

## Modeling of urban growth using cellular automata (CA) optimized by Particle Swarm Optimization (PSO)

M. H. Khalilnia<sup>a,\*</sup>, T. Ghaemirad<sup>b</sup>, R. A. Abbaspour<sup>c</sup>

<sup>a,c</sup> Department of Surveying Eng., College of Engineering, University of Tehran  
{hkhalilnia.eng, abaspour} @ut.ac.ir

<sup>b</sup> Faculty of Geodesy and Geomatics Eng., K.N.T. University of Technology  
tehereh\_ghaemirad@yahoo.com

**KEY WORDS:** Urban Growth, Cellular Automata, Particle Swarm Optimization, Logistic Regression

### ABSTRACT:

In this paper, two satellite images of Tehran, the capital city of Iran, which were taken by TM and ETM<sup>+</sup> for years 1988 and 2010 are used as the base information layers to study the changes in urban patterns of this metropolis. The patterns of urban growth for the city of Tehran are extracted in a period of twelve years using cellular automata setting the logistic regression functions as transition functions. Furthermore, the weighting coefficients of parameters affecting the urban growth, i.e. distance from urban centers, distance from rural centers, distance from agricultural centers, and neighborhood effects were selected using PSO. In order to evaluate the results of the prediction, the percent correct match index is calculated. According to the results, by combining optimization techniques with cellular automata model, the urban growth patterns can be predicted with accuracy up to 75%.

### 1. INTRODUCTION

Due to the high concentration of population in urban areas and rapid growth of industries and services, the areal developments has had a high speed in recent decades in urban areas of Iran. Urban development has often the meaning of urban growth or creating new towns or both of them. Thus, in this regard, the rapid change in the pattern of urban land use within a short period of time can be seen. on the other hand, understanding the mechanisms of urban development is crucial for planning and urban management in order to achieve sustainable urban development. Therefore, many researchers have developed the models for study of urban growth. By using these models, it can be possible to predict different scenarios for urban development before it happens in order to be aware of the cons and pros in the decision-making (Rezazade, 2009).

One of the major methods to study urban growth is using cellular automata (CA). The concept of automated cells, for the first time, was initiated in the field of computer science in 1940 by Von Neumann and Stanislaw Ulam (Junfeng, 2003; Singh, 2003). Then this concept was developed by Conway as the theory of game of life. The entrance of the CA models into the geography, is the outcome of the Tobler's work in the 1970s at the University of Michigan (Ziaeiyan Firoozabadi, 2009). 1990s can be a successful decade in the development of urban CA models. In the past two decades, CA has been used by many researchers in the urban studies (cf. Batty et al., 1997, 1999; Wu and Webster, 1998; Li and Yeh, 2000, 2001, 2002; Clark et al., 1997; White and Engelen, 1997). CA forms by a regular grid of cells where each cell can be given a certain value according to its location. These values can be changed according to the rules, defined in the model as the transition rules, and the values of neighboring cells in a discrete time intervals (Wolfram, 1984, 2002). Due to the bottom-up approach of this model, it can simulate the overall behavior of the urban growth by considering the behavior of all the urban cells which are affected by the current conditions of the central cell and the neighboring cells. In the cellular automata, the behavior of a complex system, such as an urban system, can be simulated using transition rules. Transition rules calculate the probability of the conversion of a cell as a function of the driving forces of

urban growth and position of all neighboring cells (Feng et al., 2011). Spatial parameters of the transition rules have often a severe spatial correlation which can reduce the precision and accuracy of modeling. Nowadays, how to manage these variables and propose solutions to minimize the effects of correlation has become one of the most important research topics (Li & Yeh, 2002a; Al-kheder et al., 2008). Extensive researches have been carried out recently in the field of optimizing transition rules using evolutionary algorithms such as genetic algorithm and ant colony algorithm. Min et al (2010) showed that the proposed model of developing an artificial neural network based on genetic algorithm can better identify the spatial factors and also simplify transition rules in order to optimize them for achieving more accurate simulated model. Yang, & Liu (2007) also used a genetic algorithm to optimize the parameters derived from logistic regression.

PSO optimization method is another way of evolutionary computation. But it has some differences with the genetic algorithm, thus in this method, the optimal particles produced in each stage will be allowed to transmit their data to the next generation (Andries, 2006; Eberhart & Shi, 1998). Therefore, in this method, unlike the genetic method, the probability of transferring defective particles is almost zero and the accuracy of optimization will significantly increase.

In this paper, the urban growth of Tehran during 1998 to 2010 has been modeled after determination of the effective elements in urban growth and optimizing the coefficients of the regression equations, as transition function in a cellular automata approach, with PSO method. Then the accuracy of the simulated model was evaluated by calculating the error matrix and the parameter of percent correct match (PCM).

### 2. PROPOSED METHODOLOGY

In recent years, various researchers have shown that CA is an appropriate tool for modeling urban change and they have also used it in a variety of urban phenomenon such as traffic simulation, land use change, land cover change, and urban growth. The main elements of the cellular automata methods include cellular networks, cell status, neighborhoods, time, and transition rules.

