India. The results show that both models, which were tested using the same explanatory variables, produced promising results in terms of predicting the size and extent of future urban areas. Although a close match between RBFN and MLP is observed, RBFN demonstrates higher spatial accuracy of prediction. Accordingly, RBFN was utilized to simulate urban change for 2020 and 2030. Overall, the study provides evidence that RBFN is a robust and efficient ML technique and can therefore be recommended for land use change modeling.

Keywords: radial basis function network; multi-layer perceptron network; urban change; spatial accuracy assessment; GIS

1. Introduction

Unprecedented increase in urban areas and concerns regarding the potentially devastating effects of rapid urban growth on environmental resources, climate and biodiversity, among others, have been highlighted in several studies (e.g., Lambin *et al.* 2003, Basawaraja *et al.* 2011, Triantakonstantis and Mountrakis 2012, Shafizadeh-Moghadam and Helbich 2015). This makes the monitoring and modeling of urban change's spatial patterns critical to urban planners, decision makers, and environment-protection activists. In addition, decision makers and urban planners require precise and detailed information about potential urban growth and land conversion in order to assess new development needs, their location, characteristics, as well as consequences of urban development (Jiang and Yao 2010).

*Corresponding author. Email: shafeezadeh@uni-heidelberg.de

Urban change is the result of a set of complex processes. This complexity comes from the fact that the many processes that influence Land Use Change (LUC) interact with each other and depend on the local cultural, socioeconomic, and biophysical context (Lambin *et al.* 2001). While the complexity and nonlinear nature of spatio-temporal drivers restrict the understanding of LUC in global change studies (Lambin et *al.* 2003, Kolb *et al.* 2013), understanding, conceptualizing, quantifying and measuring the mechanisms and the interaction among the drivers of urban change are required for a precise modeling of urban growth.

To date, various methods for modeling LUC and urban change are increasingly being used for various purposes. These methods encompass a wide spectrum, including cellular automata (CA) (e.g., Batty et *al.* 1999, Feng and Liu 2013), artificial neural networks (ANNs), (e.g., Pijanowski et *al.* 2002, Wang and Mountrakis 2011), agent-based models (ABMs) (e.g., Hosseinali et *al.* 2013), statistical methods such as logistic regression (Hu and Lo 2007, Shafizadeh-Moghadam and Helbich 2015), and other hybrids that integrate several methods, for example, ANNs and CA (Almeida *et al.* 2008). Proposed models commonly aim to produce accurate spatio-temporal LUC predictions, but their differences in assumptions, number of possible land use categories, spatial dependency, cross-scale linkages, and data requirements result in varying outputs and precision (Pontius and Malanson 2005). In this respect, an evaluation of 13 LUC models by Pontius *et al.* (2008) shows that for 12 of these models, the percentage of cells for which a change was correctly predicted was lower than the percentage of incorrectly classified cells.

While some methods are more suitable for understanding LUC's driving forces, for example, logistic regression (Hu and Lo 2007), other methods such as ANNs are more appropriate for prediction purposes (Pijanowski *et al.* 2014). The empirical contribution of each variable with regard to the restriction or stimulation of urban growth can be drawn from logistic regression methods, whereas ANNs are performed in a black box manner (Hu and Lo 2007, Lin *et al.* 2011, Wang and Mountrakis 2011). However, when prediction is the primary aim, it is essential to develop computational models that account for the intricate and nonlinear relations between drivers of LUC.

Despite the promising success and the unique key abilities of some LUC models, there exist other untested or less acknowledged techniques from the machine learning (ML) domain that could produce highly accurate predictions. Machine learning approaches, which are part of the artificial intelligence field, include statistical techniques like Bayesian classification, neural networks, and classification/regression trees (Kobler and Adamic 2000). They are considered to be more precise and efficient compared to conventional parametric algorithms (e.g., maximum likelihood) when dealing with large volumes of data and complex relationships (Foody 1995, Rogan *et al*. 2008).

ANNs are non-parametric techniques that employ interconnected mathematical nodes to form a network that can model complicated functional associations (Sha and Edwards 2007). ANNs have proven to be powerful tools for predicting urban growth pattern (Pijanowski *et al.* 2002, Almeida *et al.* 2008, Tayyebi et *al.* 2011, Wang and Li 2011, Rienow and Goetzke 2014). They have also been used to generate transition probability maps to support other LUC models such as CA (Li and Yeh 2002, Almeida *et al.* 2008, Wang and Li 2011). The ability of ANNs to model LUC was also compared with other empirical models. Lin *et al.* (2011) evaluated the performance of three models, that is, logistic regression, auto-logistic regression, and a model combining ANNs with the CLUE-S model (Conversion of Land Use and its

Effects, Verburg *et al.* 2002) for modeling LUC. They observed that the combination of ANNs with the CLUEs' model generates slightly better results than auto-logistic regression, and better results than logistic regression. In another study, Tayyebi and Pijanowski (2014) employed ANNs, classification and regression trees (CART), and multivariate adaptive regression splines (MARS) for simulating multiple land use types. They found that ANNs outperformed the two other techniques for simulating three land use classes (urban, agriculture, and forest). Recently, Rienow and Goetzke (2014) combined support vector machine (SVM) and Slope, Land use, Exclusion, Urban extent, Transportation and Hillshade (SLEUTH-S) urban CA in an attempt to improve the spatial accuracy of predicted changes. They found that the performance of SLEUTH-S in terms of allocation and quantity of change meaningfully improved when this model was combined with a SVM-based probability map. However, studies comparing the performance of other non-parametric models, including ANNs, have rarely been conducted.

Distinguishing features of ANNs, such as their ability to consider nonlinear relationships, make them an appropriate tool for modeling urban change. However, while having promising capacities in terms of prediction, especially compared to linear models such as logistic regression, ANNs are prone to overfitting and their results are typically difficult to interpret. More specifically, the overfitting means that the model performs better at the local level, but is less generalizable to unseen situations or other regions.

Faced with nonlinear processes, various types of ANNs such as MLP, RBFN, and self-organizing maps (SOM) can be calibrated for modeling LUC (Wang and Li 2011). Among them, MLP was successfully employed in several studies to support binary and multiclass LUC simulation (Pijanowski *et al.* 2002, Almeida *et al.* 2008, Lin *et al.* 2011). However, studies on the predictive ability of the other types of ANNs are quite limited. As of today, RBFN has mainly been used for purposes such as interpolation, time series analysis, remote sensing and mining (Potts and Broomhead 1991, Samanta 2010, Wang *et al.* 2014), as well as for the computation of conversion probabilities in LUC modeling (Wang and Li 2011). Simplicity and fast execution are two key features of RBFNs compared to feed-forward models such as MLP which are computationally expensive (Samanta 2010). In addition, the architecture and training algorithms associated with RBFNs are simpler and clearer than those of MLP implementations (Lin and Chen 2004) which are more complicated and could become trapped in a local minima solution (Wang and Li 2011).

