URBAN GROWTH MODELING USING GENETIC ALGORITHMS AND CELLULAR AUTOMATA; A CASE STUDY OF ISFAHAN METROPOLITAN AREA, IRAN

Ehsan, FOROUTAN[1], Mahmoud Reza, DELAVAR[2]

[1]Dept of Surveying and Geomatics Eng. College of Eng. University of Tehran, North Kargar St, 11155-4563, Tehran, Iran

eforoutan@ut.ac.ir

[2]Center of Excellence in Geomatics Eng. and Disaster Management, Dept of Surveying and Geomatics Eng. College of Eng. University of Tehran, North Kargar St, 11155-4563, Tehran, Iran

mdelavar@ut.ac.ir

Abstract

This study integrates cellular automata (CA) and genetic algorithms (GAs) to model urban growth in the Isfahan Metropolitan Area in Iran. The simulation of urban growth through cellular automata models brings improved understanding of the complex dynamic process of land use change, which can not be achieved through conventional models. The cellular automata (CA) as a powerful spatial dynamic modeling tool are designed as a function of parameters whose calibration plays a crucial role for obtaining a suitable set of parameters in order for precise and reliable modeling. Genetic Algorithms are useful tools for decreasing the search space for finding the optimal solution of transition rules in cellular automata and reducing the simulation uncertainties and improving its locational accuracy in urban modeling. The considered objective function in this algorithm is percent correct match (PCM) obtained from error matrix between the simulated and the reference map. Historical land use/cover data of Isfahan Metropolitan Area were extracted from the 1990 and 2001 Landsat ETM$^+$ images at 30m spatial resolution. Three different Moore neighborhood sizes have been considered for cellular automata model and simulation of urban growth for the year 2001 is performed. The simulation outcomes, evaluated with kappa statistic of 74.15% demonstrate that the integration of GA and CA could be suitable for dynamic urban growth modeling.

KEYWORDS: Cellular Automata, Genetic Algorithms, Calibration, Neighborhood Size

INTRODUCTION

Developing methods for assessing different urban growth planning scenarios regarding the future consequences for land use and the progress of current spatial plans and policies is critically important for urban and regional planners(Al-Ahmadi et al., 2009). An urban land use system is dominated by human activities with complex spatio-temporal dynamics (Hu and Lo, 2007). Many efforts have been made to model urban growth using different models and algorithms but amongst them cellular automata models have proved rather popular as frameworks for modeling and simulating the physical growth of cities. The simulation of urban growth through cellular automata models brings improved understanding of the complex dynamic process of land use change, which cannot be achieve through conventional models (Batty et al., 1999). CA was originally developed for simulating complex systems in physics and biology. CA systems were first developed in the late 1940s by S. Ulam and J. von Neumann and then at first serious research developed by (Wolfram, 1984) demonstrated that complex natural phenomena can be modeled by CA models. The first effort to employ CA for modeling urban growth refers to a pioneer work of (Tobler, 1970) who proposed the application of cellular space models to geographic modeling. Because of spatial nature of CA they could gain attention between researchers and urban planners, as a large and increasing volume of work shows that CA are proper tools for modeling spatial dynamics (e.g. (Al-Ahmadi, et al., 2009; Alkheder and Shan, 2005; Couclelis, 1985; Dragicevic, 2004; Itami, 1994; Portugali and Benenson, 1995; Wu, 1998a, 1998b; Yeh and Li, 2002)). However, even though technically there is little limitation developing CA models within a GIS environment, it remains a research issue to urban modelers to identify suitable transition rules and their defining parameters in CA based urban modeling (Feng et al., 2011). In CA, many variables are involved and each variable is usually associated with a parameter that indicates its importance in simulation (Li et al., 2007). These parameters significantly affect the outcomes of urban simulation (Yeh and Li, 2002). It is essential to define proper parameter values when CA is used to simulate realistic cities. Only through calibration the cellular automata model can produce an urban level and urban pattern close enough to reality (Shan et al., 2008). In other words, the purpose of calibration is to establish the relationship between land use change and the factors that affect probability of land conversion (Wu, 2002). Calibration in cellular automata modeling intends to find the best value combination for transition rules to match simulated and real urban phenomenon quantitatively and qualitatively (Al-Kheder, 2007). A number of cellular automata calibration methods have been developed for urban growth modeling. They achieved various levels of success and efficiency. Generally, calibration methods are divided into three categories including statistical, visual and artificial intelligence tools (Al-Kheder, 2007). (Clarke et al., 1997) calibrated the SLEUTH model by using visual and statistical tests to find the best values for the five growth parameters (slope, land use, exclusion, urban extent, transportation, and hill shade). In addition (Goldstein, 2004) uses brute force method to find the best parameter of SLEUTH model of urban growth. Such calibration methods take expanded CPU time to reach the most appropriate parameter set in the search space and may get trapped in a local maximum. In recent years, artificial intelligence techniques have attracted considerable attention research in urban modeling for many reasons. Intelligence techniques such as artificial neural networks (ANNs), fuzzy logic, and genetic algorithms (GAs) are popular tools, since they can deal with complex engineering problems which are difficult to solve by classical methods. In case of ANNs,(Yeh and Li, 2002) proposed a model which employed ANNs as calibration tool of cellular automata but unlike traditional CA models it generate implicit rules that could not be interpretable. In case of Gas, initial efforts tried to attach genetic algorithms to cellular automata urban model design for performance improvement. (Colonna et al., 1998) modeled the changes in land-use for Rome, Italy using genetic algorithms to produce a new set of rules for the cellular automata model. In addition Genetic algorithms were used to find the optimal set of possibilities of land-use planning for Provo, Utah (Balling et al., 1999). A recent study tried to formalize genetic algorithms as a calibration tool for the SLEUTH model (Goldstein, 2004). In (Shan, et al., 2008) proposed a model to enhance the efficiency of transition rule calibration in cellular automata urban growth modeling using GA for Indianapolis.