The main focus of this article is to build a new model in order to predict the urban growth of the city of Tehran. To achieve this aim, a homogeneous cellular network with two-states, as urban and non-urban, over two classified satellite images with a period of 10-years difference is used. The Logistic regression is used as the transition function to assess the urban growth. This function has the ability to model the external effects and nonlinear interactions between the relevant variables for its nonlinear nature.

Because there are high levels of dependencies between variables in urban growth while it is almost impossible to remove these dependencies, researchers have to use only some of the factors affecting the urban growth with less dependency. Table 1, shows the factors considered in the modeling of urban growth in Tehran.

Variables	Type of variables	Definition of variables	Range of variables
$D_{Urban}$	Spatial	Distance of each cell from urban cells	$0 \leq D_{Norm} \leq 1$
$D_{Main Road}$	Spatial	Distance of each cell from road cells	$0 \leq D_{Norm} \leq 1$
$D_{Agriculture}$	Spatial	Distance of each cell from agriculture cells	$0 \leq D_{Norm} \leq 1$
$D_{Neighborhood}$	Spatial	The number of urban cells in the neighborhood of each cell	depends on the Neighbourhood definition

Table 1. factors affecting the urban growth in Tehran

The probability for a cell in the position of  $(i,j)$  in the grid to convert from non-urban to urban state can be calculated from the following equation (White & Engelen, 1993; Wu,2002):

$$p_{ij}^t = (p_l)_{ij} \times (p_\Omega)_{ij} \times con(.) \times p_r \quad (1)$$

where  $(p_l)_{ij}$  shows the local value of the cell for conversion from non-urban to urban and  $(p_\Omega)_{ij}$  shows the value of the cell for conversion from non-urban to urban due to its neighbourhood cells. Due to the limiting condition of a cell, the variable,  $con(.)$ , can be 0 or 1. In this research, placing the cell on the main roads and highways is considered as a constraint through the analysis process. Factor  $p_r$  is used in order to model the effect of random errors. There are some parameters for each of the mentioned factors. These coefficients in the equations were optimized using PSO in this paper.

PSO is one of the most basic methods in Swarm-based optimization meta-heuristics. The fundamental distinction of swarm-based with other methods is inspiration by nature and colony living of creatures. Other optimization techniques like GA do not consider any roles for members (Individuals), while the swarm-based methods consider memory for its members (particles). There is a fundamental concept in the Swarm named Stigmergy, which means that a way to convey information to others (other Particles) indirectly. In fact, there is another infrastructure except individual, and no particle transfers the information directly.

In this method, each particle tries to improve its current position using the following information:

- Current status
- Current speed
- The gap between the current status and  $P_{best}$

• The gap between the current status and  $G_{best}$  where  $P_{best}$  is the best position which each particle has achieved and  $G_{best}$  is the best position which has been obtained by the entire particles.

Equation (2) presents the fitness function used in this study.

$$F(a) = \sum_{i=1}^N \sum_{j=1}^M (p_{ij}(a) - f_{ij}^0)^2 \quad (2)$$

where  $M \times N$  is the total number of urban cells,  $P_{ij}(a)$  is the probability of the urbanization of each cell (equation (1)) and  $f_{ij}^0$  represents the observed condition of each cell that can take the values 0 or 1 for urbanization or non-urbanization state of the cell in the target year, respectively.

After the optimization of the coefficients in the logistic regression equations by the proposed method, the error matrix of the simulated images were calculated and according to the results of the error matrix, the parameter of Percent Correct Match (PCM) was calculated using equation (3) (Pontius and Schneider,2001).

$$PCM = \frac{A + D}{A + B + C + D} \quad (3)$$

### 3. EVALUATION AND RESULTS

#### 3.1 Dataset for the study area

In this research the process of urban growth is modeled for the city of Tehran. Tehran is the largest city and capital of Iran according to statistics provided by the Statistical Center of Iran, the city of Tehran, with an area of about 730 square kilometers and a population of about 8.5 million, is the twenty-fifth most populous city in the world. The city of Tehran, is located in the southern part of Alborz mountains between east longitude from  $51^{\circ} 2'$  to  $51^{\circ} 36'$  with a length of approximately 50 km and north latitude from  $35^{\circ} 34'$  to  $35^{\circ} 50'$  with a length of approximately 30 km.