To the best of our knowledge, with the exception of Wang and Li (2011) who used RBFN to extract transition probabilities and coupled the result with the CA to simulate multiple land use classes, there exists no study that uses RBFN to simulate urban growth. In this paper, we address this issue and we employ RBFN to model urban change using the procedure that was adopted by Pijanowski *et al.* (2002). They used a feed-forward back propagation MLP to model and assign the location of future change based on a set of drivers. As a result, a transition probability map is generated. Assuming a constant rate of change between the successive time stamps, the model allocates the future change to the cells with greatest transition probability values. The accuracy of the prediction is assessed by comparing it with a reference map using statistical metrics that include figure of merit (FoM), overall accuracy (OA), producer accuracy (PA), and average spatial distance deviation (ASDD).

The objectives of this study are threefold: conducting a performance analysis of RBFN and MLP; simulating urban change using RBFN and MLP; and then comparing and

contrasting their performance against each other using a corresponding reference (observed) map. The comparison of these two models helped to assess their performance in modeling urban change.

The remainder of this paper is organized as follows: Sections 2 and 3 describe the study area, data set and the main preprocessing steps. Section 4 describes the RBFN and MLP methods. Section 5 presents the results and discusses the main findings. Finally, Section 6 provides conclusions and future research directions to enhance the employed methods and results.

2. Study area

Mumbai is one of the largest megacities in India and one of the most populous cities in the world (Figure 1). It is the commercial and financial capital of India and has recorded high population growth rates over the years (Pethe *et al.* 2014). According to the United Nations' Department of Economic and Social Affairs (2012), the population of Mumbai will reach 27 million by 2025, while it was approximately 12.5 million in 2011.

The average elevation of the city is 39 meters above sea level and the highest points, as extracted from the digital elevation model (DEM) layer, reach over 480 m. However, the physical shape and morphology of Mumbai reduced the affluence of land, which led to a high density population in contemporary Mumbai (Pethe *et al.* 2014).

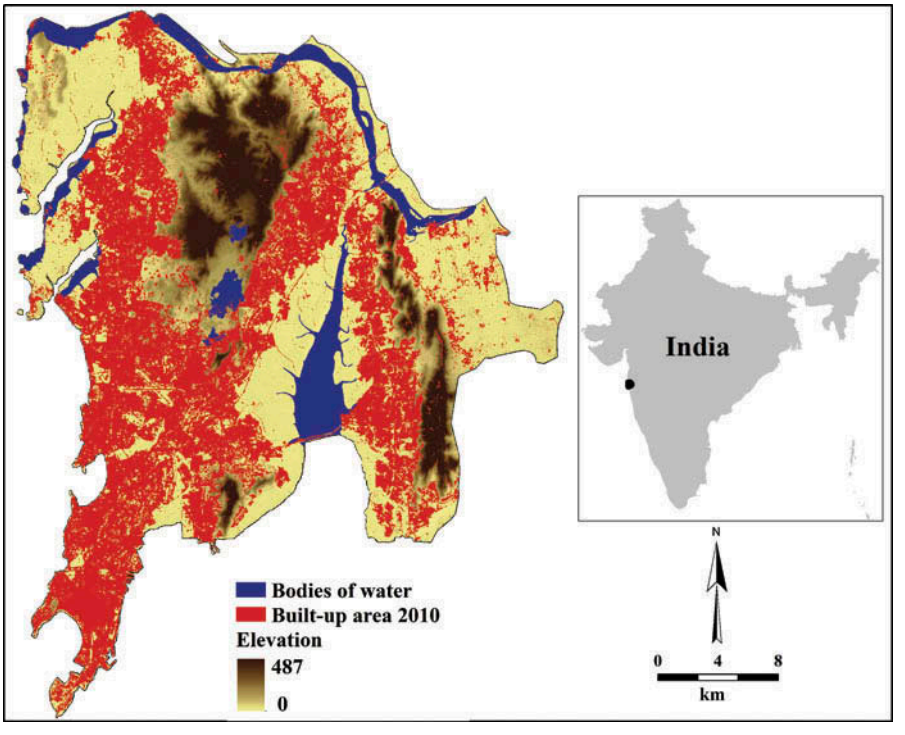

Figure 1. Study area, Mumbai, India.

A recent study conducted by Pethe *et al.* (2014) shows disorganized and haphazard patterns of current land use in Mumbai and its deviation from the planned land use, while Shafizadeh-Moghadam and Helbich (2013) predict further urban expansion throughout Mumbai in the next two decades. Therefore, to assist authorities and urban planners, the development of robust methods for predicting LUC in Mumbai is crucial.

3. Data sources and preprocessing

The main data sources used in this study are Landsat imagery dated 2001 (ETM) and 2010 (ETM^{+}). Urban footprints and most spatial attributes were derived from these. The 30 m spatial resolution imageries were acquired from the US Geological Survey (USGS). Moreover, transportation network data were accessed from the OpenStreetMap (OSM) collaborative database. The total study area covers 950,374 cells of 30 × 30 m.

Built-up areas and other important classes, including wetlands, bodies of water, forest, crop land, and open land, were extracted with a supervised scheme called maximum likelihood, as provided by the ENVI 4.7 software (Boulder, Colorado, USA). A post-classification procedure was conducted, whereby results were compared with reference maps, Google Earth and maps extracted by Taubenböck *et al.* (2012), and improved accordingly. To map built-up areas, they applied object-oriented and pixel-based classification image analysis techniques on Landsat imagery.

To predict urban growth, a set of 10 spatial variables were defined. While there is no consensus among scholars on a standard set of factors driving urban change, the majority of the spatial variables used in this paper are also used in most urban growth studies (Hu and Lo 2007, Wang and Li 2011, Feng and Liu 2013). However, the driving factors of LUC might vary across regions, given that people, LUC procedures, as well as government policies and plans, among others, affect the importance of these factors (Thapa and Murayama 2011). Next, distance maps were created using an Euclidean distance function provided in ArcGIS software (Redlands, California, USA), for the water bodies, forest, built-up areas, main roads, railway, and central business district (CBD) variables. Using a 7 × 7 neighborhood function (Hu and Lo 2007), two density maps (i.e., density of built-up areas and density of open land and crop land) were also created. These density maps can explicitly account for the influence of neighboring pixels. The DEM layer was acquired from the US Shuttle Radar Topography Mission (SRTM) website, and a slope map was derived from the same layer. All maps were then made consistent with satellite images by resampling them to a 30 m spatial resolution and by referencing them to the UTM coordinate system zone 43 N. It is worth noting that resampling the DEM layer from a 90 m to 30 m resolution does not have a significant influence on the analysis, because the study area is rather flat.

