Comparing ANN and CART to Model Multiple Land Use Changes: A Case Study of Sari and Ghaem-Shahr Cities in Iran

M. Ahmadlou[1], M. R. Delavar*[2], A. Tayyebi[3]

[1]M. Sc. Student, GIS Dept., School of Surveying and Geospatial Eng., College of Eng., University of Tehran, Tehran, Iran
m_ahmadlou@ut.ac.ir

[2]Assoc. Prof., Center of Excellence in Geomatic Eng. in Disaster Management, School of Surveying and Geospatial Eng., College of Engineering, University of Tehran, Tehran, Iran
mdelavar@ut.ac.ir

[3]University of California-Riverside, Center for Conservation Biology, Riverside, CA, USA
amin.tayyebi@gmail.com

(Received: January 2016, Accepted: May 2016)

Key Words: Classification and Regression Tree, Artificial Neural Network, Variable Importance Index, Multiple Land Use Changes, Receiver Operating Characteristic, Total Operating Characteristics

Abstract

Most of the land use change modelers have used to model binary land use change rather than multiple land use changes. As a first objective of this study, we compared two well-known LUC models, called classification and regression tree (CART) and artificial neural network (ANN) from two groups of data mining tools, global parametric and local non-parametric models, to model multiple LUCs. The case study is located in the north of Iran including cities of Sari and QaemShahr. Urban and agricultural changes over a period of 22 years between 1992 and 2014 have been model. Results showed that CART and ANN were effective tools to model multiple LUCs. While it was easier to interpret the results of CART, ANN was more effective to model multiple LUCs. In earlier studies, despite using CART, the extraction of effective factors of LUCs using a precise index has not been considered efficiently. As a second objective, this study performed a sensitivity analysis using variable importance index to identify significant drivers of LUCs. While ANN was a black box for sensitivity analysis, CART identified significant delivers of LUCs easy. The results showed that the most important factors were distance from urban areas and rivers while aspect was the least effective factor. As a third or final objective of this study, the recently modified version of receiver operating characteristics (ROC) called total operating characteristic (TOC) as well as ROC were used for accuracy assessment of CART and ANN. The area under the ROC curves were 78% and 75% for urban changes for ANN and CART models, respectively. The area under the ROC curves were 72% and 65% for agricultural changes for ANN and CART models, respectively. We found that although TOC and ROC were similar to each other, TOC proved more informative than conventional ROC as a goodness-of-fit metric. The outcome of this study can assist planners and managers to provide sustainability in natural resources and in developing a better plan for future given the needs to understand those contributing factors in urban and agriculture changes.

* Corresponding Author

1. Introduction

A city is a complex system where humans interact through their activities with the environment [1]. The reaction of cities against these activities has accelerated expansion of urban areas introducing a phenomenon known as land use change (LUC). It is widely known [2, 3, 4, 5] that LUC is a very complex process with multiple drivers of LUC operating on a variety of spatio-temporal levels including demographic (e.g., population growth), economic (e.g., gross domestic product, [6]), bio-physical (e.g., elevation and soil), institutional (e.g., policies) and cultural issues [7].

LUC leads to changes in climate [8], economy [9], food security [10], biodiversity and ecosystem, which are threats to human life and well-being. Therefore, many earth, environmental and atmospheric science applications are concerned about spatial distribution of land use such as agriculture, forest and urban [11, 12]. During the last three decades, urbanization in Iran has led to the expansion of housing and industry into previously open, low-populated areas that were originally natural areas and agricultural lands [13, 14]. As a result, disturbance of agricultural and forest areas has affected food safety of human populations and reduced biodiversity [15]. Some recovery of forests from shrub lands and abandoned agricultural land has occurred recently [16].

Different disciplines (e.g. economics, engineering, and environmental science) have applied a variety of data mining tools to extract underlying patterns of data [17]. Data mining methods generally include four groups [18]: global and local models each of which can be either parametric or non-parametric. Global models perform modeling using all available data while local models divide the data into separate subsets and fit separate models on each of the subsets. Parametric models have a fixed structure before the modeling process starts and they are model driven [18] while non-parametric models are data driven and usually do not have a fixed model structure or their model structure is unknown before the modeling process [18]. These four groups of models have been applied to quantify the relationship between dependent and multiple independent variables of LUC.

Among these four groups of models, global parametric models (GPMs) and local non-parametric models (LNPMs) have often been applied in various disciplines. As an example of GPMs in land change science, Tayyebi et al. (2008) [4] and Pijanowski et al. (2014) [19] used artificial neural network (ANN) to model and predict urbanization in Tehran, Iran and the United States, respectively. Clarke et al. (1997) [20] used a cellular automaton (CA) model, called SLEUTH, to predict urbanization in the San Francisco Bay. Shan et al. (2008) [21] used genetic algorithms to enhance the efficiency of transition rule calibration in CA urban growth modeling. Finally, Pontius and Batchu (2003) [22] used IDRISI software to simulate land cover change in India and assess the power of IDRISI software. Classification and Regression Tree (CART, [23]) is a LNPM that has been widely used in data mining, including predicting business failure [24] and hypertension in people [25].

GPMs and LNPMs have their advantages and disadvantages. GPMs can characterize LUC as earlier studies have shown [26]. It has been shown, if the accuracy of LUC model is the only concern, that GPMs usually provide better performance than LNPMs [27]. However, GPMs suffer from some limitations as well. For example, most of the GPMs assume that the spatial predictors have to follow a normal distribution for proper modeling while drivers of LUC rarely have such distributions [28]. Furthermore, most of the functions that GPMs use require prior knowledge about the relationship between input and output. Auto-correlation is another common issue concerning spatial predictors that affect the goodness-of-fit of LUC models [29]. GPMs usually suffer from auto-correlation since input variables are not independent from each other. Further problems arise with GPMs when more spatial predictors are included in the modeling process of LUC [30].

Compared to GPMs, LNPMs do not need to have prior knowledge about the distribution (i.e., normal distribution) of data, and the form and parameters of the functions (i.e., linear or non-linear, means and standard deviations; [31]). LNPMs work easily with outliers and have the ability to detect them in modeling [32]. Moreover, the model structures of the LNPMs are not fixed and the model typically grows to fulfill the complexity in data [33]. LNPMs are able to detect non-linear relationships in data, variable selection, data transformation and variable reduction [34]. LNPMs usually have a simple structure which can be easily understood and interpreted [35]. To date, very few studies have compared the potential of LNPM and GPM to model LUC [18].

