A comparative study of classification and regression tree, multivariate adaptive regression spline and random forest models to simulate land use changes: A case study of Shirgah in Iran

Mohammad Ahmadlou*[1], Mahmoud Reza Delavar[2], Anahid Basiri[3], Mohammad Karimi[4]

[1] GIS Department, Geodesy and Geomatics Faculty, K.N. Toosi University of Technology, Tehran, Iran, *E-mail:* mahmadlou@email.kntu.ac.ir

[2] *Center of Excellence in Geomatics Engineering in Disaster Management, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran, Email:* mdelavar@ut.ac.ir

[3] *Centre for Advanced Spatial Analysis, University College London, London, UK, E-mail:* a.basiri@ucl.ac.uk

[4] *GIS Department, Geodesy and Geomatics Faculty, K.N. Toosi University of Technology, Tehran, Iran, E-mail:* mkarimi@kntu.ac.ir

Abstract

Design and development of a practical land use change (LUC) model require both of a high prediction accuracy, to predict the future changes, and a well-fitted model reflecting and monitoring real-world. In this regard, many models follow the three phases of training, testing and validating in the modelling process to maximise both the accuracy and fitness. Therefore, the choice of model for different application is still a valid and important question. This paper applies and compares three widely-used data mining models of Classification And Regression Tree (CART), Multivariate Adaptive Regression Spline (MARS), and Random Forest (RF) to simulate urban LUCs of Shirgah in Iran. The results of these three phases for the three models of CART, MARS, and RF for the study area of Shirgah, in the North of Iran, verify that having the highest accuracy in the testing run does not necessarily guarantee the highest accuracy in the validating run. And so, with respect to the purpose of each project, such as modelling the current situation or predicting the future, the best model with the highest accuracy at the relevant phase or a combination of some/all should be selected. For example, in this study, MARS can provide with the best accuracy in validation run while with the lowest level of accuracy in the testing run. RF provides with the highest accuracy in testing run and the lowest level of accuracy in the validation run.

Key Words: Land Use Changes, Classification and Regression Tree, Multivariate Adaptive Regression Spline, Random Forest, Total Operating Characteristics.

* Corresponding Author

1. Introduction

In many parts of the world, the rapid population growth is the major cause of the land use changes (LUCs) (Marshall et al. 2005). This is basically due to the fact that the larger population can have a broader range of needs and demands that require some new infrastructures (Xu et al. 2007). In the recent decades, the population of urban areas has increased due to the larger net migration towards the cities. It is expected that urban areas host 80% of the whole population by 2030 (Han et al. 2009). Growing towards the existing farmlands and several other land use-land covers (Dai et al. 2001), especially in developing countries, can have several environmental, economic and societal impacts. LUCs can influence the future habitats (Jantz et al. 2015), the soil properties (Biro et al. 2013; Mengistu and Waktola 2016), water resources (Tong et al. 2012), biodiversity and ecosystem services (De Baan et al. 2013; Koellner and Geyer 2013) and hydrology (Cuo et al. 2013). This shows how important is to carefully and continuously monitor, model and predict the LUCs.

Several modelling techniques have been applied to model land use-land cover change (LUCC). Some of them are essentially statistical models such as logistic regression (Hu and Lo 2007). Some other are based on the machine learning techniques (e.g. cellular automata (Berberoğlu et al. 2016; Wahyudi and Liu 2016), artificial neural networks (Pijanowski et al. 2002), genetic algorithms (Shan et al. 2008)), data mining models (e.g. classification and regression tree (Ahmadlou et al. 2016b; Tayyebi et al. 2014), random forest (Ahmadlou et al. 2016a; Kamusoko and Gamba 2015), multivariate adaptive regression spline (Ahmadlou and Delavar 2015; Ahmadlou et al. 2015; Tayyebi et al. 2014)) and agent-based models (Matthews et al. 2007). Also, there are some hybrid models integrating two or more methods, e.g. cellular automata and support vector machine (Yang et al. 2008), genetic algorithms and support vector machine (Shafizadeh-Moghadam et al. 2017) and cellular automata and genetic algorithms (Shan et al. 2008), to have a better fitness for the purpose. Three phases of training, testing and validating are a good practice for a modelling process to followed optimising accuracy and re-usability while giving a fair estimation of performance and scalability.