Future urban growth is limited to specific areas, mainly open lands and crop lands. Locations where urban growth is not allowed to occur are known as exclusive areas (Pijanowski *et al.* 2002). In our study, built-up areas in 2001, bodies of water and wetlands were considered as exclusive areas. Beside these areas, the remaining regions were included in the modeling process and presented to the network as places where future change could occur. Table 1 describes the explanatory variables of the data set.

Table 1. Spatial explanatory variables of urban change between 2001 and 2010.

Variables	Range	Mean	Data type
Distance to CBD (m)	0–33,000	18,069	Continuous
Distance to railway (m)	0–11,500	2385	Continuous
Distance to main roads (m)	0–6200	515	Continuous
Distance to built-up areas (m)	0–4300	536	Continuous
Distance to bodies of water (m)	0–23,000	2636	Continuous
Distance to wetlands (m)	0–22,000	4188	Continuous
Distance to forest (m)	0–16,000	2793	Continuous
DEM (m)	0–487	39	Continuous
Slope (°)	0–60	5.86	Continuous
Density of built-up areas (cell)*	0–49	29	Ranked
Density of available open land/crop land (cell)*	0–49	24.8	Ranked

Note: *Each cell covers 900 sqm.

4. Method

After the preparation of explanatory variables, we employed two methods, RBFN and MLP, for prediction of urban change for 2010 in Mumbai using the process depicted in Figure 2. The next process involved evaluating the models' output using a set of statistical indices. Through this process an ideal method was selected for simulating urban change for 2020 and 2030. The next sections will describe the methods that were used for the prediction process.

4.1. Radial basis function network (RBFN)

In this paper, we used a RBFN model for enumerating urban change and their underlying dynamics. The RBFN is basically a feed-forward neural network consisting of three layers of neurons (Moody and Darken 1989) with N inputs, L nodes in the hidden layer, and M nodes in the output layer; however, the parameter initialization of RBFN (such as centers and widths of functions) is obtained according to the distribution of data. This is in contrast to the feed-forward back propagation algorithm where parameters are typically initiated randomly (Li and Hori 2006). In RBFN, each neuron of the input layer is connected with each neuron of the hidden layer, and each neuron of the hidden layer is connected with each neuron of the output layer. Neurons within the same layer are not connected. RBFN can be implemented using either supervised or unsupervised learning procedures.

The input layer is not directly involved in the calculation of the RBFN's output. The purpose of it is to distribute the N-dimensional input vectors that are our explanatory variables to the neurons of the hidden layer, which consecutively transform the input vectors. For this purpose, each neuron of the hidden layer has a prototype vector assigned to it. This vector represents the center of a Gaussian function, the radial basis function, which is associated with the neuron.

$$\varnothing = \exp - \left(\frac{|x - \mu_c|^2}{\sigma^2} \right), \tag{1}$$

where μ_c is a basis center and σ is the width or scale parameter.

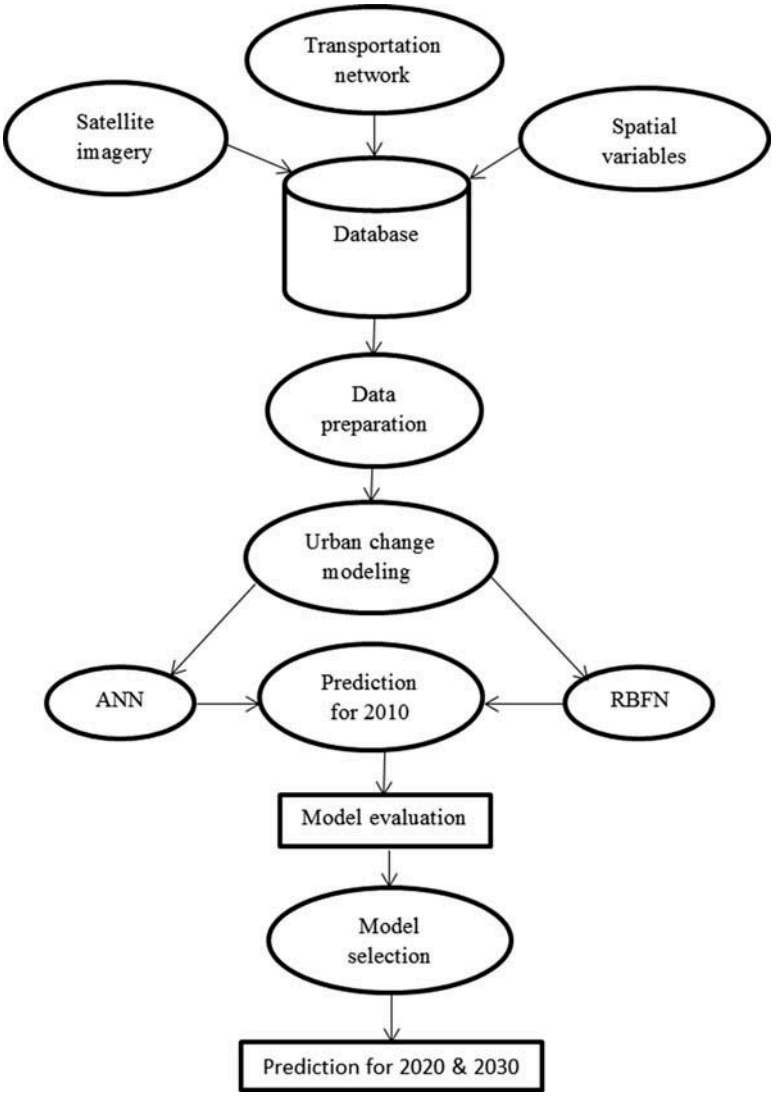

Figure 2. Flowchart of the urban change modeling and evaluation.

The transformed input vectors are then propagated forward through the weighted connections to the neurons of the output layer, which calculate the sum of incoming signals as follows:

$$y = w_o + \sum_{i=1}^{M} w_i \varnothing i_i, \tag{2}$$

where y is the output from the network for the input x, w_o is the bias, w_i are the weights of the linear combiner, and M is the total number of basis functions.

Since the number of neurons of the output layer usually is the output space dimension and the number of neurons of the input layer is the dimension of the input vectors, the number of neurons of the output layer must be appropriately chosen beforehand. The output of the network contains continuous values ranging from 0 to 1 which is a linear combination of radial basis functions of the inputs and neuron parameters.

In contrast, the centers and width of the radial basis functions and the connection weights are determined through training, which, as described in this paper, is typically performed in three consecutive stages that are dependent on each other.