The first objective of this study is to model multiple LUCs using CART and ANN in order to explore the interactions and competitions between multiple LUCs. The study was conducted in the north of Iran including cities of Sari and QaemShahr. We specifically modeled urban and agriculture changes over a period of 22 years between 1992 and 2014. Few studies have been focused on multiple land use change modeling (e.g., Verburg et al., 2001 [36]; Li and Yeh 2002 [37]; Ballestores et al., 2012 [38]; Ralha et al., 2013 [39]; Tayyebi and Pijanowski 2014 [27]). While the previous studies have utilized the CA, ANN, and CART models, the ability of CART in variable selection has not been considered as yet. This advantage of CART would help to avoid collinearity issue, which is the non-independence of predictor factors [40], among the explanatory variables as well as would help decision makers to identify the most important influential factors [41]. Therefore, the second objective of this study is to perform sensitivity analysis using variable importance index (VIM) to variable selection.

There is no single calibration metric for LUC models that can provide unbiased interpretation of model performance [42]. Thus, many calibration metrics have been proposed by the LUC community. Among them, receiver operating characteristic (ROC) is one of the most common accuracy metric that has been used in land use science to calibrate LUC models. ROC compares the simulated map with the reference map for each given threshold using contingency table. However, ROC fails for cases where some types of error are more important than other types of error [43]. ROC also fails to reveal the size of each entry in the contingency table for each threshold [44]. Pontius and Si (2014) [45] recently introduced the total operating characteristic (TOC) to rectify the limitations of ROC. The third objective of this study is to calibrate CART and ANN for modeling multiple LUCs using ROC and TOC indices. We specifically compared ROC and TOC for calibration of urban and agricultural changes. Figure 1 shows the structure of the paper.

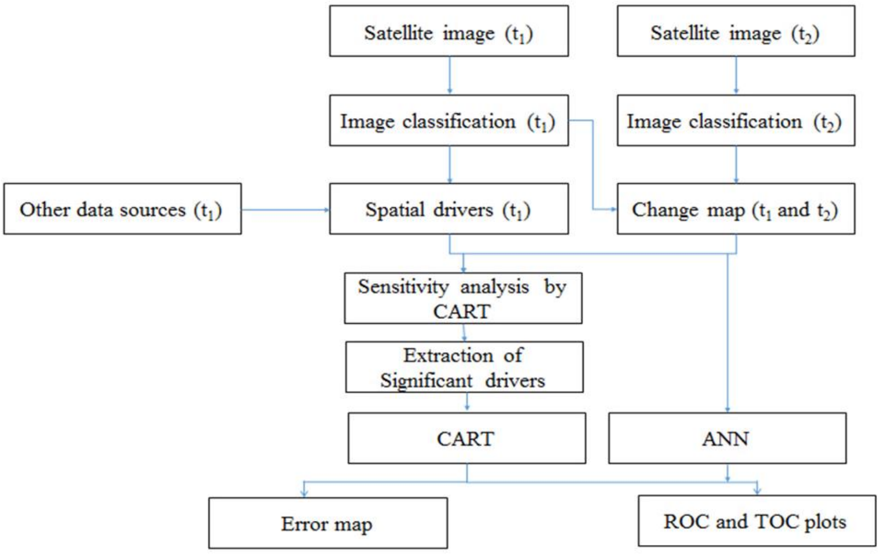


Fig. 1: Paper structure

2. Study Area and Methods

2.1 Study Area

The study area is part of Mazandaran Province located in northern Iran including two cities of Sari which is the capital of Mazandaran Province and QaemShahr. These cities are located between the northern slopes of the Alborz Mountains and southern coast of the Caspian Sea and have a hot-summer Mediterranean climate. Winters are cool and rainy whilst summers are hot and humid. The region's economy is mostly based on food production (e.g. milled rice, dairy products, canned meat, etc.). Sari is the most populous city of Mazandaran. According to a recent survey in 2011, the population of Sari almost doubled (from 141,020 in 1986 to 296,417 in 2011). Also, the population of QaemShahr almost doubled from 109,288 in 1986 to 196,050 in 2011.

2.2 Classification and Regression Tree

The logical steps of CART model are as follows: It starts with the entire set of input drivers at the root node (node at the top of the tree) and recursively divides the input data according to the independent variables that introduce the highest purity, which refers to the degree to which the leaf nodes (nodes without children) are made up of cases with the same land use class, and homogeneity within the internal nodes. Among the various variables, a variable is selected for a given node that can increase node purity. There are several criteria for data division in each of the nodes the most notable of which is the Gini index which is one of the attribute selection measures for the target variable with nominal values [46]. The Gini index at node t was determined using Eq.1:

$$Gini\ (t) = \sum_{i \neq j} P(W_i) \times P(W_j) \tag{1}$$

$P(w_i)$ in Eq. 1 is the relative frequency of class i. The process of tree growth continues until the highest purity at the leaf nodes (nodes without children) is achieved.

2.2.1 Identification of Significant Drivers

To determine the importance of each LUC driver in the modeling process and extract the most important drivers for decision-making, we used the VIM. CART implements Eq. 2 to identify the relative importance of variables [47]:

$$VIM(x_j) = \sum_{t=1}^{T} \frac{n_t}{N}\ \Delta Gini\ \left(S(x_j, t)\right) \tag{2}$$

Where n_t is the number of observations that belong to node t of the tree, N is the total number of observations, n_t/N is the proportion of observations that belong to node t of the tree, T is the total number of nodes and x_j are spatial drivers of LUCs. $\Delta Gini$ is the difference between the calculated Gini index at node t and its parent node (Eq. 3, [48]).

$$\Delta Gini\ \left(S(x_j, t)\right) = Gini(t) - P_l Gini(t_l) - P_r Gini(t_r) \tag{3}$$

P_l and P_r are a fraction of the samples which go which go to the left and right node of node t.