The model development can have the following phases: (a) the training run (calibration), (b) the testing run, and (c) the validating run. The calibration of the LUC models adjusts the input parameters in order to guarantee the best fitness. LUC drivers with a specified target variable (i.e. land use classes) are used for this purpose. The training run of the calibration run works with a subset of the entire dataset between time t_1 and t_2 which is randomly selected to train the model. The testing run of the calibration run uses the model produced by the training run to simulate the change from t_1 to another subsequent point at t_2. Then the simulated map, i.e. map at t_2, is compared to a reference map, as ground truth. The validation phase may give a better estimation of the potential achievable accuracy from the same model applied to a new subject domain (Rykiel Jr 1996). The validating run applies the model, which have been trained and tested, on a new dataset corresponding another time (i.e. t_3) to predict the LUC between t_2 and t_3. Then the predicted results at t_3 are compared to a reference map of t_3 to assess the prediction ability of the model (Pontius et al. 2008).

One of the most important aims of the LUC modelling is to accurately predict the changes. It is thought that each model can have different level of accuracy at each phase of modelling process (van Vliet et al. 2016). For example, while one model may be have the highest accuracy in the testing run, yet another model may provide with the highest accuracy level in the validating run. Given the variety and the wide spectrum of available LUC models, it is important to select and use the most suitable model and for making such decisions having a better understanding of the performance, accuracy, and also other strengths and limitations can be very useful (Pontius et al. 2008). Several research projects compared a pair of models as the pairwise comparison could be easier to perform, particularly for a large dataset. In addition, most of the studies considered the accuracy of testing runs and did not consider prediction ability of the models, which indicate the re-usability of the trained and tested models (Cheng and Masser 2003; Hu and Lo 2007; Tayyebi et al. 2014). For example, Tayyebi and Pijanoviski, (2014) compared CART, MARS and artificial neural network (ANN) in three different regions, they just considered training and testing run i.e. a single interval, and did not consider prediction ability of the three models. Thus, the main objectives of this research are (a) to implement several of the most widely used models, including CART, MARS and RF, and (b) to compare their capabilities in testing and validating phases in LUCs modelling.

This paper is organized as follows. Section 2 explains the study area, data used to apply our model to simulate of LUC and the methods including MARS, CART and RF. We present the core results including the error maps and performance measure and discuss the findings in Section 3. Finally, in section 4 the conclusions are presented.

2. Study area

The study area is Shirghah, in the north of Iran, located at 35° 56´ N and 52° 57´ E (see figure 1). Shirgah dominantly consist of agricultural lands and so has faced some significant changes in the land use and land cover. This area is particularly important in terms of agriculture and sustainable development, and so the LUCC are needed to be well studied.

Figure 1. Study area is Shirgah zone (right panel) located in North of Iran (left panel)

2.1. Dataset and methodology

In this study, most of the required information was captured from the Landsat satellite images. The road network is obtained from OpenStreetMap (OSM) database, which is publicly and freely available source of data contributed by the crowd. Landsat images in 1991 (TM, May 1991), 2001 (ETM+, May 2001) and 2011 (ETM+, April 2011) were processed and projected to UTM Zone 39 North with 30m spatial resolution. The selection of the study years depends on the availability of data, the reduction of sessional effects, the regular time intervals between the images, and the possibility of obtaining the cloud-free images.

The classification procedure was performed through the maximum likelihood classification followed by a post-classification phase, which can improve the accuracy of the classified maps. The land uses are classified into four classes of built-up areas, agricultural areas, water bodies, and forest. The classification accuracies using the Kappa index (Cohen 1968) are 87%, 86% and 88% for 1991, 2001 and 2011, respectively. The driving forces of urban changes and their related information are shown in Table 1. The LUC maps can be generated by comparing the two sequent land use maps. Next, a set of 10 driving forces (Table 1) is considered. Figure 2 represent the implemented methodology for this research.

Table 1. Spatial explanatory variables of urban change between 1991 and 2001.

Figure 2: The proposed methodology workflow

2.2. Classification and regression tree

CART is a rule-based data mining technique that can handle both classification and regression tasks (L Breiman et al. 1984). The learning process of the CART model consists of two stages of selecting the tree structure and determining the predictions at the leaf nodes (i.e. nodes without children). CART does not need to have any

Assumption such as normal distribution for the relationship between the dependent and independent variables. CART recursively divides the input data with respect to the independent variables that introduce the highest purity, where the leaf nodes are formed from the same land use class. Among all existing variables, the variable is selected that can increase the node purity. There are several criteria for data partitioning at each node, among which, the Gini index can handle nominal values (Olson et al. 2007). The Gini index at node *t* is determined using Eq.1 (L Breiman et al. 1984):

$$Gini\ (t) = \sum_{i \neq j} P(w_i) \times P(w_j) \qquad (1)$$

Where $P(w_i)$ is the relative frequency of class *i-th*. The process of tree growth continues until the highest purity at the leaf nodes is achieved. If the decision tree performs the modelling using a target variable with nominal values, it is called a classification tree, and if it performs the modelling using a target variable with continuous values, it is called a regression tree (L Breiman et al. 1984).