(1) In the first stage, the centers of the radial basis functions are determined. This is typically done using vector quantization algorithms such as the neural gas or k-means (Moody and Darken 1989). In this paper, we used the neural gas algorithm. Neural gas is a well-known unsupervised method for learning the topologies associated with data. It is broadly used as an efficient algorithm in learning tasks, for example, clustering, dimensionality reduction, and vector quantization (Shi *et al.* 2014). Vector quantization determines a set of prototype vectors that approximate the probability density function of the input space. Each prototype vector represents a Voronoi region, which consists of the nearest input vectors.
(2) In the second stage, the kernel widths of the functions are calculated. A convenient approach is to set the width of each kernel function to the distance between its center and the center of the nearest radial basis function (Saha and Keeler 1990). In addition, Benoudjit and Verleysen (2003) suggested scaling the determined widths by a constant factor in order to preserve a natural overlap between the kernel functions.
(3) In the last stage, the weights of connections between the hidden and output layer are obtained. This can be done by iteratively presenting the input vectors to the network and moving the weights in a direction opposite from the gradient of the error function, effectively minimizing the error.

4.2. Multi-layer perceptron (MLP)

Another approach for estimating spatial urban change transition probabilities was a MLP model trained with a back propagation algorithm (BP). MLP is a widely used design in ANN that has been applied successfully in many disciplines, including LUC modeling. It typically includes a set of interconnected layers of neurons, that is, an input layer, one or more hidden layer of computational nodes, and an output layer. In a layer-by-layer fashion, the input layer propagates through the network in a forward manner. In many cases including this paper, MLP is trained by a BP algorithm (Chauvin and Rumelhart 1995). The BP algorithm operates on the basis of error-correction rule (Haykin 1994) that includes a forward pass and backward pass step to adjust the calculated weights. In the forward pass, the input vector is presented to the network and its effect propagates through the network on a layer-by-layer fashion, with a set of output layers produced as a result. The weighting process is initiated by randomly assigning and calculating the mean squared error (MSE), after which the iteration continues until a certain amount of accuracy is met (Rumelhart *et al.* 1986). The MSE is measured by subtracting the output value of the network from the actual output presented in the training phase (Tayyebi and Pijanowski 2014). The process of optimizing calculated weights in order to reduce error begins if the acceptable level of performance is not achieved. In that case, the error is

circulated back to the neurons in the hidden layer, allowing them to update the weightings and alleviate the error. Once the validation run confirms that performance of the training phase has been completed successfully (e.g., using a data set exclusively set aside for validation), the model can be considered for running on other data sets for further implementations.

While in the forward pass, the estimated weights are fixed, in the backward pass, they are adjusted based on the error-correction rule, and this process is pursued until a certain error-correction rate is reached. Based on Haykin (1994), each neuron in the network includes a nonlinear activation function. A commonly used form of nonlinearity that was also employed in this paper is sigmoidal nonlinearity, which is defined as follows:

$$y_j = \frac{1}{1 + \exp(-v_j)}, \tag{3}$$

where v_j is the weighted sum of all inputs and bias term and y_j is the output of the neuron.

To prevent overfitting, training data was presented randomly to the network and after 5000 cycles the network started to converge, a situation which shows the learning process stabilizes the error level to a minimum value (Pijanowski *et al.* 2002).

For a detailed discussion on the use of MLP in LUC, please refer to Li and Yeh (2002) and Pijanowski *et al.* (2002).

5. Results and discussion

5.1. Training process

Following the removal of exclusionary zones including built-up areas, bodies of water and wetlands in 2001, 489,743 pixels remained in the data layer. Comparison of the Landsat imageries for 2001 and 2010 identifies that a subset of 92,483 cells transformed into the built-up class and 397,160 cells remained unchanged. To begin with, unchanged and changed cells were considered as dependent variables versus the 10 spatial explanatory variables described in Section 2 as independent variables. Dependent variable was coded as either 0 or 1. Pixels that were transformed to the built-up class between 2001 and 2010 were coded as 1, and those remained unchanged were considered as 0. For both RBFN and MLP, the same set of spatial drivers was used. For the implementation of our model, we used the R programming language; all variables were converted to the comma-separated values file format (compatible with R).

Prior to introducing data to the network, to avoid inconsistency due to the potential effects of different ranges and units of drivers on the model output, the variables were normalized between 0 and 1 (Feng and Liu 2013) using the following equation:

$$x_i^{'} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{4}$$

For each spatial variable, $x_i^{'}$ is the transformed value of the ith row, x_i is the original value, $x_{\max}$ is the maximum value, and $x_{\min}$ is the minimum value. Subsequently, following the normalization, the spatial variables were then split into two parts, one portion for training (70%) and the remaining (30%) for validation. The validation process helps to ensure that the model has not memorized the training sample, which would reduce the generalization

ability of the model for unseen data. For both models, 10 nodes equal to the number of explanatory variables (input layer) were selected for the hidden layer. The Gaussian function was chosen as the activation function for the RBFN model and the sigmoid activation function for the MLP model. The result of the model execution was a transition probability map, ranging from 0 to 1, with 0 indicating a minimal likeliness that urban growth will occur, and 1 indicating a maximal likeliness.

The next step consisted in assigning the location of changes for 92,483 cells that had changed to built-up class between 2001 and 2010. Following the approach of Pijanowski *et al.* (2002) and Wang and Li (2011), cells with the greatest values are supposed to transition first. Accordingly, we have assigned them to the cells that have the highest likelihood value in the transition probability map obtained from the training process. Thus, 92,483 cells which had the greatest change likelihood values were selected from the output transition probability map obtained from the MLP and RBFN models and classified as new built-up areas. Figure 3 shows the observed changes between 2001 and 2010 along with the predicted changes obtained by RBFN and MLP models.

After model validation, this process continues for simulating urban change for the upcoming years using the relatively well-performing model. The rational is that the amount of future urban growth will be constant similar to the previous explored time intervals. For example, the converted 92,483 cells between 2001 and 2010 indicate that the annual average growth was approximately 10,276 cells and this amount is multiplied by 10 and 20 for simulating urban growth in the selected years of 2020 and 2030, respectively.

5.2. Model validation and comparison

From the total amount of 92,483 cells that were converted to built-up class between 2001 and 2010, RBFN and MLP could correctly predict the spatial location of 59,143 (63.95%) and 56,064 (60.62%) of the changed cells, respectively. From the total amount of 397,160 cells that remained unchanged, the percentage correctly predicted as unchanged were correspondingly 369,167 (92.95%) and 360,790 (90.84%) for RBFN and MLP models, respectively. These proportions indicate that both models succeeded to correctly predict the change of over half of cells. While the same region and data set were used for both models, quantifying the likeliness that change could occur for each cell is, to some extent, done differently in each model, because of the different architectures and methods they are based on.

