Journal of Geomatics Vol. 14, No. 1, April 2020

Integration of Cellular Automata-Markov Chain and Artificial Neural Network model for urban growth simulation

Kriti Rastogi* and Shashikant A. Sharma
Space Applications Centre, ISRO, Ahmedabad-380015
*Email: kritirastogi@sac.isro.gov.in

(Received: Jan 08, 2020; in final form: June 30, 2020)

Abstract: The urban sprawl and growth modelling helps in identifying areas of potential urban expansion. This will help in exploring alternatives in urban design and human-environment interactions. Also, it minimizes the negative impact of urban sprawl for sustainable and environment-friendly futures. In this study, a spatio-temporal urban growth modelling for year 1992-2032 of Ahmedabad city is performed by simulating historical urban built-up data and predicting future urban growth. The Indian Remote Sensing satellite data for year 1992, 2005 and 2018 is used for generating urban built-up land cover and change map. These are analysed along with the auxiliary data such as proximity from major roads, slope, population density which acts as the driving force for urban expansion. The integration of Cellular Automata-Markov Chain (CA-MC) model is used along with Artificial Neural Network for understanding the relationship between the driving forces and urban built-up. The urban built-up has increased from 146.67 km^2 to 193.82 km^2 during 1992 to 2005 which has further increased to 229.52 km^2 in 2018 with increase of 36% in urban built-up in past 27 years. The 2018 predicted urban expansion is validated using overall accuracy of 97% and with kappa value 0.94. The prediction of 2032 urban expansion as per the optimised ANN CA-MC model has an area of 275.76 km^2.The results show that the increase in built-up area are closely associated with the existing built-up areas.

Keywords: Urban Sprawl, Cellular Automata**,** Markov Chain, Artificial Neural Network

1. Introduction

Rapid increase in population lead to unprecedented urban growth stemming variety of complex problems such as traffic congestion, air pollution, deforestation, farmland decrease and chaotic urban settlements with poor infrastructure. Urban growth in sub-urban areas of the metropolitan cities are scattered development with low density, poor accessibility involving massive change in urban land cover affecting the ecosystem, biodiversity and natural resources. In order to ensure sustainable development, decision makers and urban planners needs precise information regarding the urban growth and its future expansion. Urban land cover and change analysis and urban growth prediction is an important input, for assessing the amount and impact of development and its consequence on the environment (Jiang and Yao, 2010). Also, it is helpful for understanding and developing theories of urban morphologies and its interaction with other land use classes for developing environmental models such as urban climate models.

India has taken major initiative to start smart cities mission for establishing sustainable cities through urban planning and proper management plans. Remote sensing and Geographical Information System (GIS) together offer a powerful tool for spatial and temporal analysis of urban growth and it can provide regular data of urban expansion. Urbanisation being a complex process needs comprehensive mathematical model, along with additional socio-economic and demographic variables for simulating its process. Several GIS-based mathematical models have been used for forecasting with spatial scope, such as Cellular Automata (CA) (Batty et al., 1999; Li and Yeh 2000; Sudhira et al., 2004; Aburas et al., 2016), Multi-Agent Model (Arsanjani et al., 2013; Zhang et al., 2015), Land Transformation Model (Pijanowski et al., 2002) and SLEUTH (Jat et al., 2017). Among these models, Cellular Automata (CA) is one of the most popular model used for simulation and prediction of urban growth.

CA simulations and predictions are governed on the assumption that previous urban growth will affect the future pattern through local and regional interactions among different land use classes (Sante et al, 2010). Because of its 'bottom-up' structure, CA can simulate the emergent macro-scale phenomenon (urban expansion) by micro-level interactions (cell state change) (Xu et.al., 2019). In allocating changes under the predefined conditional rules, CA always starts with the cells of the highest probability of change. Therefore, it is capable of predicting the most probable sites for development, estimating the probability of amount of change, as well as allocating the estimated quantity of change within a study area. However, the conventional CA approach considers only the neighbourhood effect in spatial allocation without quantitatively considering the role of urban expansion drivers (Aburas et al., 2016; Li et al., 2017; Mustafa et al., 2017).