2.3 Artificial Neural Network

Artificial neural networks (ANNs) are data mining tools capable of capturing non-linear associations underlying land use transformations [49]. The multi-layer perceptron, as used in this paper, typically includes an input, a hidden layer, and an output layer. These three layers are connected to each other in a feed-forward manner [18, 49, 51]. During the training, the initial model parameters are modified repeatedly until model outcomes represent the observed output as accurate as possible. ANNs start by randomly assigning the weightings and calculating the mean squared error (MSE), after which the cycle continues until a terminating criterion is met [52]. The MSE is measured by calculating the difference between the output from the network and the actual output presented in the training phase [19]. If the level of MSE 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.

2.4 Accuracy and Error Estimation

The primary output of the CART and ANN models are a suitability map after the training run. The suitability map shows the suitability of cells for either persistence or change. The suitability maps from both models were then converted to the simulated maps using the quantity of changes between the two dates. For this purpose, two referenced land use maps separated in time (t1 and t2) were compared to quantify the amount of LUCs and obtain the referenced change map. To assess the model performance, the simulated map was compared with the referenced change map. We used two well-known calibration metrics including ROC [22] and TOC [45] to assess the performance of the both models.

2.4.1 Receiver Operating Characteristic

ROC uses a series of thresholds to convert the suitability map to the simulated map. The values greater than these thresholds in the suitability map are set to 1 (meaning LUC in the desired cell) and the other values are set to zero (meaning non-change in the desired cell). ROC then compares the simulated map with the reference map in t_2 for each given threshold using the contingency table (Table 1). The disagreement (false positives, FP and false negatives, FN) and agreement (true positive, TP and true negative, TN) rates are the measure of model performance. After generating the contingency table (Table 1), sensitivity (X_t) and specificity (Y_t) are calculated for the entire thresholds to plot the ROC curve (*t* indicates the threshold values). Where sensitivity is the proportion of correctly predicted change cells, and specificity is the proportion of correctly predicted no-change cells [54]. ROC curve consists of connecting the series of calculated ratios (X_t and Y_t) for each threshold. The area under curve (AUC) represents accuracy of the model.

Table 1: Using contingency table to compare simulated and referenced land use maps

		Reference map	
		Change	Non-change
Simulated map	Change	True Positive (TP)	False Negative (FN)
	Non-change	False Positive (FP)	True Negative (TN)
		P = TP + FP	Q= FN + TN

2.4.2 Total Operating Characteristic

Although ROC has been frequently applied in LUC science [15], there have been numerous studies arguing against using ROC over the last two decades [45]. One of the most important limitations is that ROC fails to reveal the size of each entry in the contingency table for each threshold. Pontius and Si (2014) [45] recently suggested total operating characteristic (TOC) as an alternative to overcome the limitation of ROC. Although TOC creates a curve very similar to ROC, TOC curve is different from ROC curve in different ways. For example, TOC shows four members of the contingency table at once on TOC curve for each threshold [55, 53]. The TOC is more intuitive than ROC since it provides results based on the actual units in the contingency table instead of unit-less values such as ROC curve. Therefore, TOC shows the four members of the contingency table in real units rather than in percentages as in ROC.

TOC plot shows the total number of cells on the horizontal axis, P + Q, and the total number of changed cells in the reference map on the vertical axis, P. Although TOC has maintained all the properties of ROC, for each threshold in the TOC curve, there are four entries in the contingency tables ([45], Table 1). The X_t and Y_t axes in the TOC curve are equal to true positive + false negative and true positive, respectively. To plot the TOC curve (Figure 2), we must first draw a parallelogram, vertices of which are equal to (0, 0), (P, P), (Q, 0), and (P+Q, P). P and Q indicate the number of observations which underwent changes or no changes between the two dates, respectively. The unit on both axes indicates the number of observations, considered as the number of cells for the given case study. For example, the horizontal axis varies from 0 to 250 thousand cells, which is the size of the study area (P + Q). Moreover, the coordinate on the horizontal axis refers to the number of changed cells in the simulated map for a given threshold (True Positive + False Negative). In the same way, the vertical axis varies from 0 to 30 thousand cells, which is the number of changed pixels in the reference map. Furthermore, the coordinate on the vertical axis refers to the number of cells which are predicted as change and have been actually changed in the reference map (TP). Then, TOC curve is drawn by X_t and Y_t which were calculated using a contingency table. For each threshold, four numbers in the entries of probability table are given in Figure 2. The area under ROC curve is equal to the ratio of the area of TOC curve within the parallelogram to the whole area of the parallelogram.

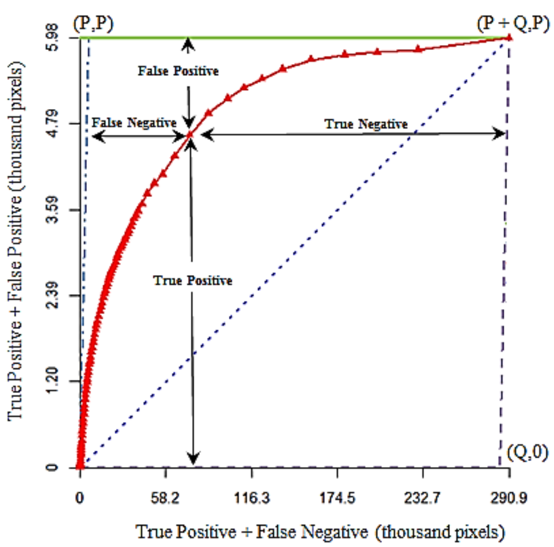


Fig. 2: Total Operating Characteristic (Pontiuse and Si, 2014)

3. Data Preparation and Model Implementation

3.1 Data Preparation

Since Mazandaran Province has abundant natural and agricultural resources to feed the country every year, it is essential to study the LUC which is in fact a threat to Iran's natural resources. The main data sources used in this study were two Landsat images with 30m by 30m spatial resolution in 1992 (TM) and 2014 (ETM^{+}). The two images were classified into urban, agricultural, river, forest and barren classes using maximum likelihood classification method. The overall accuracies (kappa index) of image classification for 1992 and 2014 were 85.5% and 89.3%, respectively. Then, land use change map was obtained by comparing the two reclassified images (output of CART and ANN).