In order to evaluate our LUC models, we used relative operating characteristic (ROC) curve which measures similarity between the observed and simulated change (Pontius and Schneider 2001). RBFN and MLP models resulted in 85.2% and 89.1%, respectively (Figure 4). The results show that both models operated better than a random model (a model with 50% ROC value). However, in a comparative view, the ROC curve informs us that the RBFN model generates better probability maps. The ROC curve does not however indicate the spatial configuration of the model's accuracy and errors and therefore a visual comparison and supplementary indices of association that account for spatial arrangement are required (Costanza 1989, Turner *et al.* 1989).

To assess the spatial accuracy of both models, the predicted maps and the reference map (which is the urban extent of 2010, as extracted from the Landsat data) were overlaid so that four diagnoses classes were obtained: (1) true positive: cells that were identified by

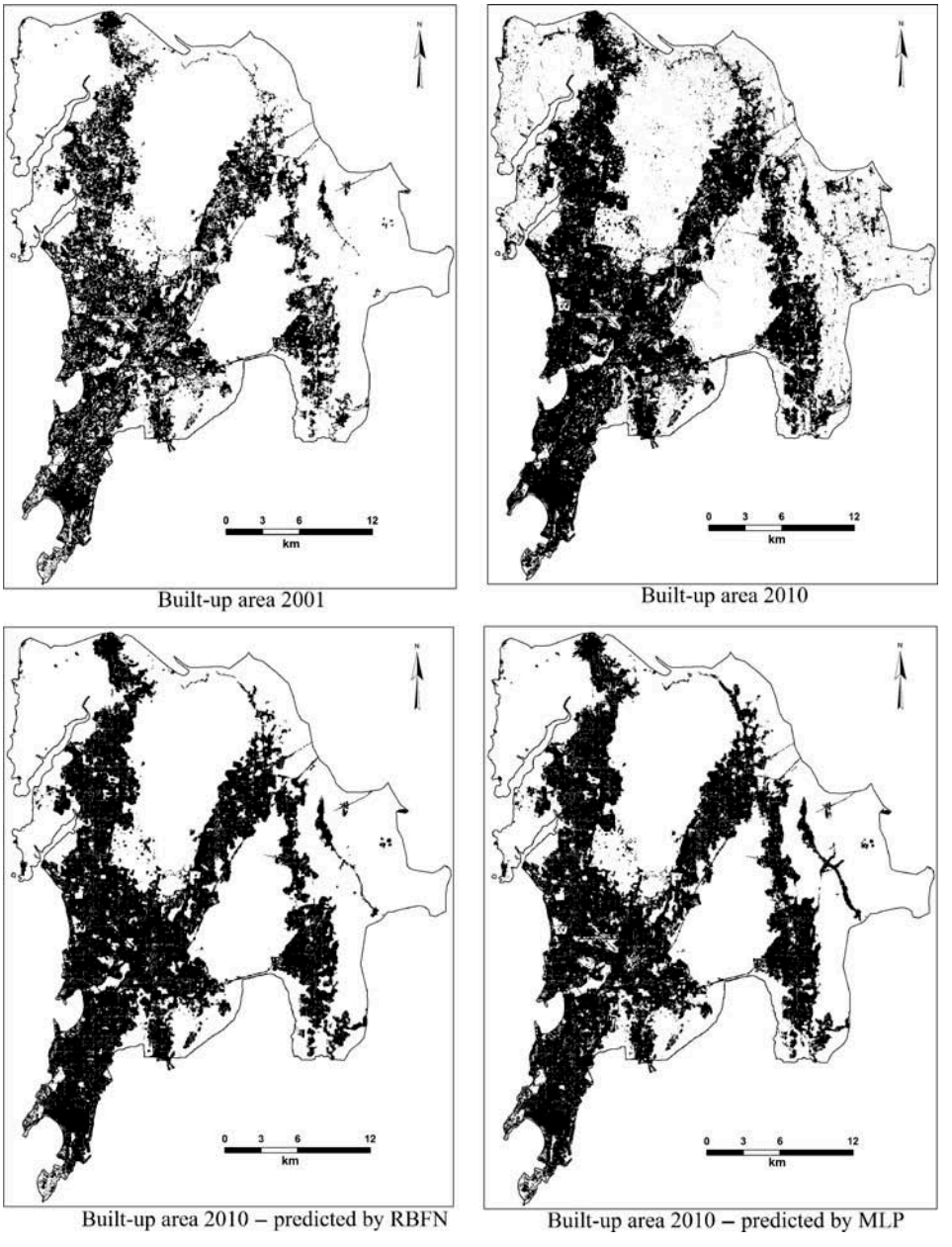

Figure 3. Observed urban changes versus changes predicted by RBFN and MLP.

both model and on the reference map as cells where change occurred; (2) true negative: cells that were identified by both models and on the reference map as cells where no change occurred; (3) false positive: cells that were identified by both model as cells where change occurred, but that are identified on the reference map as cells where no change occurred; and (4) false negative: cells for which no change was predicted by both models, but that were identified on as cells where change occurred. Figure 5 shows the spatial

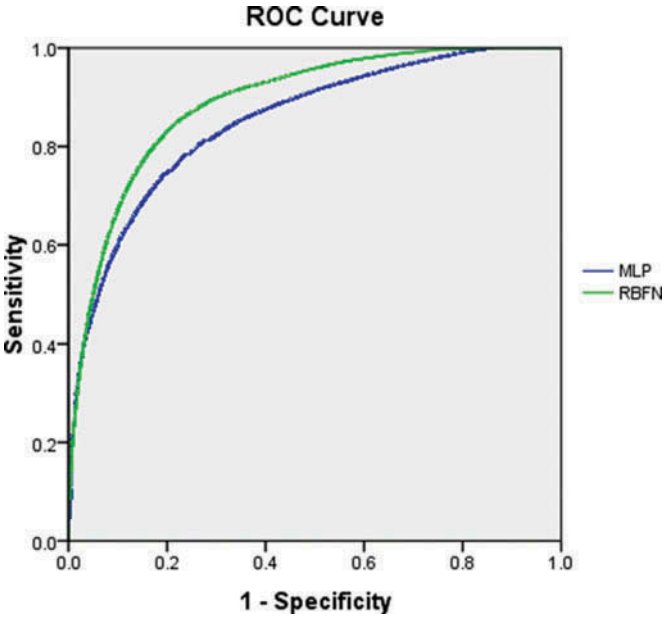

Figure 4. ROC curve for assessing the performance of MLP and RBFN models.