2.3. Multivariate adaptive regression spline

MARS is a data mining model that splits data into several partitions and formulates the relationship between the target variables, here the land use change classes, and the explanatory variables (Fridedman 1991). This paper builds up this relationship using the piecewise polynomial functions, which are called *basis functions*. While the non-linear models can only fit one set of coefficients to the data, MARS can fit separate piecewise polynomial functions to each region and so MARS can generate a distinct set of coefficients (Fridedman 1991). Eq. (2) shows the general equation of MARS (Fridedman 1991):

$$\hat{Y} = \hat{f}(x) = \sum_{m=1}^{M} \alpha_m B_m(x) + e \qquad (2)$$

where m, M, α, x, Y, and e are, respectively, the number of spatial drivers of LUC, the number of sub-regions, the basis function coefficients, spatial drivers of LUC, land use classes which will be modelled, and the error term. B is the basis function which, itself, can be represented as (Fridedman 1991):

$$B_m(x) = \prod_{i=1}^{N_m} [S_{i,m}\ (X_{v(i,m)} - t_{i,m})]_+^q \qquad (3)$$

where N, $S_{i,m}$, $X_{v(i,m)}$, $t_{i,m}$ and q are, respectively, the interaction order of the m-th basis function, ±1, the *v*-th variable where $1 \leq v\ (i, m) \leq k$ (where k is the total number of spatial drivers), a knot location of the spatial drivers and the power of the basis function. N can be specified by the users as prior domain knowledge. When q = 1, the simple linear splines are determined. The subscript '+' is following phrase (Fridedman 1991):

$$[S_{i,m}\ (X_{v(i,m)} - t_{i,m})]_+^q = \begin{cases} [S_{i,m}\ (X_{v(i,m)} - t_{i,m})]^q & S_{i,m}\ (X_{v(i,m)} - t_{i,m}) \geq 0 \\ 0 & otherwise \end{cases} \tag{4}$$

When dispensable basis functions are added to the equation, there is a possibility to over-fit. To avoid the over-fitting, the basis functions with the least influences are excluded using generalised cross validation (GCV) as shown in Eq. 5 (Friedman and Silverman 1989). The aim of MARS is minimising the GCV.

$$GCV = \frac{\left(\frac{1}{n}\right) \sum_{i=1}^{n} [y_i - f(x_i)]^2}{[1 - (C(m)/n)]^2} \tag{5}$$

In Eq. (5) *n*, *y* and *f* are the number of total observations, the response variable, and the estimated function by MARS, respectively. C is defined as follow (Fridedman 1991):

$$C(M) = M\ \times d \tag{6}$$

where d and M are the cost for each basic function and total number of basic functions, respectively.

2.4. Random Forest

RF was first introduced by Breiman (2001) as a model with a larger number of CART trees. In RF, each tree is based on a random subset of the observations and the split within each tree is made based on a random subset of candidate variables (Leo Breiman 2001). For each observation, all the individual trees vote for one class and the overall prediction of the RF is the average of predictions from each of the trees (Leo Breiman 2001). Trees in the RF can use the Gini index as the splitting criterion. The root node, i.e. node at the top of each tree, in the forest contains a bootstrap sample from the total data as the training dataset. The observations that are not in the training set, roughly one third of the total data set, are referred as the out-of-bag (OOB) observations (Leo Breiman 2001). OOB is used to assess the model performance.

2.5. Model validation

Model validation is a performance assessment process that compares the predicted and simulated values (here maps). In this study, three Landsat images in two time intervals were used for the training, testing and validating phases. Data corresponding to the changes between1991 and 2001 were used for training and testing run and the data corresponding to the changes between 2001 and 2011 were used for the validation run of the models. Many calibration metrics such as kappa (Cohen 1968), Figure of Merit (Pontius et al. 2008) and percent correct match (R. Pontius Jr and Schneider 2001) have been used for accuracy assessment of the land use change models. One of the most widely used is relative operating characteristic (ROC), which uses a series of thresholds to convert the probability map to simulated map (R Gil Pontius Jr and Batchu 2003). Pontius Jr and Si (2014) recently presented the total operating characteristic (TOC) to rectify the limitations of ROC. For example, ROC fails to reveal the size of each entry in the contingency table for each threshold (Robert Gilmore Pontius Jr and Si 2014). This paper uses TOC to compare the accuracy of CART, MARS and RF models.