Factors considered as drivers of change between 1992 and 2014 were elevation, aspect, slope, distance to urban, agriculture, river and road. The spatial drivers used in this study are common drivers used in previous LUC studies [56, 57]. Urban areas and land use classes, including urban, forest, river and agricultural lands, were extracted from land use map in 1992. Also, transportation network data were taken from the OpenStreetMap (OSM) collaborative database. Digital Elevation Model (DEM) was downloaded from the Iran National Cartographic Center website (ASTER with a spatial resolution of 30 m). Distance maps in1992 were formed using Euclidean distance function provided in GIS environment, for urban, forest, river, road, agricultural variables (inputs of CART and ANN). Moreover, slope and aspect maps were calculated using DEM in GIS.

3.2 Model Implementation

CART and ANN use spatial drivers in 1992 and change and no-change areas between 1992 and 2014 as output (Figure 3). The entire study area contains 269,937 cells. Salford Predictive Modeler and LTM software were used to model multiple LUCs and conduct sensitivity analysis of variables [27]. We divided the entire study area into training (60%) and testing (40%) datasets using stratified random sampling approach [18]. We used 60% of data to train both models. We then applied the trained models to the rest of the data to test the performance of ANN and CART.

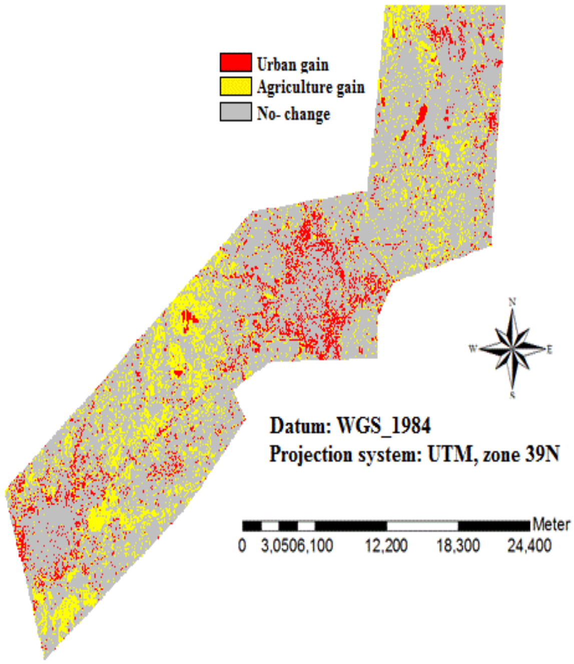


Fig. 3: Land use change map of the study area between 1992 and 2014

4. Results

4.1 Training run of ANN

ANN contains an input layer including explanatory variables, a hidden layer, and an output layer. The input layer has seven nodes corresponding to seven predictor variables. According to Kolmogorov's theorem [58], if n is the number of neurons in the input layer, $2n + 1$ hidden nodes can guarantee the perfect fit of any continuous functions so fifteen units was chosen for hidden layer. The output layer includes two nodes corresponding to our target land use classes, urban and agriculture. The MSE starts around 0.2 and drops thorough 100 cycles. We stopped training at 400 cycles where the MSE was 0.015.

4.2 Variable Importance Index

VIM index numbers have been calculated in the range of 0 to 100 (Figure 4). Distance to urban areas was the most important factor in the LUC. This variable is followed by the distance from the river factor. Likewise, the drivers of distance to agricultural areas, distance to roads and elevation are ranked in that order. The aspect variable has no influence on urban and agricultural changes and thus was removed from the list of predictor variables.

4.3 CART Simulation

CART model has identified some variables to stand for land use changes including slope, elevation, distance to agriculture, distance to urban, distance to road and distance to river. The percentage of each land use can be seen at leaf nodes (Figure 5). Few leaves included only two land use classes (e.g. nodes 1, 2, 8, 9, 10, 11, 12 and 13), whereas others had three land use classes (e.g. nodes 3, 4, 5, 6, 7, 16, 17). This indicated the fact that drivers which classified leaf nodes with two land use classes were more significant than drivers which classified leaf nodes with three land use classes. Nodes 3, 5, 11 and 17 had the highest concentration of urban gain. Similarly, nodes 1, 7, 15 and 19 had the highest concentration of agricultural gain (Figure 6).

CART model had nineteen rules in the tree structure (equal to the number of leaves). To make each of the rules associated with each of the leaf nodes, users should start from the root node and reach leaf nodes through internal nodes. For example, in Figure 5, the rule related to leaf node 1 can be expressed as follows: if the distance to urban is less than 157 m, if the distance to urban is less than 15 m, and if elevation is less than 21.5 m, then44% of cells belonging to leaf node 1 are no-change class and 56% of them are under agricultural land use class. In order to generate the simulated map with multiple land use classes, the probability of cells can be ranked in descending order based on the number of changed land use classes between the two reference dates. Using the number of referenced changed cells and the range of possible values from low to high, cells were assigned to the change and no-change land use classes.

4.4 Model Evaluation

The area under the ROC curves were 78% and 75% for urban changes for ANN and CART models, respectively (Figure 7a). The area under the ROC curves were 72% and 65% for agricultural changes for ANN and CART models, respectively (Figure 7b). Results showed that ANN was more successful in simulating urbanization and agricultural changes compared to CART. In addition, ROC curve for urbanization reached a higher TP rate compared to agricultural gain ROC curve. This showed the fact that the model trained better for urbanization rather than agricultural gain.

The area under ROC curve is equal to the ratio of the area under the TOC curve within the parallelogram to the total area of the parallelogram (Figure 8). From TOC curve, we can find out all the four parameters for each threshold. For examples, for a threshold of 0.62 in the TOC curve for urbanization (Figure 8a), the values for TP, TN, FP and FN were 23279, 185694, 9789 and 52284, respectively. Similarly, for a threshold of 0.76 in the TOC curve for agricultural gain (Figure 8b), the values for TP, TN, FP and FN were 43869, 145296, 19388 and 103493, respectively. Also, urban and agriculture uses were predicted by ANN and CART for 2036, respectively (Figure 9).

5. Conclusion and Discussion

In this paper, we used CART and ANN models to model multiple LUCs. We illustrated the predictive ability of CART and ANN for modeling multiple land-use changes for a study area in the north of Iran. The performance of the CART and ANN models were evaluated using ROC and TOC indexes. Overall, the obtained simulated maps show reasonable agreement with the observed data. We found ANN performed better than CART to model multiple LUCs. However, it was easier to interpret the results of CART due to the simplicity of CART structure and its effectiveness in simulating multiple LUCs.