The limitation of conventional CA model is resolved by its open structure, which can be integrated with other models such the logistic regression, fuzzy logic and Markov Chain (Wang et al., 2013; Bihamta et al., 2015) to stimulate urban growth pattern (Clarke, 1997). Among common integration of models Cellular Automata-Markov Chain (CA-MC) has been one of the popular methods used to model urban expansion. It requires comparatively smaller number of driving factors and physical constraints for adequately simulate and predict urban expansion (Arsanjani et al., 2013; Liu, 2012). For understanding the relation of urban growth with its drivers, logistic regression could be used for creating optimum set of variables for the CA-MC model. However, it may not work well if the relationship between urban expansion and its drivers is non-linear (Mustafa et al., 2017) and also working with large datasets. In order to incorporate non-linearity between the urban expansion and its drivers, Artificial Neural Network is used. Integrated ANN-CA-MC has the ability to model complex non-linear relationships between dependent (such as urban growth) and independent variables (such as distance to roads, demographics etc) with fewer statistical assumptions and most importantly without knowing prior relation between these variables. For this reason, ANNs have been incorporated into other models such as CA-MC to simulate and predict urban expansion despite the difficulties in properly parameterizing and optimally configuring an ANN model.

In this study, we integrate an optimised ANN with CA-MC to stimulate and predict the urban expansion of Ahmedabad city in 2018 and 2032. This study attempts to identify the urban sprawl patterns using remote sensing and GIS technique and predicts the urban growth for future. The model is validated with the reference data, quantified using overall accuracy and kappa coefficient.

2. Study Area

Ahmedabad city is in the central part of Gujarat, India, which is located in the western part of India. It is a semi-arid region. It has an area of 720 sq. km (Figure 1). Ahmedabad is located on the banks of Sabarmati River. It emerges as an important economic and industrial hub in India. Historically, it is called "Manchester of East" and recently declared as Heritage City of India by UNESCO. The city is governed by Ahmedabad Municipal Corporation (AMC) and Ahmedabad Urban Development Authority (AUDA). AUDA is responsible for both land use planning and strategic planning of the city. As per the AMC, 2006, the inner city is considered as Central Business District (CBD) with dense building structures. Area east of Sabarmati river has old and dense building structures with several industries and area west of Sabarmati river is predominantly having residential, commercial and isolated buildings around farm lands.

3. Spatial Data Base Creation

Remote sensing data from Indian Remote Sensing satellite (IRS-1A) and Resouresat-1 and 2 for the year 1992, 2005 and 2018 for the month between Octobers to December have been used in this study. These images cover the entire area study area with same season, which is important for change detection analysis as it minimises seasonal vegetation differences and the effects of varying sun positions. Each time frame is classified into two classes urban and non-urban. Non-urban consists of classes such as vegetation, fallow land, soil and water bodies. These images were classified using semi-automated and object-based image analysis method with the accuracy of 89% (Jain and Sharma, 2019).

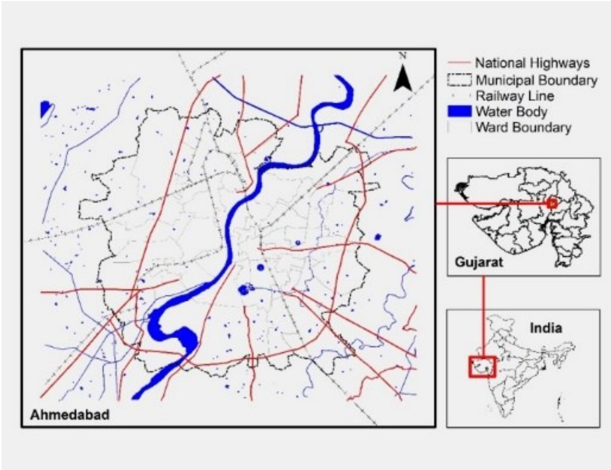
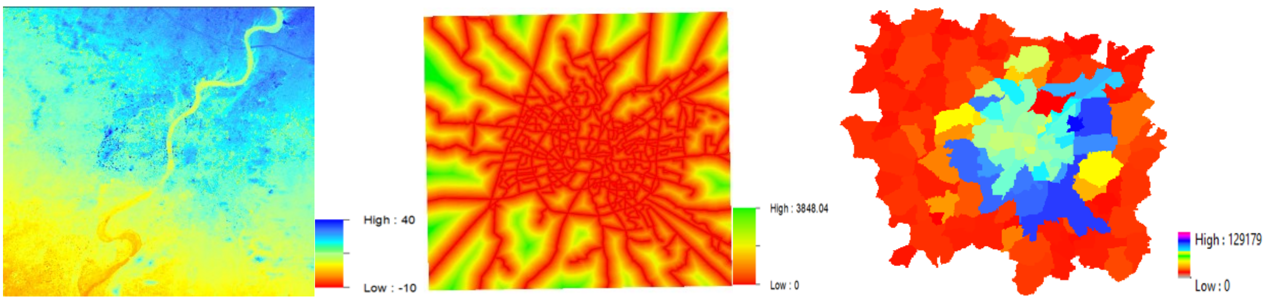