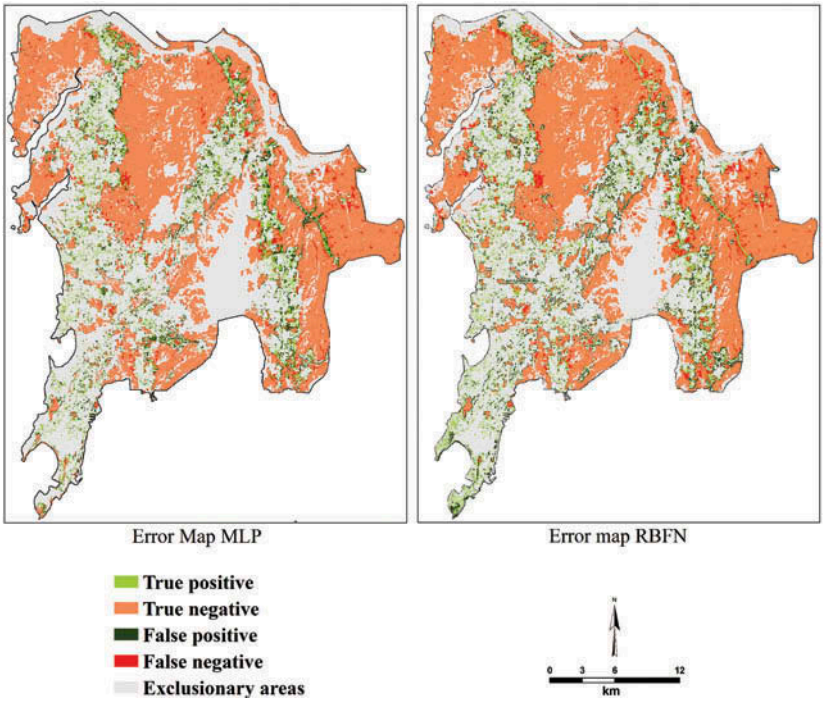

Figure 5. Predicted maps for 2010 obtained by overlaying the 2010 reference map and the maps predicted by RBFN and MLP models.

distribution and visualization of the agreements and discrepancies between the reference map and the predictions of the RBFN and MLP models.

We used four statistical measures, namely FoM, PA, OA, and ASDD, to assess different aspects of both models versus the reference map. The selected measures are among the recommended and widely used tools for assessing urban growth models (Pontius *et al.* 2008, Thapa and Murayama 2011). The FoM is the ratio of the intersection of the observed change and predicted change to the union of the observed change and predicted change; it ranges from 0% (no overlap) to 100% (full match between observed and predicted changes) (Pontius *et al.* 2008).

$$\mathrm{FoM} = \frac{B}{(A + B + C + D)} \tag{5}$$

PA and OA are defined as follows:

$$\mathrm{PA} = \frac{B}{(A + B + C)} \tag{6}$$

$$\mathrm{OA} = \frac{(B + E)}{(A + B + C + D + E)} \tag{7}$$

In Equations 5–7, A is the fraction 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 fraction of error cells due to observed change modeling – predicted as a wrong gaining category, and D is the fraction of error cells due to observed persistence – predicted as change. E denotes the area of correctness due to observed persistence – predicted as persistence.

ASDD is calculated as follows:

$$\mathrm{ASDD}(A, P) = \frac{1}{K} \sum_{k=1}^{K} D(A_k, A_k.\mathrm{nearest}(P)), \tag{8}$$

where A_k is the observed built-up cell, P is the map layer of predicted built-up areas, $A_k.\mathrm{nearest}(P)$ denotes the nearest predicted location to A_k, and K is the number of actual built-up cells. Furthermore, the ASDD approach allows to observe how far the cells have deviated from the baseline situation where change was predicted.

FoM, PA, OA, and ASDD were 46.92, 63.95, 87.47, and 21.17 for the RBFN, and 43.5, 60.62, 85.13, and 20.09 for the MLP, respectively. As it can be seen from these results and visualized in Figure 6, both models performed well and achieved promising results, but the RBFN model appeared better in terms of spatially explicit prediction of location of the changes. However, boxplot exploration of the two models shows that both models have predicted that changes would occur in locations where it was unlikely to occur. This issue is evident in both the observed changes, as well as in the simulated maps. In reality, there are some new constructions that, depending on their usage, should be built at a specific distance from the already built-up areas, which gives them lower likelihood value in comparison with the majority of other built-up areas. In addition, other isolated constructions that do not fall in any of the main land use classes, such as

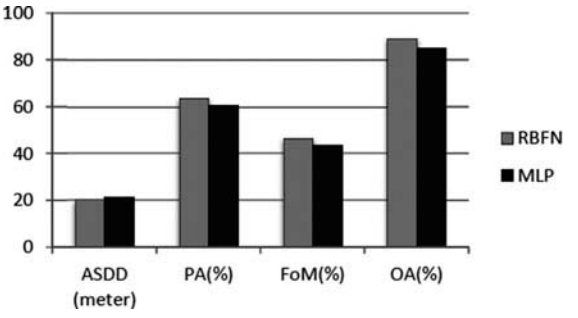

Figure 6. Statistical indices showing the intersection of the observed change and the predicted change.

residential, commercial, or industrial, also contributed to this issue. Therefore, to avoid such inconsistency, it would be better to filter out such island pixels.

5.3. Urban change predictions

Given the superiority of RBFN over MLP in predicting urban change, the spatio-temporal future pattern of growth in Mumbai for 2020 and 2030 was simulated using the RBFN model. In doing so, we assumed a constant rate of change for the next years, as observed in the previous time interval. The previous time span was 9 years (between 2001 and 2010. The total number of changes was divided by the number of years, and the average rate of growth obtained is 10,276 cells per year, or approximately 925 hectare per year. This average growth rate has been used for the 2020 and 2030 predictions. Then, the simulated patterns for 2020 and 2030 were added to the observed urban built-up area in 2010 (Figure 7). However, it should be noted that such assumption always carry some uncertainty. This is due to the fact that the function and relation of spatial factors of urban change are not necessarily fixed over time, as in previous years. In addition, some drivers will probably change for example, it is not far-fetched to see that future pattern of transportation network will change as it is a leading driver in most urban change models.

6. Conclusion

Spatial accuracy of predicted urban changes is of great importance to urban planners and decision makers. Thus, developing more accurate spatially explicit LUC models is necessary. Compared to other well-known techniques, for instance CA and logistic regression, little research on modeling urban growth with ANNs has been conducted.

In this paper, we analyzed the performance of two neural-network-based approaches, namely RBFN and MLP, for modeling urban change. We applied both models on the same case study (Mumbai, India), and used the same set of drivers. We used two Landsat images dated 2001 and 2010 for mapping the urban extent and for extracting spatial factors. The predicted urban change for 2010 was compared with a 2010 reference map and the results were evaluated using a set of statistical indices, including FoM, PA, OA, and SDMM.