Error (FN and FP) were higher for agricultural change modeling compared to urban change modeling. The reason could be due to the patterns of agricultural gain in the given region which was more dispersed around the existing houses. This fact makes it difficult for both models to simulate this complex pattern. Other reasons might be the presence of other factors (e.g., land values and population growth) that impact agricultural LUC which have not been considered in this study due to data limitation in Iran. CART was more successful in simulating urbanization rather than agricultural gain since urbanization happened mostly around the existing urban areas.

Although ANN was a black box to identify significant drivers of LUC, CART was very effective to rank significant drivers of LUC. CART model has recognized slope, elevation, distance to urban, agriculture, roads and water as the most important influential factors for modeling multiple land use change. The importance of distance to urban areas is undeniable due to the placement of all facilities within the city and access to required places for people including administrative and recreational facilities as well as markets. Given that water is one of the most important factors for agricultural production (converting the lands around the lands around the river to agricultural lands). We found that TOC was more informative than ROC for model evaluation and tackled the limitations of ROC. The advantage of TOC over ROC was that TOC presented four members of the contingency table at the same time on the curve for each threshold. Scientists can calculate a variety of accuracy metrics for each threshold using the entries in the contingency table for each threshold. Moreover, entries in the contingency table can help to identify segments of the curve that are important for a specific application. For example, LUC modelers are interested in identifying the threshold at which the ROC curve becomes flat. This provides information about the distribution of suitability values in suitability map. In land change modeling, there are always errors linked with LUC modeling. The first type of error, which is the most common one, comes from the classification of satellite image data. In this images introduced errors into CART and ANN study, we used Landsat imageries to create land use maps at two dates. Errors in both classified modeling since data error can affect both inputs and outputs of LUC models. The second type of error occurs during the modeling process since LUC models are not perfect. In the future, attempts should be made to minimize the effect of these errors using more advanced classification algorithms and high resolution images. In addition, this study only used two referenced land use maps to calibrate both models. Although the prediction rate is satisfactory for the current date, it might not be satisfactory for another date. Since the application of this study is intended for the planners, using a third time land use map for a better evaluation is highly recommended [59].