Figure 1: Study Area of Ahmedabad City

3.1 Auxiliary Data used

Driving factors are responsible for initiating the LULC change. These can be the categorised into, slope, elevation and infrastructure's proximity. Infrastructure's proximity such as proximity from road, highway is used. Slope is an important factor that drives the land cover changes, steep slope could become limiting factor for built-up area. However, most of built-up area is found to be in relatively flat area because of its relative easiness for building construction. All of driving factors are used for CA-MC model based on Artificial Neural Network. The auxiliary data for this study includes physical factors (DEM, slope), Euclidian distance to road networks and population data obtained from AMC. The Digital Elevation model (DEM) has been generated from Cartosat-1 stereo pair with 5 m spatial resolution.



(a) (b) (c)

Figure 2: Auxiliary data used for urban growth modelling. (a) Slope (b) Proximity to road (c) Population density

4. Methodology

In this section, the essential characteristics of the integrated model are discussed (Figure 3.) First, the land use maps of 1992, 2005 and 2018 were produced using IRS data and temporal change in land use were evaluated. Second, the main driving force for urban expansion were investigated and trained by ANN for optimal network. Then the CA-MC model is applied for simulating urban growth. In order to verify the results, the land use map was validated as per the reference maps using kappa index. Finally, the model is used to simulate future land use maps of 2018 and 2032.The detailed description of models are following.