This paper applies genetic algorithm as an artificial intelligence technique for calibration of cellular automata transition rules for urban growth modeling in Isfahan Metropolitan Area of Iran from 1990 to 2001. The model validation is performed employing kappa statistic and overall accuracy to test the model ability to simulate urbanization pattern of reality.

STUDY AREA

The Isfahan Metropolitan Area in the center of Iran is considered in this study for modeling urban growth. The data that have been used for calibration and simulation included two historical satellite images covering a period of ten years. These raw images include two 28.5 m resolution ETM$^+$ images 1990 and 2001.

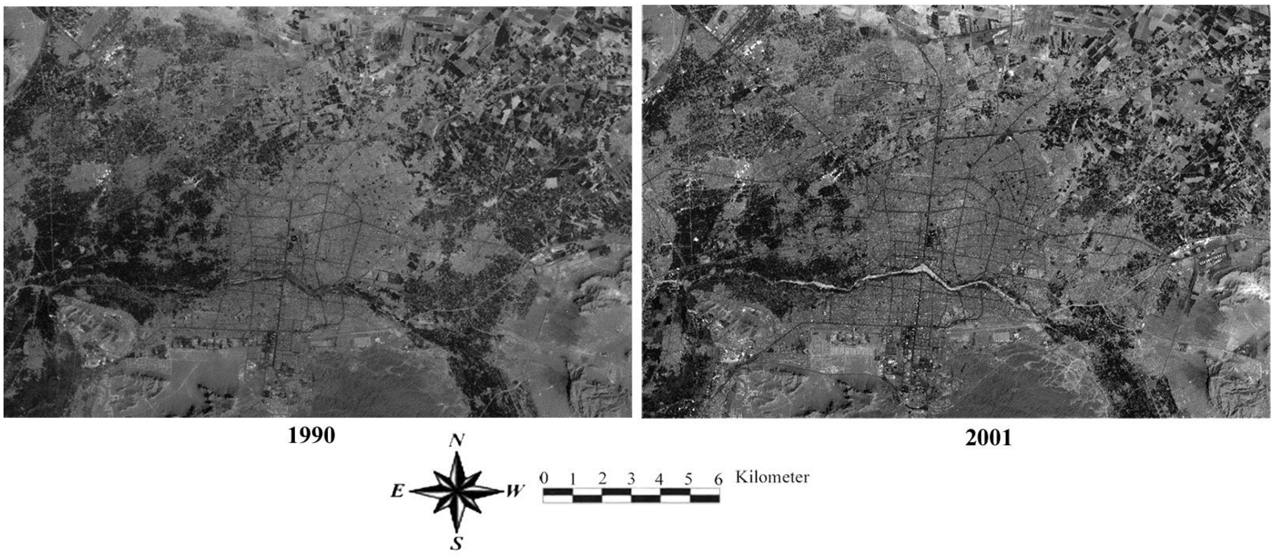


Fig. 1. Satellite images of Isfahan in 1990 and 2001, respectively.

Both images were geometrically rectified and registered to the Universal Transverse Mercator (UTM) WGS 1984 zone 39N. Registration errors were about 0.50 pixels. In addition, combinations of RGB bands of Landsat images were performed to prepare satellite imageries for better classification. Fig. 1 shows the final results for the 1990 and 2001 Landsat images of Isfahan. After rectifying and registering the images, the next step is classification of the outcomes which are the inputs to a cellular automata model. All land use classes of Isfahan were also reclassified from their original classification to Anderson Level