Despite the fact that both models performed well in terms of spatial accuracy of predicted changes, and despite their promising results in comparison with other LUC

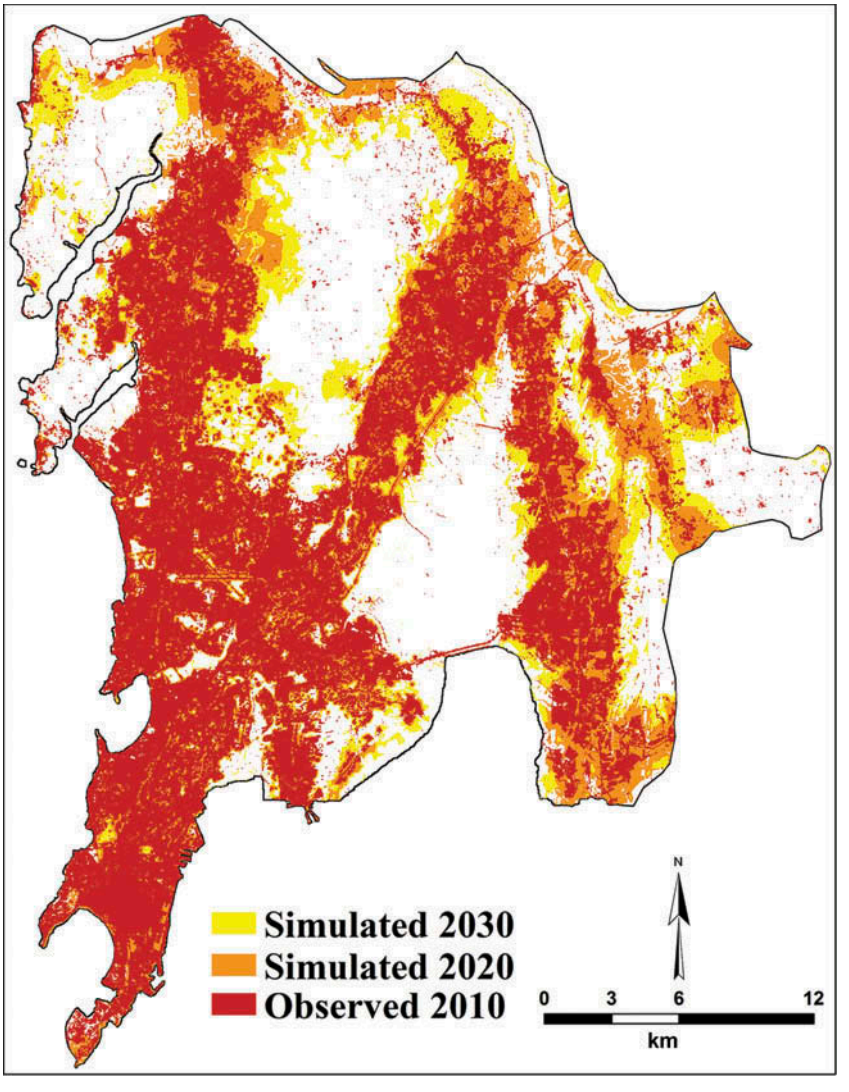

Figure 7. Actual built-up areas for the year 2010 and simulated built-up areas for 2020 and 2030 by the RBFN model.

models, given the relative superior performance of RBFN, this method was employed for simulating urban change in 2020 and 2030.

ANN-based models can learn essential characteristics of urban change dynamics including complexity and nonlinear relationships through a training process, and can generalize the discovered pattern to the unseen data. However, depending on specific characteristics of different geographical regions and the role of master plans and policy makers, spatial accuracy of the employed models may vary. When compared to other models such as logistic regression, they are not capable of quantifying and separating the magnitude effect of urban change driving forces in an explicit manner.

In the same vein, as discussed by Pijanowski *et al.* (2002), there are some assumptions and limitations in this paper. First, the projection assumes that spatial patterns and the extent of urban drivers will remain constant. For variables such as the transportation network, this assumption might not be valid and thus including updated and planned infrastructure information for the future studies is highly recommended. Another important issue is related to the assumption of a constant relationship between urban dynamics and their underlying drivers that may vary over time. Future studies could also consider relaxation of this assumption, potentially by developing hybrid models.

In conclusion, the results of this study show the ability of RBFN for predicting urban change and its potential as a robust tool for spatial-oriented planning and protecting the environment towards sustainable urban development. Further studies could focus on performance analysis of RBFN with other ML approaches, such as SVM, and on its comparison with other empirical LUC models. Also, adding further socioeconomic variables that could lead to more precise predictions but that were not available for this study is highly recommended.

References

Almeida, C.M., *et al.*, 2008. Using neural networks and cellular automata for modelling intra-urban land-use dynamics. *International Journal of Geographical Information Science*, 22 (9), 943–963. doi:10.1080/13658810701731168

Basawaraja, R., *et al.*, 2011. Analysis of the impact of urban sprawl in altering the land use, land-cover pattern of Raichur City, India, using geospatial technologies. *Journal of Geography and Regional Planning*, 4, 455–462.

Batty, M., Xie, Y., and Sun, Z., 1999. Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems*, 23, 205–233. doi:10.1016/S0198-9715(99)00015-0

Benoudjit, N. and Verleysen, M., 2003. On the kernel widths in radial-basis function networks. *Neural Processing Letters*, 18 (2), 139–154. doi:10.1023/A:1026289910256

Chauvin, Y. and Rumelhart, D.E., 1995. *Backpropagation: Theory, Architectures, and Applications*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Costanza, R., 1989. Model goodness of fit: a multiple resolution procedure. *Ecological Modelling*, 47, 199–215. doi:10.1016/0304-3800(89)90001-X

Feng, Y. and Liu, Y., 2013. A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing. *International Journal of Geographical Information Science*, 27 (3), 449–466. doi:10.1080/13658816.2012.695377

Foody, G.M., 1995. Land-cover classification by an artificial neural network with ancillary information. *International Journal of Geographical Information Systems*, 9 (5), 527–542. doi:10.1080/02693799508902054

Haykin, S., 1994. *Neural networks: a comprehensive foundation*. Upper Saddle River, NJ: Prentice Hall PTR.

Hosseinali, F., Alesheikh, A.A., and Nourian, F., 2013. Agent-based modeling of urban land-use development, case study: simulating future scenarios of Qazvin city. *Cities*, 31, 105–113. doi:10.1016/j.cities.2012.09.002

Hu, Z. and Lo, C.P., 2007. Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems*, 31 (6), 667–688. doi:10.1016/j.compenvurbsys.2006.11.001

Jiang, B. and Yao, X., eds., 2010. *Geospatial analysis and modelling of urban structure and dynamics*. Dordrecht: Springer. ISBN 9048185718.