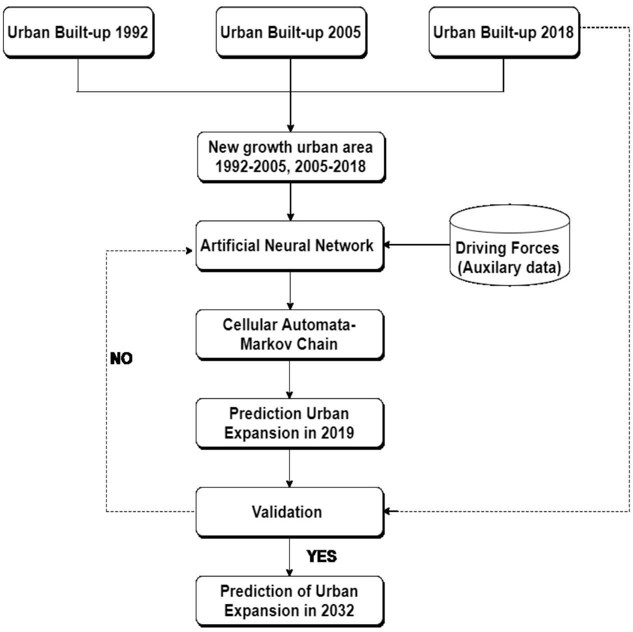


Figure 3: Flowchart of the ANN-CA-MC method in simulating urban expansion

4.1 Artificial Neural Network

Artificial Neural Networks (ANNs) are widely used for modelling with self-adapting, self-organizing, and self-learning abilities (Pijanowski et al., 2002; Park et al., 2011; Berberoğlu et al., 2016). The most frequently used and efficient feed-forward, error Back-Propagation Three-Layer Perceptron (BP-TLP) is adopted to simulate urban expansion owing to its simplicity, ease of training, and its abilities for reasonable associative memory and prediction (Rumelhart et al., 1986). The selected three-layer ANN is composed of an input layer, a hidden layer, and an output layer with 3 input nodes, 50 hidden nodes, and 2 output nodes (Figure 4). Each input node represents an independent variable and the number of hidden nodes significantly affects ANN performance (Hagan et al., 1996). Too few nodes will cause a significant prediction error, while too many will prolong the training process and lead to overfitting. In this study, the number of optimal hidden nodes was set to 50 by trial and error method. The selection of this number was based on model performance and network simplicity. The expected output has two possibilities of (1, 0) and (0, 1). The former indicates that the cell in question meets the expectation of conversion to urban, and the latter signifies non-urban cells. The network was trained stepwise iteratively with a targeted Root Mean Square Error (RMSE) of 0.001 between the model output and the reference data. However, reaching this RMSE threshold might cause overfitting, which was avoided by setting the number of training epochs to 500. These two parameters reduced the possibility of overfitting by early stopping. After an initial weight was assigned to each input variable, the ANN started to 'learn' from the training samples by adjusting the weights between neurons in response to the RMSE between the modelled output and the observed value. The training was terminated according to the targeted RMSE threshold, or the number of training epochs, whichever was reached first.

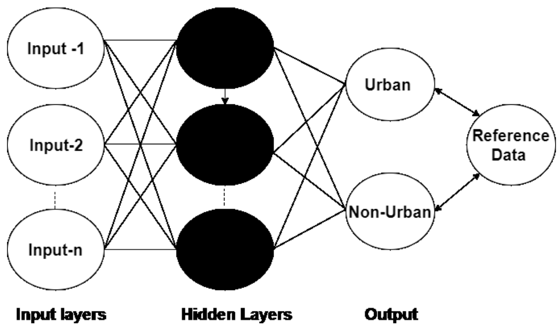


Figure 4: The architecture of the ANN adopted in this study.

4.2 Markov Chain-Cellular Automata

Markov chain is employed to predict the probability of urban land cover class change from one state to another by taking into account the past land cover change trend. Markov chain is a series of random values whose probabilities at a time interval depends on the value of the previous time (Surabuddin et al., 2013). Markov chains output describes as transitional probability matrix (equation 1):

$$p = (P_{ij}) = \begin{matrix} p11\ p12 \ldots\ldots p1n \\ p21\ p22 \ldots\ldots p2n \\ pn1\ pn2 \ldots\ldots pnn \end{matrix} \qquad (1)$$

The probability of changes from (i^{th}) class into (j^{th}) class is described as a transformation probability (P_{ij}); n is the number of classes with the constraint below (equation 2). The transition probability matrix is a set of conditional probabilities for the cells in the model to go to a particular new state. (Akin et al., 2014)

$$0 \le P_{ij} \le 1 \ (i, j = 1,2,3, \ldots, n) \ (2)$$

$$\sum_{i=1}^{n} P_{ij} = 1 \ (i, j = 1,2,3, \ldots, n)$$

Markov chains model is obtained by (equation 3):

$$P(n) = P(n-1)P_{ij} \qquad (3)$$

Where (n) is state probability of any times and $P(n-1)$ is preliminary state probability. Markov chain calculates probability of changes but does not represent the spatial explicit and location of changes. This limitation is fulfilled by combing with CA in order to minimise the weakness of the method. But for incorporating drivers of urban growth ANN is combined along with CA-MC model.

4.3 Accuracy Assessment

Accuracy test of urban growth model is an important part of the model whether the result of the model can be used by policy maker or not. Simulation and prediction techniques are ineffective and will have no scientific importance, if these techniques or models are not validated. Hence, the accuracy assessment of projected urban growth maps using validation methods is an extremely important step in urban growth modelling. The accuracy of the predicted land cover model was assessed using the kappa statistics. Overall accuracy and Kappa index are the most significant coefficients used to validate urban growth simulation (Yang et al., 2011; Al-sharif and Pradhan, 2013).