3. Results and Discussion

The result of the LUCs is a binary map, either unchanged or changed between 1991 and 2001. The changed areas (converted to built-up area) are coded with 1 and the others are coded with 0. The LUC variable, i.e. the target variable, and a set of ten LUC drivers including distance from the nearest main road, the shortest distance to the nearest built-up area, distance from the nearest agricultural land, distance from the nearest forest, slope, elevation, aspect, northing and easting were established as the explanatory variables. In addition, the built-up areas in 1991 and water bodies are excluded from the modelling process as exclusionary zone. In the phase of model calibration, 60% of the cells are used for training and 40% are reserved for validation. The whole process of modelling and validation was programmed in the MATLAB software. The results of each model are discussed in detail in the following sub-sections.

3.1. MARS Modelling

Having run MARS model, it starts converging after 15 basis functions shown and the error rate falls below 0.019 (Fig.3). The regression equation is created using these 15 basic functions and begins with a constant, which is the average of the target variable (i.e. land use classes) and the basis functions are gradually added to this term. The equation has 15 basis functions with their coefficients. Each includes one explanatory variable and is split at one knot. For example, the first basis function considers distance to urban variable, which is split in 108.167 meters. The coefficient of the first basic function is -0.0000032. In this basis function, for each cell, the maximum of zero and the difference between the distance to the urban variable and 108.176 meters are calculated and multiplied by 0.0000032.

Figure 3: GCV across adding radial basis functions to MARS

3.2. CART Modelling

CART model includes eleven nodes contributing. Fig. 4 shows the relative cost of the training phase, which measures the misclassification error with respect to the tree size. This plot starts around 0.55 in tree with two nodes and then decreases dramatically in tree with six nodes. The most accurate tree is shown by the green bar marking in Fig. 4.

The best tree in this process has a set of eleven leaf nodes, which reached a relative cost of 0.426, see Fig. 4. Having eleven leaves means that CART model has got eleven rules in the tree structure, see Fig. 5. To construct each of the rules that are associated with each of the leaf nodes, the users should start from the root node and reach to the leaf nodes through the internal nodes.

Figure 4. Relative cost of the training run

Figure 5. Tree navigator of CART

3.3. RF Modelling

RF model calculates the output by taking the average of the total number of spatial divers of the trees. As a best practice, in order to make decisions at each tree node, only the square root (m) of the total number of predictors (indicated by *p*) is used. Since there are ten independent variables to predict the target variable, the m parameter is set to be four.

The number of trees to be fitted is 400. To assess the model performance, OOB approach is used, see Fig. 6.

Figure 6. Decrease in error as a function of number of trees for RF model

3.4. Suitability, simulated and error maps

LUC models can also produce a by-product called suitability map, i.e. a map showing the suitability of change for each cell which varies from low to high. Then, A simulated map, i.e. a map showing the membership value of each cell to either changed or unchanged classes can be create. Cells in suitability map have values varying from 0 to 1 where values closer to 0 are more likely for no-change and values closer to 1 are more likely for change, While cells in simulated map have either 0 (unchanged cells) or 1 (changed cells). Error maps are created by overlaying the simulated maps and the observed change maps of 2001 and 2011. Fig. 7 shows the error maps obtained from this process using CART, MARS and RF models. True Positive (TP) or Hits shows the cells which are correctly predicted as change verified by the observed map. False Negative (FN) or False Alarms shows the cells which are predicted as change while are actually non-changed cells according to the observed map. False Positive (FP) or Misses shows the cells which are predicted to remain with no changes, while they are actually changed cells in the observed map. And finally, True Negative (TN) or Correct Rejections shows the cells which are correctly predicted to have no changes, verified by the observed map.