I (Anderson et al., 1976) for the modeling exercises. Four classes are defined based on maximum likelihood classification system namely water, road, urban, vegetation area and barren. These land uses considered as effective land uses for urban growth as shown in Fig. 2.

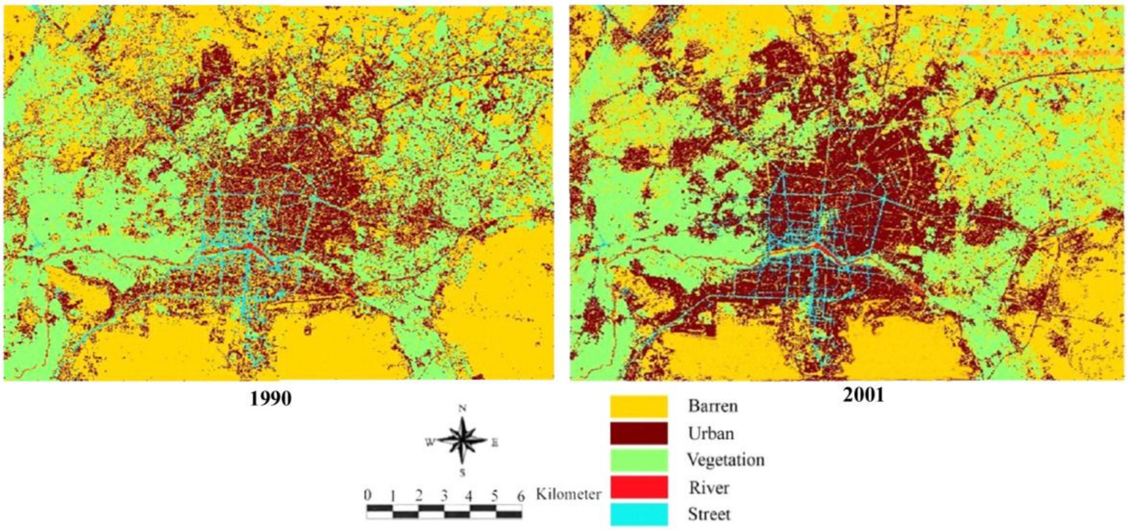


Fig. 2. Isfahan classified land use map in 1990 and 2001, respectively

METHODOLOGY

Calibration of cellular automata plays a crucial role in a precise and reliable modeling. In this section we address the structure and function of the proposed model. Land uses extracted from satellite images in previous section are: urban, road, vegetation, water and non-urban (barren) areas. Each of these land uses has its effect on urban growth that importance of each, varies from a region to another. These preferences have been considered in a set of rules. For example some people prefer to inhabit in areas near to the vegetation areas while the other prefers have an access to streets. Using cellular automata transition rules, we will be able to extract these rules and use them for modeling and forecasting urban growth. The proposed transition rules for the cellular automata urban growth model are designed using the above mentioned input to identify the minimum urbanization conditions needed in a Moore neighborhood (square) for a test pixel to become urban or not. These transition rules are summarized as follows:

1. IF the test pixel is water, road, vegetation OR urban THEN no change.

2. IF the test pixel is non-urban (barren) THEN it becomes urban if:

   - Its neighboring URBAN pixel count is ≥ No_u,

     OR

   - Its neighboring ROAD pixel count is ≥ No_s,

     OR

   - Its neighboring VEGETATION pixel count is ≥ No_v,

     OR

   - Its neighboring WATER pixel count is ≥ No_w.

where No_r, No_s, No_P and No_l are the minimum number of urban, street, vegetation and water pixels, respectively in a square neighborhood which are require for urbanization of one non-urban pixel. In this research we consider three different neighborhood sizes (3×3, 5×5 and 7×7) and finally compare the obtained results from them and select the best one for our study area. There are a total 9^4, 25^4 and 49^4 combinations of possible rule values for 3×3, 5×5 and 7×7 neighborhood sizes, respectively. These extensive search spaces entail us to use optimum search methods such as genetic algorithms for reducing the search spaces and finding the optimum solutions.

For evaluating each generated population, we have to use model validation parameters. The objective function used in this study is percent correct match (PCM) and we are going to minimize (100-PCM). The percent correct match (PCM) is a parameter that indicates the accuracy of the model using confusion matrix. This parameter assesses the following parameters: (1) the urban growth is occurred in real world and the model illustrates it, (2) the urban growth is not occurred in real world while the model has shown the growth.

The confusion matrix can summarize the results by overlaying the map of a simulated land use on the map of reality, as shown in Table 1 (Pontius and Schneider, 2001).

Table 1. Confusion Matrix

Model	Reality		
	Change	Non-change	Total
Change	A	B	A+B
Non-change	C	D	C+D
Total	A+C	B+D	A+B+C+D

The percent correct match (PCM) is calculated from (1) (Pontius and Schneider, 2001):

$$PCM = \frac{A+D}{A+B+C+D} \tag{1}$$

Genetic algorithms (GAs) were invented by John Holland in the 1960s (Holland, 1992)and were further developed by him and his colleagues at the University of Michigan in the 1960s and the 1970s (Goldberg, 1989). The general algorithm of this approach is as follows:

```
Algorithm: Genetic Algorithms
  •  Initialize population
  •  Evaluate population
  •  Do while (termination-criteria is not satisfied)
  •  Select parents for reproduction
  •  Perform recombination and mutation
  •  Evaluate
Loop
```

The first step towards applying genetic algorithms is generating initial population. Population for calibrating the proposed cellular automata model includes the number of urban, street, vegetation and water pixels in a defined square neighborhood. As an example, a rule as a member of population could be 2, 4, 5, 7 which means that a non-urban pixel will become urban if in its neighborhood there are at least 2 or 4 or 5 or 7 urban, street, vegetation or water, respectively. Proposed population size is 20 which mean that the population includes 20 of the mentioned rules that are randomly generated. Then the cellular automata model is run for each member in the population and the fitness of the population is calculated through the objective function. If each of the rules satisfies the termination criterion, it will be selected as the best answer and algorithm will be stopped; otherwise the rules with the lower objective function are given a lower rank and vice versa. For using genetic operation such as elitism, crossover and mutation, population must be sorted descending in term of its fitness. The next population will be generated according to elitism selection, crossover and mutation. Based on the proposed algorithm, 10% of the sorted population in the previous step is copied directly into the next generation before crossover and mutation. This step allows us to keep the best answers. The next step in genetic algorithms is crossover and mutation applied to population from the previous step. Crossover is the process in which two chromosomes (strings) combine their genetic materials to produce a new offspring which possesses both their characteristics. In Fig. 3 crossover operation on the two parents are shown that leads to producing two offsprings. In this paper crossover rate was 0.60 which means that crossover applies on the 60% of population.