4.4 Model Implementation

The model was implemented as shown in figure 4. First, land cover maps in 1992, 2005 and 2018 were used to create two urban expansion maps (1992–2005, 2005–2018) through spatial overlay. Both transformed and untransformed samples were used as inputs to the ANN. Second, the ANN was cross-validated at least 500 times to optimize the selected calibration variables. The predicted result in 2018 was validated against the observed 2018 urban area using kappa index. With kappa value equal to 0.94 and mean square error as 0.083, the trained model indicates reasonable results comparison to other random allocation, this particular model was then applied to predict urban expansion in 2032. After the model passed the validation assessment, it was used to simulate urban area in 2032 with the land cover in 2018 as the baseline scenario under the assumption that future urban expansion would behave identically to what had occurred in the past. CA-MC model was applied to create the projection of 2018 land cover based on 2005 and 1992 land cover and projected model is being evaluated with kappa statistics based on reference image.

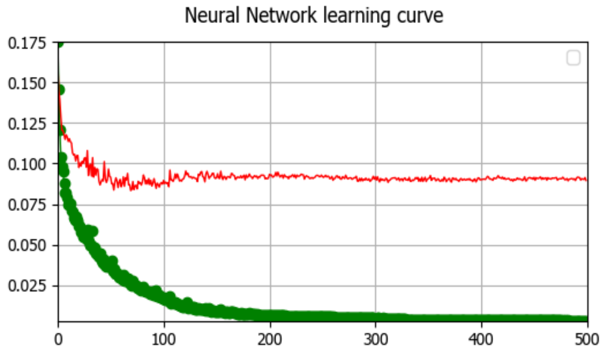


Figure 5: Optimised neural network adopted learning curve and loss function with learning curve (red) and loss function curve (green).

5. Results

The urban built-up has increased significantly between 1992, 2005, 2018 (Table 1 and Figure 6). It has indeed increased by 36% in past 27 years. Built-up area increased from 146.67 km^2 in 1992 to 193.82 km^2 in 2005 and 230 km^2 in 2018. The increase in built-up area can be attributed to the population growth and settlement expansion, these scenarios had culminated into conversion of natural vegetation and open spaces to built-up areas. Based on the rate of change between 1992 and 2018, the predicted urban built-up areas revealed that by 2032, built up area would increase to 275.76 km^2 in 2032. This is increase of 19.5 % of urban built-up area. The simulated results show that the increase in built-up area are closely associated with the existing built-up areas. Urban sprawl has mostly occurred in north-west and south-west part, it has attracted many automobiles industries and builders for construction of new housing colonies, hence there is low-density urban sprawl on the city's western periphery. The projected urban growth in 2032 indicates in future more industries and housing societies will come to this place making it denser (Figure 7). The growth in this area is attributable to population growth and also people shifting from the interior of city to the outskirts' due to inflation of price of land and denser built-up surrounding with less green spaces and more polluted air.

Table 1 Change in built-up area during (1992-2018)

Land cover	1992 Area (km^2)	2005 Area (km^2)	2018 Area (km^2)
Urban	146.67	193.82	230.0
Non-Urban	574.14	526.99	491.29