3.5. Performance of the MARS, CART and RF in the testing and validating run

Having trained the models, the remaining 40% of the data in the first interval, as well as land use change for 2011, are used to be predicted (testing run and validating run). Six suitability maps are created for the three models (i.e. CART, MARS and RF for 2001 and 2011). Each threshold in TOC curve creates a two-by-two contingency table, which has four numbers including TP (Hits), FN (False Alarms), FP (Misses) and TN (Correct Rejections). 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 (Fig. 8). From TOC curve, all of the parameters for each threshold can be found, see Fig. 8. The ratio of the areas under the TOC curve which is within TOC’s bounding parallelogram to the area of TOC’s bounding parallelogram are 94.87%, 84.37% and 98.99% for testing run in 2001 for CART, MARS and RF models, respectively (Fig. 7a,7c,7e). CART, MARS and RF models find the ratios, for validating run in 2011, 76.55%, 78.16% and 73.60%, respectively.

The results indicate that RF is a the most successful model for simulating urbanisation changes in testing run while MARS has the least accuracy in this phase. In the validating run, however, the highest and the lowest levels of accuracy belong to MARS and RF, respectively. This shows that having the highest accuracy in the testing run does not guarantee having the highest accuracy in validating run.

Figure 7. The error map obtained by overlaying the reference maps for the years 2001 and 2011 with the map for 2001 and 2001 predicted by the CART, MARS and RF models

Figure 8. TOC curves for assessing the performance of CART (2001 (a) and 2011 (b)), MARS (2001 (c) and 2011 (d)) and RF (2001 (e) and 2011 (f)) for testing and validating runs.

4. Conclusion

The prediction accuracy is one of the most challenging aspects of developing LUCs models. While one model may provide the highest accuracy in the training phase, it may provide an unacceptable level of accuracy in testing phase. Therefore, finding the best balance and also having a better understanding of the accuracy of each phase would help better design and development of the LUC models with respect to the purpose of use, e.g. prediction or monitoring. Review of research background indicated that comparative studies on prediction accuracy of the LUC models can be enhanced by investigating how they can be applied to different scenarios of LUC. This paper compared three widely used data mining models of CART, MARS and RF in two phases of testing and validation. The results showed that having the highest accuracy in the testing run does not guarantee having the highest accuracy in the validating run. This highlights the importance of considering the two phases of the testing and validating runs in developing and evaluating LUCs models.

5. References

Ahmadlou, M., & Delavar, M. (2015). Multiple Land Use Change Modeling Using Multivariate Adaptive Regression Spline and Geospatial Information System. *Journal of Geomatics Science and Technology, 5*(2), 131-146.

Ahmadlou, M., Delavar, M., Shafizadeh-Moghadam, H., & Tayyebid, A. (2016a). Modeling urban dynamics using random forest: Implementing roc and toc for model evaluation. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, 285-290.

Ahmadlou, M., Delavar, M., & Tayyebi, A. (2016b). Comparing ANN and CART to model multiple land use changes: A case study of Sari and Ghaem-Shahr cities in Iran. *Journal of Geomatics Science and Technology, 6*(1), 292-303.

Ahmadlou, M., Delavar, M., Tayyebi, A., & Shafizadeh-Moghadam, H. (2015). Using Multivariate Adaptive Regression Spline and Artificial Neural Network to Simulate Urbanization in Mumbai, India. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40*(1), 31.

Berberoğlu, S., Akın, A., & Clarke, K. C. (2016). Cellular automata modeling approaches to forecast urban growth for adana, Turkey: A comparative approach. *Landscape and urban planning, 153*, 11-27.

Biro, K., Pradhan, B., Buchroithner, M., & Makeschin, F. (2013). Land use/land cover change analysis and its impact on soil properties in the northern part of Gadarif region, Sudan. *Land Degradation & Development, 24*(1), 90-102.

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

Breiman, L., Friedman, J., Stone, C., & Olshen, R. (1984). Classification Algorithms and Regression Trees. CRC press.

Cheng, J., & Masser, I. (2003). Urban growth pattern modeling: a case study of Wuhan city, PR China. *Landscape and urban planning, 62*(4), 199-217.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin, 70*(4), 213.

Cuo, L., Zhang, Y., Gao, Y., Hao, Z., & Cairang, L. (2013). The impacts of climate change and land cover/use transition on the hydrology in the upper Yellow River Basin, China. *Journal of hydrology, 502*, 37-52.

Dai, F., Lee, C., & Zhang, X. (2001). GIS-based geo-environmental evaluation for urban land-use planning: a case study. *Engineering Geology, 61*(4), 257-271.

De Baan, L., Alkemade, R., & Koellner, T. (2013). Land use impacts on biodiversity in LCA: a global approach. *The International Journal of Life Cycle Assessment, 18*(6), 1216-1230.