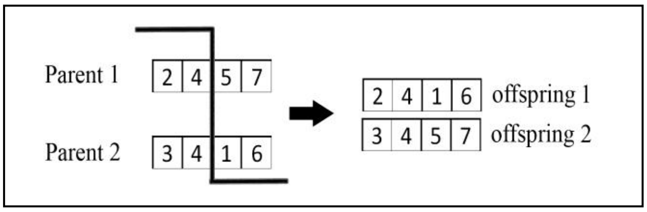


Fig. 3. Cross Over operator

The last step in finalizing a new population is mutation operation. It is the process by which a string is deliberately changed so as to maintain diversity in the population set and prevents the solution from local minima. In this study, mutation applied to the 30% of the population by generating a set of random solutions. By applying elitism selection, crossover and mutation, new population is produced and could be evaluated using the objective function. This process is repeated until termination criterion is satisfied.

RESULTS AND DISCUSSION

Table 3 shows the results acquired through applying genetic algorithms as a tool for calibrating cellular automata transition rules. Calibrated rules column in Table 3 indicate the minimum required number of urban, street, vegetation and water pixels in a defined square neighborhood, respectively.

Aside from calibration of cellular automata, validation is other issues that must be addressed. Validation procedure measure the performance of CA model to simulate a reliable model that can be reproduced urbanization pattern of reality. A well known measurement procedure is based on the use of error matrices for mapping comparisons (Norte Pinto and Pais Antunes, 2007).For evaluating the model accuracy, Kappa statistics (KS) is used. The Kappa statistics is much used to assess the similarity between the observed and predicted results. The calculation of Kappa is based on contingency table (Monserud and Leemans, 1992) (Table 2).

Table 2. The Contingency Table

	Model					Total
	Class	1	2	…	c	
Reality	1	P_{11}	P_{12}	…	P_{1c}	P_{1T}
	2	P_{21}	P_{22}	…	P_{2c}	P_{2T}
	…	…	…	…	…	…
	c	P_{c1}	P_{c2}	…	P_{cc}	P_{cT}
Total		P_{T1}	P_{T2}	…	P_{Tc}	1

On the basis of the contingency table, many statistics can be derived as follow (Hagen, 2003):

$$P(A) = \sum_{i=1}^{c} P_{ii}$$
$$P(E) = \sum_{i=1}^{c} P_{iT}.P_{Ti} \qquad (2)$$

And finally Kappa statistics could be calculated from (3) (Hagen, 2003):

$$KS = \frac{P(A) - P(E)}{1 - P(E)} \qquad (3)$$

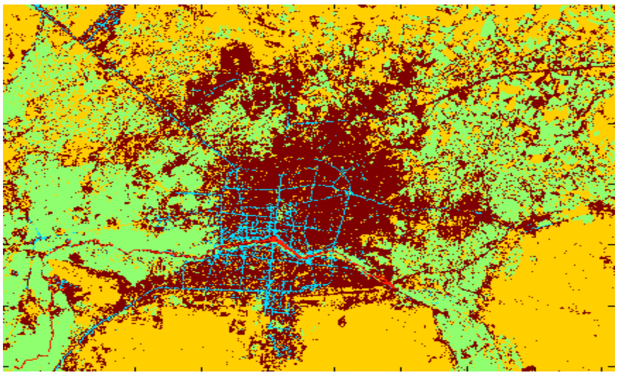
It is generally considered that Kappa values for map agreement are: >0.8 is excellent; 0.6-0.8 is very good; 0.4-0.6 is good; 0.2-0.4 is poor and <0.2 very poor (Pijanowski et al., 2005).