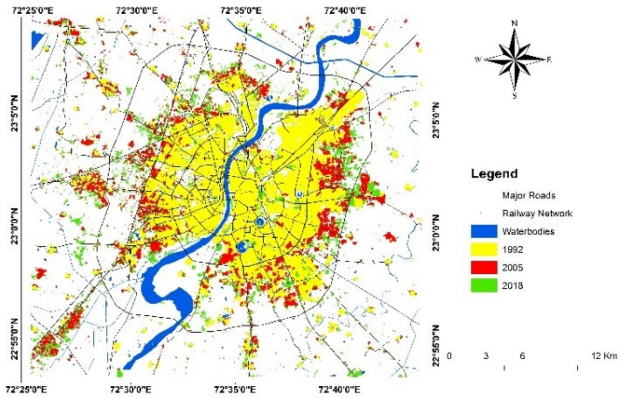


Figure 6: Urban Built-Up Change during (1992-2005) and (2005-2018) period

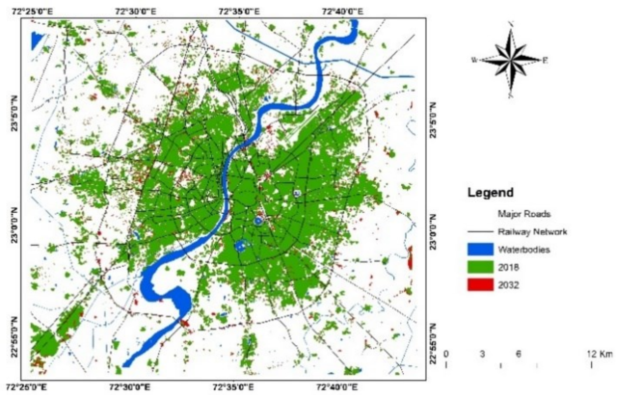


Figure 7: Predicted urban growth in 2032 and existing 2018 urban area for Ahmedabad city.

6. Conclusions

At global level, urban sprawl has become a challenge for sustainable development in cities and it has negative impact on environment as well as human society. Urban built-up growth has increased by 36% from 1992 to 2018 in Ahmedabad and if this trend continues, the other land cover classes such as agricultural land, vacant land will be changing to built-up areas. The validation results demonstrate the possibility of using machine learning, such as ANN, to improve the capability of CA-MC simulation of urban expansion. In the simulation, the ANN plays the most important role for incorporating the relationship between the urban growth and socio-economic variables affecting urban expansion. It is predicted that urban areas will increase to 275.76 km^2 by 2032. The predicted result from this model shows that newly urbanized areas in 2032 are based on the urban growth in the past. The most affected regions for future urban growth will be western part of Ahmedabad. Overall, this model can be used to provide relevant and useful information for urban planners and local government decision makers.

Acknowledgements

The author would like to thank Shri D.K. Das, Director, Space Applications Centre for the Institutional support and encouragement during the course of the study. We would like to express our sincere gratitude to Dr. Raj Kumar, DD, EPSA and Dr. Markand Oza, Head, CGDD, Space Applications Centre for their valuable comments and encouragement.

References

Aburas, M.M., Y.M. Ho., Mohammad F.R., and Z.H Ashaari (2016). The simulation and prediction of spatio-temporal urban growth trends using cellular automata models: a review, International Journal of Applied Earth Observation and Geoinformation, 52, 380–389. doi: 10.1016/j.jag.2016.07.007.

Akin A., S. Aliffi, and F. Sunar (2014). Spatio-temporal urban change analysis and the ecological threats concerning the third bridge in Istanbul City, Int. Arch Photogramm Remote Sensing Spat Inf SCI 40 (7), 9.

AMC (2006). City Development Plan Ahmedabad 2006-2016. Ahmedabad Municipal Corporation, Ahmedabad.

Al-sharif, A.A.A., and B. Pradhan (2013). Urban sprawl analysis of Tripoli metropolitan city (Libya) using remote sensing data and multivariate logistic regression model, Journal of the Indian Society of Remote Sensing, 42(1), 149–163.

Arsanjani, J.J., M. Helbich, W. Kainz, and A.D. Boloorani (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion, International Journal of Applied Earth Observation and Geoinformation, 21, 265–275.
doi: 10.1016/j.jag.2011.12.014.

Batty, M., Y. Xie, and Z. Sun (1999). Modelling urban dynamics through GIS-based cellular automata, Computers, Environment and Urban Systems, 23 (3), 205–233. doi:10.1016/S01989715(99)00015-0.

Berberoğlu, S., A. Akin, and K.C. Clarke (2016). Cellular automata modelling approaches to forecast urban growth for adana, Turkey: a comparative approach, Landscape and Urban Planning, 153, 11–27.
doi: 10.1016/j.landurbplan.2016.04.017

Bihamta, N., A. Soffianian., S. Fakheran. and M Gholamalifard. (2015). Using the SLEUTH urban growth model to simulate future urban expansion of the Isfahan metropolitan area, Iran, Journal of the Indian Society of Remote Sensing, 43 (2), 407–414. doi:10.1007/s12524-014-0402-8.