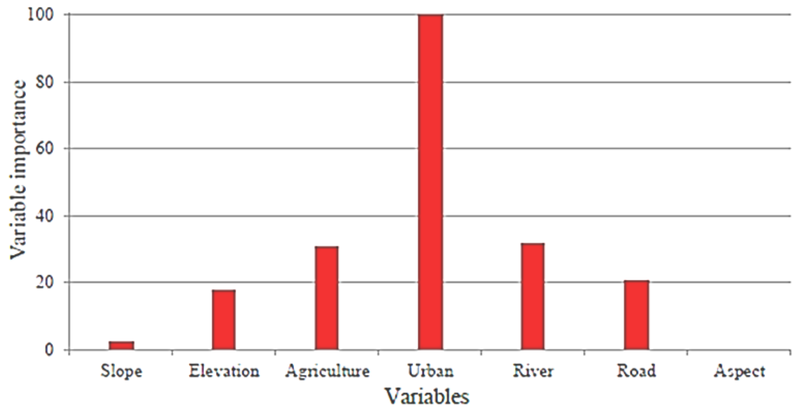


Fig. 4: Comparing the importance of predictor variables

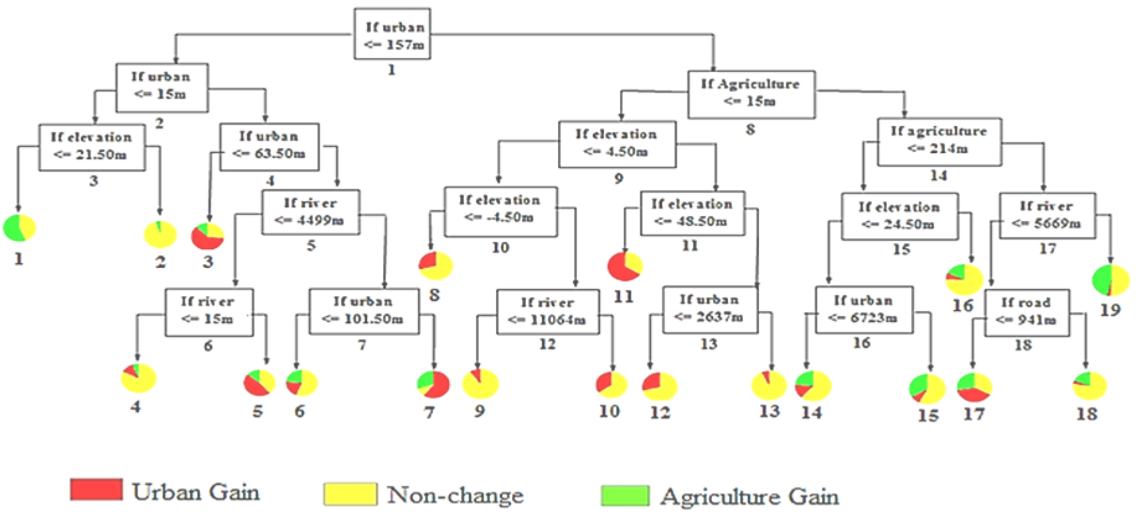


Fig. 5: Tree navigator of CART for importance drivers

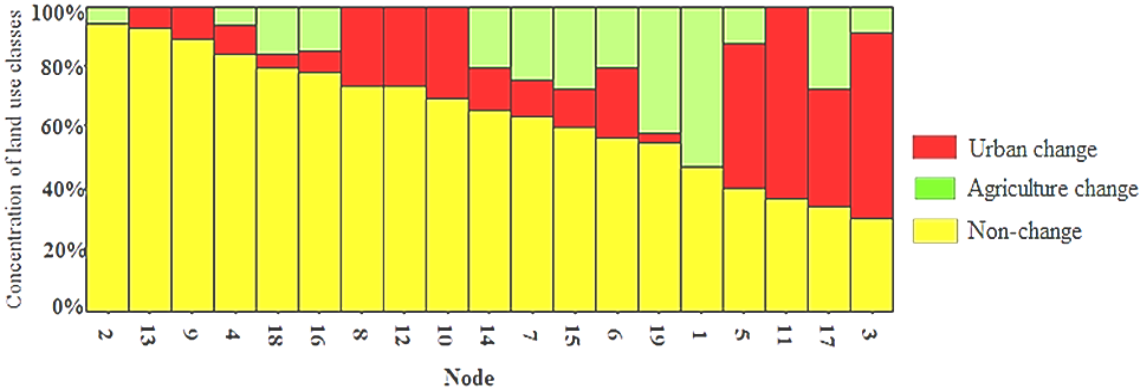
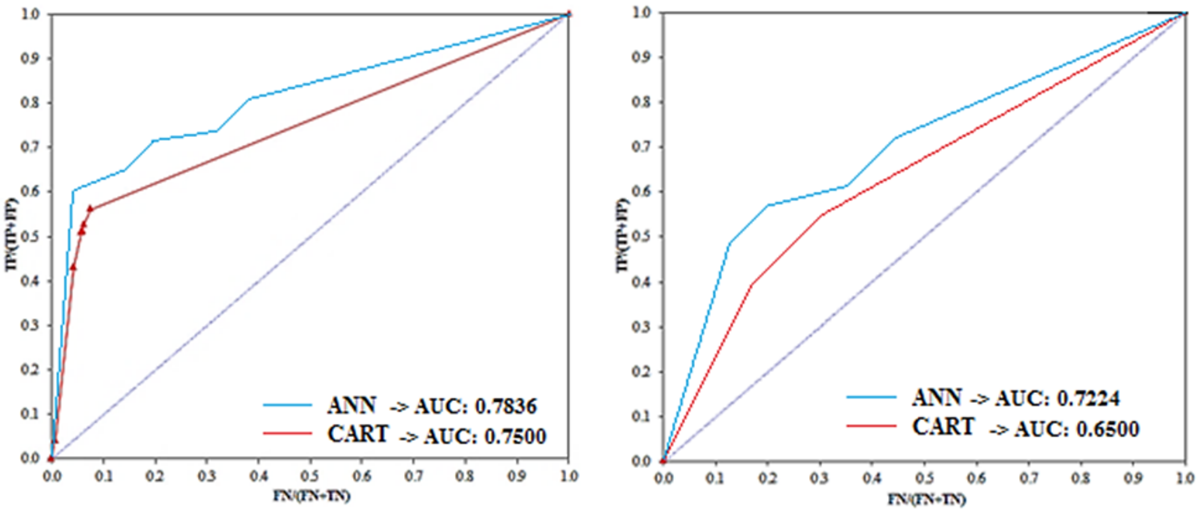
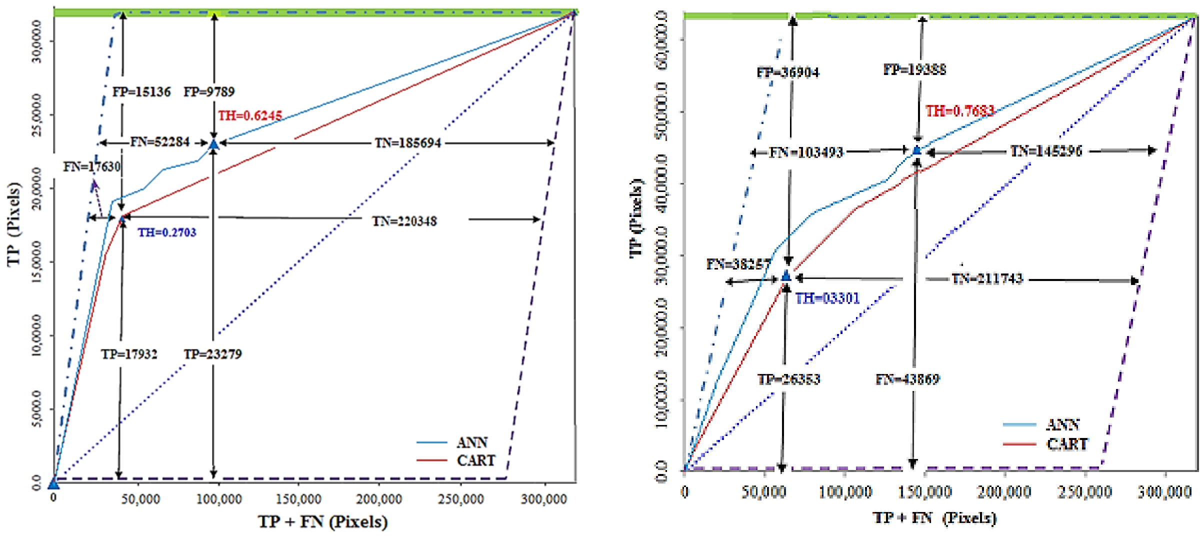


Fig. 6: Terminal nodes in CART



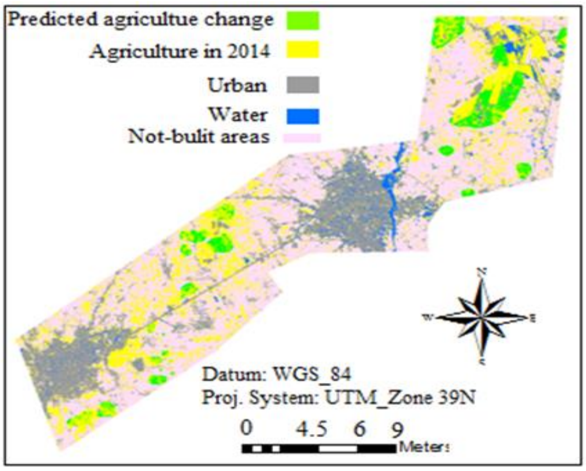
(a) **(b)**

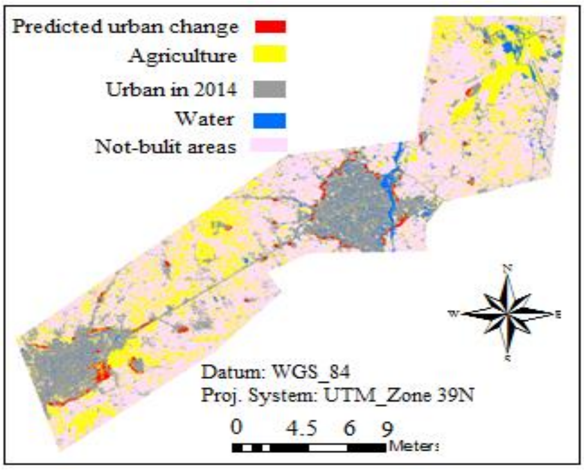
Fig. 7: The area under the ROC curves a) for urban change and b) for agriculture change