Table 3. Results of Cellular Automata Calibration

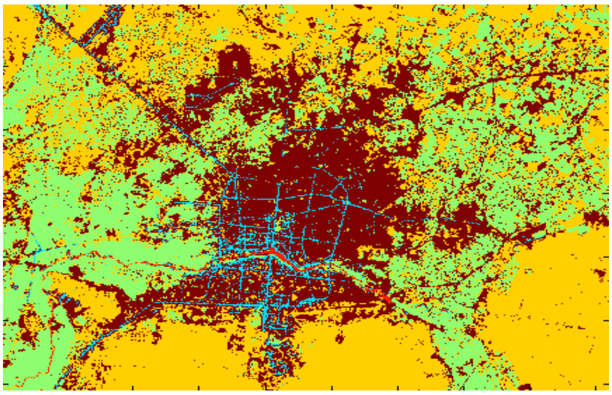
Windows Size	PCM%	KS%	Calibrated Rules
3	86.28	72.38	4 , 1 , 8 ,2
5	87.17	74.15	11 , 3 , 23 , 12
7	86.49	72.65	25 , 12 , 43 , 25

As the results show (Table 3), percent correct match (PCM) and Kappa statistics calculated from confusion matrix and contingency table, indicates that using a square neighborhood size of 5×5 i.e. in a neighborhood of 150×150m is more suitable than 3×3 or 7×7 neighborhood size. Based on the earlier work (Pijanowski, et al., 2005) Kappa values calculated for simulated land use map seems very good. The simulated land use map of Isfahan Metropolitan Area is shown in Fig. 4.

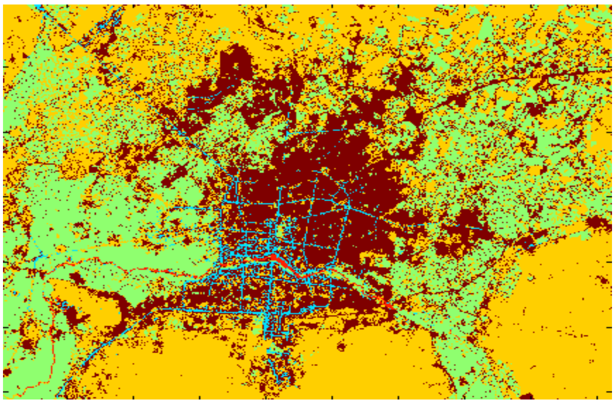
As mentioned, the calibrated rules are the minimum required number of pixels of land uses for urbanization of a non-urban pixel. These rules as a set of numbers can be interpreted as follows: In case of window size of 5 the calibrated cellular automata transition rules are 11, 3, 23, and 12. It means that in the Moore neighborhood of a non-urban pixel must be at least 11 pixels of urban land use, or 3 pixels of street land use, or 23 pixels of vegetation land use and or 12 pixels of water land use. Therefore, it can be drawn that the order of urbanization preference is street, urban, water and vegetation land uses.



Simulated land use map using 3×3 neighborhood size



Simulated land use map using 5×5 neighborhood size



Simulated land use map using 7×7 neighborhood size

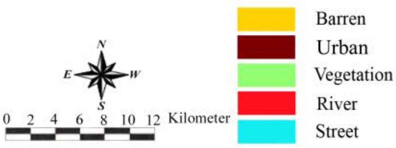


Fig. 4. Isfahan simulated land use maps in 2001 using different neighborhood sizes.

CONCLUSIONS

In this paper, the calibration procedure of the cellular automata urban model was outlined along with the results for the Isfahan Metropolitan Area, Iran. Since cellular automata urban models involve large number of variables, it needs an effective method for calibration. This exhaustive search for finding the optimum solutions will be reduced only through metaheuristics methods such as genetic algorithms. Genetic algorithms due to the effective search and high performance could be a suitable method for calibrating urban cellular automata. The paper verified that the neighborhood size of the cellular automata model depends on the urban structure of study area and varies from one region to another. Also a larger neighborhood size generates a smoother pattern compared to the smaller one. Future works encourage doing the calibration process in spatial units smaller than townships and acquiring the model parameters to test the effect of spatial modeling unit size on the reliability of modeling.

REFERENCES

Al-Ahmadi, K., See, L., Heppenstall, A., & Hogg, J. (2009). Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. *Ecological Complexity, 6*(2), 80-101.

Al-Kheder, S. A. (2007). *Urban growth modeling with artificial intelligence techniques.* PURDUE UNIVERSITY.

Alkheder, S., & Shan, J. (2005). Cellular Automata Urban Growth Simulation and Evaluation-A Case Study of Indianapolis, Geomatics Engineering, School of Civil Engineering, Purdue University.