Clarke, K.C., S. Hoppen, and L. Gaydos (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area, Environment and Planning B: Planning and Design, 24 (2), 247–261.doi:10.1068/b240247.

Jain, G.V. and S.A. Sharma (2019). Spatio-temporal analysis of urban growth in selected small, medium and large Indian cities, Geocarto International,34:8,887-908, DOI:10.1080/10106049.2018.1450450

Jiang, B. and X. Yao (2010). Geospatial Analysis and Modelling of Urban Structure and Dynamics. Springer, Dordrecht, The Netherlands, ISBN 9048185718.

Jat, M.K., M. Choudhary, and A. Saxena, (2017). Application of geo-spatial techniques and cellular automata for modelling urban growth of a heterogeneous fringe, The Egyptian Journal of Remote Sensing and Space Science, 223-241, doi:10.1016/j.ejrs.2017.02.002

Hagan, M.T., H.B Demuth, and M.H. Beale (1996). Neural network design. Boston, MA: PWS Publishing.

Li, X., Y. Chen, X. Liu, X. Xu, and G. Chen (2017). Experiences and issues of using cellular automata for assisting urban and regional planning in China, International Journal of Geographical Information Science, 31 (8), 1606–1629.
doi:10.1080/13658816.2017.1301457.

Liu, Y. (2012). Modelling sustainable urban growth in a rapidly urbanising region using a fuzzy-constrained cellular automata approach, International Journal of Geographical Information Science, 26 (1), 151–167.
doi:10.1080/13658816.2011.577434.

Li, X. and A.G.O. Yeh (2000). Modelling sustainable urban development by the integration of constrained cellular automata and GIS, International Journal of Geographical Information Science, 14 (2), 131–152.
doi:10.1080/136588100240886.

Mustafa, A., M. Cools, I. Saadi, and J. Teller (2017). Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (HUEM), Land Use Policy, 69 (Supplement C), 529–540.
doi: 10.1016/j.landusepol.2017.10.009.

Park, S., S. Jeon, S. Kim, and C. Choi (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. doi:10.1016/j. landurbplan.2010.09.001.

Pijanowski, B.C., D.G. Brown, B.A. Shellito, and G.A., Manik (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.

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

Santé, I., A.M. Garcia, and R. Crecente (2010). Cellular automata models for the simulation of real-world urban processes: a review and analysis. Landscape and Urban Planning, 96 (2), 108–122.

Sudhira, H.S., T.V. Ramachandra, and K.S. Jagadish (2004). Urban sprawl: metrics, dynamics and modelling using GIS, International Journal of Applied Earth Observation and Geoinformation, 5 (1), 29–39. doi:10.1016/j. jag.2003.08.002.

Surabbuddin, M.M., N. Sharma, M. Kappas, and P.K. Garg (2013). Modelling of spatio-temporal dynamics of land use and land cover of Brahmaputra river basin using geoinformatics techniques. Geocarto Int 28(7), 632-656.

Wang, H., S. He, X. Liu, L. Dai, P. Pan, S. Hong, and W. Zhang (2013). Simulating urban expansion using a cloud-based cellular automata model: A case study of Jiangxia, Wuhan, China. Landscape and Urban Planning,110,99–112.doi: 10.1016/j.landurbplan.2012.10.016.

Xu.T., J. Gao, and G. Coco (2019). Simulation of urban expansion via integrating artificial neural network with Markov chain – cellular automata, International Journal of Geographical Information Science, 33(10), 1960–1983 https://doi.org/10.1080/13658816.2019.1600701

Yang, W., F. Li, R. Wang, and D. Hu (2011). Ecological benefits assessment and spatial modelling of urban ecosystem for controlling urban sprawl in Eastern Beijing, China, Ecological Complexity, 8 (2), 153–160. doi: 10.1016/j.ecocom.2011.01.004.

Zhang, H., X. Jin, L. Wang, Y. Zhou and B. Shu (2015). Multi-agent based modelling of spatiotemporal dynamical urban growth in developing countries: simulating future scenarios of Lianyungang city, China, Stochastic Environmental Research and Risk Assessment, 29 (1), 63–78. doi:10.1007/s00477-014-0942-z.