Fridedman, J. (1991). Multivariate adaptive regression splines (with discussion). *Ann. Statist, 19*(1), 79-141.

Friedman, J. H., & Silverman, B. W. (1989). Flexible parsimonious smoothing and additive modeling. *Technometrics, 31*(1), 3-21.

Han, J., Hayashi, Y., Cao, X., & Imura, H. (2009). Application of an integrated system dynamics and cellular automata model for urban growth assessment: A case study of Shanghai, China. *Landscape and urban planning, 91*(3), 133-141.

Hu, Z., & Lo, C. (2007). Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems, 31*(6), 667-688.

Jantz, S. M., Barker, B., Brooks, T. M., Chini, L. P., Huang, Q., Moore, R. M., et al. (2015). Future habitat loss and extinctions driven by land-use change in biodiversity hotspots under four scenarios of climate-change mitigation. *Conservation Biology, 29*(4), 1122-1131.

Kamusoko, C., & Gamba, J. (2015). Simulating urban growth using a Random Forest-Cellular Automata (RF-CA) model. *ISPRS International Journal of Geo-Information, 4*(2), 447-470.

Koellner, T., & Geyer, R. (2013). Global land use impact assessment on biodiversity and ecosystem services in LCA. *The International Journal of Life Cycle Assessment, 18*(6), 1185-1187.

Marshall, J. D., McKone, T. E., Deakin, E., & Nazaroff, W. W. (2005). Inhalation of motor vehicle emissions: effects of urban population and land area. *Atmospheric Environment, 39*(2), 283-295.

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. (2007). Agent-based land-use models: a review of applications. *Landscape Ecology, 22*(10), 1447-1459.

Mengistu, D. A., & Waktola, D. K. (2016). Monitoring land use/land cover change impacts on soils in data scarce environments: a case of south-central Ethiopia. *Journal of Land Use Science, 11*(1), 96-112.

Olson, D. L., Shi, Y., & Shi, Y. (2007). *Introduction to business data mining* (Vol. 10): McGraw-Hill/Irwin Englewood Cliffs.

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

Pontius Jr, R., & Schneider, L. (2001). Land-use change model validation by a ROC (relative operating characteristic) method. *Agric. Ecosyst. Environ, 85*, 239-248.

Pontius Jr, R. G., & Batchu, K. (2003). Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in India. *Transactions in GIS, 7*(4), 467-484.

Pontius Jr, R. G., & Si, K. (2014). The total operating characteristic to measure diagnostic ability for multiple thresholds. *International Journal of Geographical Information Science, 28*(3), 570-583.

Pontius, R. G., Boersma, W., Castella, J.-C., Clarke, K., de Nijs, T., Dietzel, C., 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.

Rykiel Jr, E. J. (1996). Testing ecological models: the meaning of validation. *Ecological modelling, 90*(3), 229-244.

Shafizadeh-Moghadam, H., Tayyebi, A., Ahmadlou, M., Delavar, M. R., & Hasanlou, M. (2017). Integration of genetic algorithm and multiple kernel support vector regression for modeling urban growth. *Computers, Environment and Urban Systems, 65*, 28-40.

Shan, J., Alkheder, S., & Wang, J. (2008). Genetic algorithms for the calibration of cellular automata urban growth modeling. *Photogrammetric Engineering & Remote Sensing, 74*(10), 1267-1277.

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

Tong, S. T., Sun, Y., Ranatunga, T., He, J., & Yang, Y. J. (2012). Predicting plausible impacts of sets of climate and land use change scenarios on water resources. *Applied Geography, 32*(2), 477-489.

van Vliet, J., Bregt, A. K., Brown, D. G., van Delden, H., Heckbert, S., & Verburg, P. H. (2016). A review of current calibration and validation practices in land-change modeling. *Environmental Modelling & Software, 82*, 174-182.

Wahyudi, A., & Liu, Y. (2016). Cellular automata for urban growth modelling. *International Review for Spatial Planning and Sustainable Development, 4*(2), 60-75.

Xu, C., Liu, M., An, S., Chen, J., & Yan, P. (2007). Assessing the impact of urbanization on regional net primary productivity in Jiangyin County, China. *Journal of environmental management, 85*(3), 597-606.

Yang, Q., Li, X., & Shi, X. (2008). Cellular automata for simulating land use changes based on support vector machines. *Computers & geosciences, 34*(6), 592-602.