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). A land use and land cover classification system for use with remote sensor data. Geological Survey Professional Paper 964. *US Geological Survey Circular, 671*, 7-33.

Balling, R. J., Taber, J. T., Brown, M. R., & Day, K. (1999). Multiobjective urban planning using genetic algorithm. *Journal of urban planning and development, 125*, 86.

Batty, M., Xie, Y., & Sun, Z. (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, environment and urban systems, 23*(3), 205-233.

Clarke, K. C., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B, 24*, 247-262.

Colonna, A., DiStephano, V., Lombardo, S., Papini, L., & Rabino, A. (1998, 07–09 October). *Learning urban cellular automata in a real world: The case study of Rome metropolitan area.* Paper presented at the ACRI'98 Third Conference on Cellular Automata for Research and Industry, Trieste.

Couclelis, H. (1985). Cellular worlds: a framework for modeling micro-macro dynamics. *Environment and Planning A, 17*(5), 585–596.

Dragicevic, S. (2004, 27-30 June 2004). *Coupling fuzzy sets theory and GIS-based cellular automata for land-use change modeling.* Paper presented at the Fuzzy Information, 2004. Processing NAFIPS '04. IEEE Annual Meeting of the.

Feng, Y., Liu, Y., Tong, X., Liu, M., & Deng, S. (2011). Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. *Landscape and Urban Planning, 102*(3), 188-196. doi: 10.1016/j.landurbplan.2011.04.004

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*: Addison-wesley.

Goldstein, N. C. (2004). Brains vs. Brawn—comparative strategies for the calibration of a cellular automatabased urban growth model (Vol. 3): Boca Raton, FL: CRC Press.

Hagen, A. (2003). Multi-method assessment of map similarity. *International Journal of Geographical Information Science, 17*(3), 235-249.

Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*: The MIT press.

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

Itami, R. M. (1994). Simulating spatial dynamics: cellular automata theory. *Landscape and Urban Planning, 30*(1-2), 27-47.

Li, X., Yang, Q. S., & Liu, X. P. (2007). Genetic algorithms for determining the parameters of cellular automata in urban simulation. *SCIENCE IN CHINA SERIES D EARTH SCIENCES-ENGLISH EDITION-, 50*(12), 1857.

Monserud, R. A., & Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling, 62*(4), 275-293.

Norte Pinto, N., & Pais Antunes, A. (2007). Cellular automata and urban studies: a literature survey. *ACE: Arquitectura, Ciudad y Entorno, núm. 3, Febrero 2007*.

Pijanowski, B. C., Pithadia, S., Shellito, B. A., & Alexandridis, K. (2005). Calibrating a neural network-based urban change model for two metropolitan areas of the Upper Midwest of the United States. *International Journal of Geographical Information Science, 19*(2), 197-215. doi: 10.1080/13658810410001713416

Pontius Jr, R. G., & Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment, 85*(1-3), 239-248. doi: 10.1016/s0167-8809(01)00187-6

Portugali, J., & Benenson, I. (1995). Artificial planning experience by means of a heuristic cell-space model: simulating international migration in the urban process. *Environment and Planning A, 27*, 1647-1647.

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.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography, 46*, 234-240.

Wolfram, S. (1984). Cellular automata as models of complexity. *Nature, 311*(5985), 419-424.

Wu, F. (1998a). An experiment on the generic polycentricity of urban growth in a cellular automatic city. *Environment and Planning B: Planning and Design, 25*(5), 731-752.

Wu, F. (1998b). Simulating urban encroachment on rural land with fuzzy-logic-controlled cellular automata in a geographical information system. *Journal of Environmental Management, 53*(4), 293-308.

Wu, F. (2002). Calibration of stochastic cellular automata: the application to rural-urban land conversions. *International Journal of Geographical Information Science, 16*(8), 795-818.

Yeh, A. G. O., & Li, X. (2002). *Urban simulation using neural networks and cellular automata for land use planning.* Paper presented at the Symposium on geospatial theory, processing and applications, Ottawa.