Kobler, A. and Adamic, M., 2000. Identifying brown bear habitat by a combined GIS and machine learning method. *Ecological Modelling*, 135 (2–3), 291–300. doi:10.1016/S0304-3800(00)00384-7

Kolb, M., Mas, J.-F., and Galicia, L., 2013. Evaluating drivers of land-use change and transition potential models in a complex landscape in Southern Mexico. *International Journal of Geographical Information Science*, 27 (9), 1804–1827. doi:10.1080/13658816.2013.770517

Lambin, E.F., *et al.*, 2001. The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, 11 (4), 261–269. doi:10.1016/S0959-3780(01)00007-3

Lambin, E.F., Geist, H.J., and Lepers, E., 2003. Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, 28, 205–241. doi:10.1146/annurev.energy.28.050302.105459

Li, W. and Hori, Y., 2006. An algorithm for extracting fuzzy rules based on RBF neural network. *IEEE Transactions on Industrial Electronics*, 53, 1269–1276. doi:10.1109/TIE.2006.878305

Li, X. and Yeh, A.G.O., 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16, 323–343. doi:10.1080/13658810210137004

Lin, G.-F. and Chen, L.-H., 2004. A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. *Journal of Hydrology*, 288 (3–4), 288–298. doi:10.1016/j.jhydrol.2003.10.008

Lin, Y.-P., *et al.*, 2011. Predictive ability of logistic regression, auto-logistic regression and neural network models in empirical land-use change modeling–a case study. *International Journal of Geographical Information Science*, 25 (1), 65–87. doi:10.1080/13658811003752332

Moody, J. and Darken, C.J., 1989. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1 (2), 281–294. doi:10.1162/neco.1989.1.2.281

Pethe, A., *et al.*, 2014. Re-thinking urban planning in India: learning from the wedge between the de jure and de facto development in Mumbai. *Cities*, 39, 120–132. doi:10.1016/j.cities.2014.02.006

Pijanowski, B.C., *et al.*, 2002. Using neural networks and GIS to forecast land use changes: a land transformation model. *Computers, Environment and Urban Systems*, 26 (6), 553–575. doi:10.1016/S0198-9715(01)00015-1

Pijanowski, B.C., *et al.*, 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*, 51, 250–268. doi:10.1016/j.envsoft.2013.09.015

Pontius, G.R. and Malanson, J., 2005. Comparison of the structure and accuracy of two land change models. *International Journal of Geographical Information Science*, 19 (2), 243–265. doi:10.1080/13658810410001713434

Pontius Jr, R.G., *et al.*, 2008. Comparing the input, output, and validation maps for several models of land change. *The Annals of Regional Science*, 42 (1), 11–37. doi:10.1007/s00168-007-0138-2

Pontius Jr, R.G. and Schneider, L.C., 2001. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, 85 (1–3), 239–248. doi:10.1016/S0167-8809(01)00187-6

Potts, M.S. and Broomhead, D.S., 1991. Time series prediction with a radial basis function neural network. *SPIE. Adapt. Signal Process*. 1565, 255–266. doi:10.1117/12.49782

Rienow, A. and Goetzke, R., 2014. Supporting SLEUTH–Enhancing a cellular automaton with support vector machines for urban growth modeling. *Computers, Environment and Urban Systems*, 49, 66–81.

Rogan, J., *et al.*, 2008. Mapping land-cover modifications over large areas: a comparison of machine learning algorithms. *Remote Sensing of Environment*, 112 (5), 2272–2283. doi:10.1016/j.rse.2007.10.004

Rumelhart, D.E., Hinton, G.E., and Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature*, 323 (6088), 533–536. doi:10.1038/323533a0

Saha, A. and Keeler, J.D., 1990. Algorithms for better representation and faster learning in radial basis function networks. *In*: *Advances in neural information processing systems*. San Francisco, CA: Morgan Kaufmann, 482–489.

Samanta, B., 2010. Radial basis function network for ore grade estimation. *Natural Resources Research*, 19 (2), 91–102. doi:10.1007/s11053-010-9115-z

Sha, W. and Edwards, K.L., 2007. The use of artificial neural networks in materials science based research. *Materials and Design*, 28 (6), 1747–1752. doi:10.1016/j.matdes.2007.02.009

Shafizadeh-Moghadam, H. and 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. doi:10.1016/j.apgeog.2013.01.009

Shafizadeh-Moghadam, H. and 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 (unpublished).

Shi, J., Chen, C., and Zhong, S., 2014. Privacy preserving growing neural gas over arbitrarily partitioned data. *Neurocomputing*, 144, 427–435. doi:10.1016/j.neucom.2014.04.033

Taubenböck, H., *et al.*, 2012. Monitoring urbanization in mega cities from space. *Remote Sensing of Environment*, 117, 162–176. doi:10.1016/j.rse.2011.09.015

Tayyebi, A. and 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. doi:10.1016/j.jag.2013.11.008

Tayyebi, A., Pijanowski, B.C., and 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 (1–2), 35–44. doi:10.1016/j.landurbplan.2010.10.007

Thapa, R.B. and Murayama, Y., 2011. Urban growth modeling of Kathmandu metropolitan region, Nepal. *Computers, Environment and Urban Systems*, 35 (1), 25–34. doi:10.1016/j.compenvurbsys.2010.07.005

Triantakonstantis, D. and Mountrakis, G., 2012. Urban growth prediction: a review of computational models and human perceptions. *Journal of Geographic Information System*, 4 (6), 555–587. doi:10.4236/jgis.2012.46060

Turner, M.G., Costanza, R., and Sklar, F.H., 1989. Methods to evaluate the performance of spatial simulation models. *Ecological Modelling*, 48, 1–18. doi:10.1016/0304-3800(89)90057-4

United Nations, 2012. *World urbanization prospects. The 2011 revision*. New York: United Nations Publication.

Verburg, P.H., *et al.*, 2002. Modeling the spatial dynamics of regional land-use: the CLUE-S model. *Environmental Management*, 30 (3), 391–405. doi:10.1007/s00267-002-2630-x

Wang, J. and Mountrakis, G., 2011. Developing a multi-network urbanization model: a case study of urban growth in Denver, Colorado. *International Journal of Geographical Information Science*, 25 (2), 229–253. doi:10.1080/13658810903473213

Wang, Q., Shi, W., and Atkinson, P.M., 2014. Sub-pixel mapping of remote sensing images based on radial basis function interpolation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92, 1–15. doi:10.1016/j.isprsjprs.2014.02.012

Wang, Y. and Li, S., 2011. Simulating multiple class urban land-use/cover changes by RBFN-based CA model. *Computers and Geosciences*, 37 (2), 111–121. doi:10.1016/j.cageo.2010.07.006