(a) (b)

Fig. 8: TOC curves a) for urban change and b) for agriculture change





(a) (b)

Figure 9. Predicted a) agriculture use by CART and b) urban use by ANN for 2036

References

[1] Batty, M., Y. Xie, and Z. Sun. (1999), “Modeling urban dynamics through GIS-based cellular automata,” Computers, Environment and Urban Systems, 23(3), pp. 205-233.

[2] Verburg, P. H., T. C. M. Nijs, J. R. Eck, H. Visser, and K. Jong. (2004), “A method to analyze neighborhood characteristics of land use patterns,” Computers, Environment and Urban Systems 28(6), pp. 667–690.

[3] Pontius Jr, R.G., W. Boersma, J. C. Castella, K. Clarke, T. de Nijs, C. Dietzel, Z. Duan, and et al.(2008), “Comparing input, output, and validation maps for several models of land change,” Annals of Regional Science, 42(1), pp. 11-47.

[4] Tayyebi, A., M.R. Delavar, S. Saeedi, and J. Amini. (2008), “Monitoring the urban expansion by multi-temporal GIS maps. In: Application of Remote Sensing and Imagery.” Integrating Generations, FIG Working Week 2008 and FIG/ UNHABITAT Seminar.

[5] Iiames, J. S., R. G. Congalton, and R. S. Lunetta. (2013), “Analyst variation associated with land cover image classification of Landsat ETM+ data for the assessment of coarse spatial resolution regional/global land cover products,” GIScience & Remote Sensing, 50(6), pp. 604-622.

[6] Tayyebi, A. H., M. R. Delavar, A. Tayyebi, and M. Golobi. (2010), “Combining multi criteria decision making and Dempster Shafer theory for landfill site selection,” International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Kyoto Japan, pp.1073-1078.

[7] Geist, H., W. McConnell, E F. Lambin, E. Moran, D. Alves, and T. Rudel. (2006), “Causes and Trajectories of Land-Use/Cover Change, Land-Use and Land-Cover Change, Global Change,” The IGBP Series 2006, pp. 41-70.

[8] Watson, R. T., I. R. Noble, B. Bolin, N. H. Ravindranath, D. J. Verardo, and D. J. Dokken. (2000), “Land Use, Land-use Change and Forestry: A Special Report of the Intergovernmental Panel on Climate Change,” Cambridge University Press.

[9] Long, H., G. K. Heilig, X. Li, and M. Zhang. (2007), “Socio-economic development and land-use change: Analysis of rural housing land transition in the Transect of the Yangtse River, China,” Land Use Policy, 24(1), pp. 141-153.

[10] Godfray, H. C. J., J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S.M. Thomas and C. Toulmin. (2010), “Food security: the challenge of feeding 9 billion people,” Science, 327(5967), pp. 812-818.

[11] NRC, (2005), “Radiative Forcing of Climate Change: Expanding the Concept and Addressing Uncertainties,” National Research Council, 208 p.

[12] NRC, (2007), “Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond.” National Research Council 456 p.

[13] Tayyebi, A., Pijanowski, B. C., & Tayyebi, A. H. (2011a), “An urban growth boundary model using neural networks, GIS and radial parameterization: An application to Tehran, Iran,” Landscape and Urban Planning, 100(1), pp. 35-44.

[14] Tayyebi, A., Pijanowski, B. C., & Pekin, B. (2011b), “Two rule-based urban growth boundary models applied to the Tehran Metropolitan Area, Iran,” Applied Geography, 31(3), pp. 908-918.

[15] Pijanowski, B. C., Tayyebi, A., Delavar, M. R., and Yazdanpanah, M. J. (2010), “Urban expansion simulation using geospatial information system and artificial neural networks,” International Journal of Environmental Research, 3(4), 493-502.

[16] Diaz, G. I., L. Nahuelhual, C. Echeverría, and S. Marín. (2011), “Drivers of land abandonment in Southern Chile and implications for landscape planning,” Landscape and Urban Planning 99(3-4), pp. 207-217.

[17] Li, M., J. Im, and C. Beier. (2013), “Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington Wildlife Forest,” GIScience & Remote Sensing, 50(4), pp. 361-384.

[18] Tayyebi, A., B. C. Pijanowski, M. Linderman, and C. Gratton. (2014a), “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, pp. 202-221.

[19] Pijanowski, B. C., A. Tayyebi, J. Doucette, B. K. Pekin, D. Braun, and J. Plourde. (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, pp. 250-268.

[20] Clarke, K.C., S. Hoppen, and L. Gaydos. (1997), “A self-modifying cellular automation model of historical urbanization in the San Francisco Bay area,” Environment and Planning B: Planning and Design, 24 (2), pp. 247–261.

[21] Shan, Jie., S. Alkheder, and J. Wang. (2008), “Genetic algorithms for the calibration of cellular automata urban growth modeling,” Photogrammetric Engineering & Remote Sensing, 10, pp. 1267-1277.

[22] Pontius Jr, R.G., and K. Batchu. (2003), “Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in India,” Transactions in GIS, 7 (4), pp. 467-484.

[23] Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. (1984), “Classification and Regression Trees,” Wadsworth, Belmont, CA.

[24] Li H. and J. Jie Sun. (2009), “Predicting business failure using multiple case-based reasoning combined with support vector machine,” Expert Systems with Applications, 36 (6), pp. 10085–10096.

[25] Ture, M., I. Kurt, A.T. Kurum, and K. Ozdamar. (2005), “Comparing classification techniques for predicting essential hypertension,” Expert Systems with Applications, 29 (3), pp. 583-588.

[26] Lambin, E.F. and H.J. Geist. (Eds.) (2008), “Land-use and Land-Cover Change: Local Processes and Global Impacts,” Springer, Berlin.

[27] Tayyebi, A. and B.C. Pijanowski. (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, pp. 102-116.

[28] Lumley, T., P. Sutherland, A. Rossini, N. Lewin-Koh, D. Cook, and Z. Cox. (2002), “Visualizing high-dimensional data in time and space: ideas from the Orca project,” Chemometrics and Intelligent Laboratory Systems, 60 (1-2), pp. 189-95.

[29] Gobim, A., P. Campling, and J. Feyen. (2002), “Logistic modeling to derive agricultural land determinants: a case study from southeastern Nigeria,” Agriculture, Ecosystem and Environment, 89 (3), pp. 213-228.

[30] Millington, J.D.A., G.L.W. Perry, and R. Romero-Calcerrada. (2007), “Regression techniques for examining land use/cover change: a case study of a Mediterranean landscape,” Ecosystems, 10 (4), pp. 562–578.

[31] Zhao, Z. (2008), “Parametric and nonparametric models and methods in financial econometrics,” Statistics Surveys, 2, pp. 1-42.

[32] De'Ath, G. and K.E. Fabricius. (2000), “Classification and regression trees: a powerful yet simple technique for ecological data analysis,” Ecology, 81, pp. 3178–3192.

[33] Hardle, W., M. Müller, S. Sperlich, and A. Werwatz. (2004), “Nonparametric and Semi parametric Models,” Springer Series in Statistics, Berlin.

[34] Stanton, R. (1997), “A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk,” Journal of Finance, 52 (5), pp. 1973-2002.

[35] Timofeev, R. (2004), “Classification and Regression Trees (CART) Theory and Applications.” Master Thesis., Humboldt University Berlin.

[36] Verburg, P.H., Soepbaoer, W., Veldkamp, A., Limpiada, R., Espaldon, V. and Mastura, S. (2002). “Modeling the spatial dynamics of regional land use: the CLUE S model.” Environmental Management, Vol. 30, pp. 391–405.

[37] 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, Vol. 16, No. 4, pp. 323–343.

[38] Ballestores Jr., F., Qiu, Z., Nedorezova, B.N., Nedorezov, L.V., Ferrarini, A., Ramathilaga, A. and Ackah, M. (2012).” An integrated parcel-based land use change model using cellular automata and decision tree.” Proc. International Academy of Ecology and Environmental Sciences, Vol. 2, No. 2, pp. 53–69.

[39] Ralha, C.G., Abreu, C.G., Coelho, C.G., Zaghetto, A., Macchiavello, B. and Machado, R.B. (2013).”A multi-agent model system for land-use change simulation.” Environmental Modeling & Software, Vol. 42, pp. 30–46.

[40] Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36(1), 27-46.

[41] Peneva-Reed, E. (2014), “Understanding land-cover change dynamics of a mangrove ecosystem at the village level in Krabi Province, Thailand, using Landsat data,” GIScience & Remote Sensing, 51(4), pp. 403-426.

[42] MacLean, M. G., and R. G. Congalton. (2011), “Investigating issues in map accuracy when using an object-based approach to map benthic habitats,” GIScience & Remote Sensing, 48(4), pp. 457-477.

[43] Dodd, L.E. and M.S. Pepe. (2003), “Partial AUC estimation and regression,” Biometrics, 59 (3), pp. 614–623.

[44] Pontius Jr, R. G., and B. Parmentier. (2014), “Recommendations for using the relative operating characteristic (ROC),” Landscape Ecology, 29(3), pp. 367-382.

[45] Pontius Jr, R. G., and K. Si. (2014), “The total operating characteristic to measure diagnostic ability for multiple thresholds,” International Journal of Geographical Information Science, 28(3), pp. 570-583.

[46] Olson, D., and S. Yong. (2007), “Introduction to Business Data Mining,” McGraw Hill International Edition.

[47] Tavakoli-Kashani, A., and A. Shariat-Mohaymany. (2011), “Analysis of the traffic injury severity on two lane, two-way rural roads based on classification tree models,” Safety Science, 49(10), pp. 1314–1320.

[48] Mitchell, T.M. (1997), “Machine Learning.” Redmond, Washington, Basic Books.

[49] 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), pp. 553–575.

[50] Pijanowski, B.C., Tayyebi, A., Delavar, M.R., Yazdanpanah, M.J., (2009), “Urban expan-sion simulation using geographic information systems and artificial neuralnetworks” International Journal of Environmental Research, 3 (4), pp. 493–502.

[51] Tayyebi, A., P. C. Perry, and A. H. Tayyebi. (2014b), “Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster–vector routines with remote sensing and GIS,” International Journal of Geographical Information Science, 28(4), pp. 639-659.

[52] Bishop, C.M. (1995), “Neural Networks for Pattern Recognition,” Clarendon Press, Oxford.

[53] Tayyebi, A., B. K. Pekin, B. C. Pijanowski, J. D. Plourde, J. S. Doucette, and D. Braun. (2013), “Hierarchical modeling of urban growth across the conterminous USA: developing meso-scale quantity drivers for the Land Transformation Model,” Journal of Land Use Science, 8(4), pp. 422-442.

[54] Fielding, A.H. and J.F. Bell. (1997), “A review of methods for the assessment of prediction errors in conservation presence/absence models,” Environmental Conservation, 24 (1), pp. 38–49.

[55] Cook, NR. (2007), “Use and misuse of the receiver operating characteristics curve in risk prediction,” Circulation, 115 (7), pp. 928–935.

[56] Wang, J. and G. Mountrakis. (2011), “Developing a multi-network urbanization model: A case study of urban growth in Denver, Colorado,” International Journal of Geographical Information Science, 25 (2), pp. 229-253.

[57] Feng, Y., and Y. Liu. (2013), “A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing,” International Journal of Geographical Information Science, 27(3), pp. 449–466.

[58] Kolmogoro, A.N., 1956. On the representation of continuous functions of several variables as superpositions of functions of smaller number of variables, in: Soviet. Math. Dokl. pp. 179–182.

[59] Chung, C. J. F., and A. G. Fabbri. (2003), “Validation of spatial prediction models for landslide hazard mapping,” Natural Hazards, 30(3), pp. 451-472.