In this study, the two satellite images of land use of Tehran in the years 1998 and 2010 are used (figure 2). The obtained images from Landsat satellite were classified by using a supervised method in the ENVI software by the introduction of 10 learning areas for each land use. Table 2 shows the characteristics of the images:

Projection system	Base level	satellite	Spatial resolution (m)	Radiometric resolution (m)	sensor	Date of taking images
UTM Zone 39	WGS84	Landsat 5	28.5	8	TM	Sep. 27, 1988
UTM Zone 39	WGS84	Landsat 7	28.5	8	ETM <sup>+</sup>	Oct. 17, 2010

Table (2): characteristics of the obtained images for process

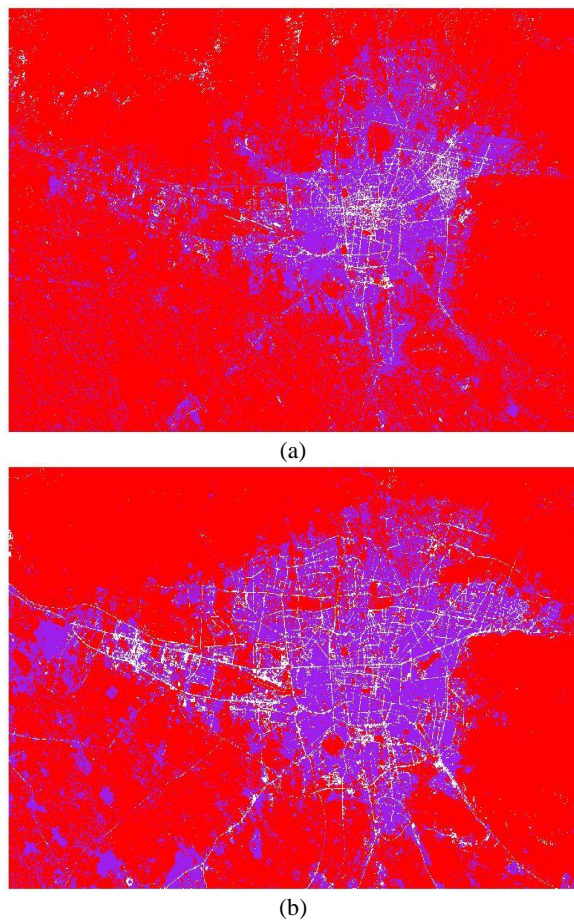


Figure 2. Two classified images used in the study, image obtained in 1998 (a) and in 2010 (b)

### 3.2 Implementation and Results

In the first step of research, the parameter of the local probability was obtained and spatial variables are defined. Then, the parameters  $P_r$  was calculated in order to adjusting Random errors. In this regard, with randomly generation of a set of weighting coefficients ( $a_p$  and  $a_0$ ) and calculating the fitness function of each one, the optimization process of the weighting coefficients with PSO method was started.

In the next step, by using the generated optimal coefficients, the probability of urbanization of each pixel was calculated via the equation (1). Finally, by defining a threshold value for the probability, the simulated image of the urban growth was generated. Figure (3) show the realistic and simulated raster of the urban growth in the year 2010.

In order to evaluate each of the anticipated patterns, PCM index is calculated. The results of calculation of this index show that by expanding the defined scope as neighborhood and increasing the influence of adjacent pixels on each pixel subsequently, the accuracy of the prediction model will increase.

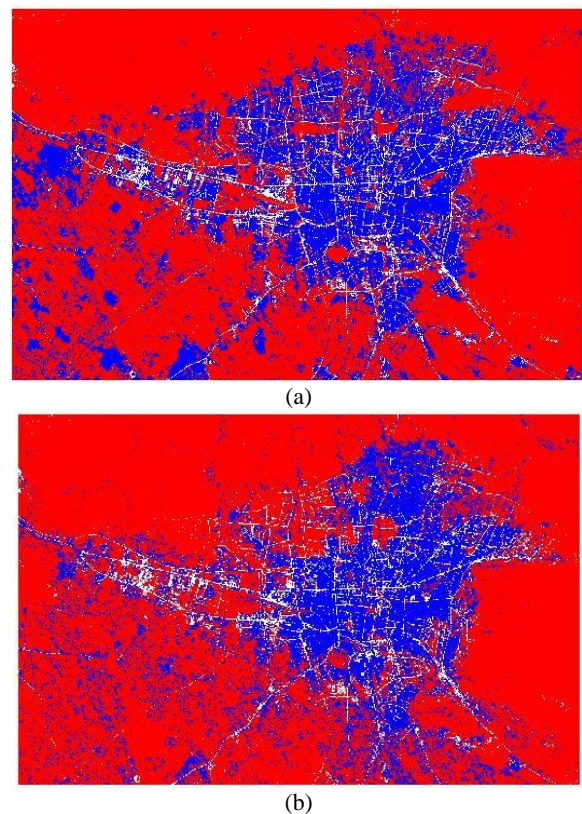


Figure 3. Real land growth of Tehran in 2010 (a) and Simulated land growth of Tehran for 2010 (b)

### 4. CONCLUSION

In this paper, two satellite images of Tehran, which were taken by TM and ETM<sup>+</sup> for years 1988 and 2010 are used as the base information layers to study the changes in urban patterns of this metropolis. The patterns of urban growth for the city of Tehran are extracted in a period of twelve years using cellular automata setting the logistic regression functions as transition functions. Furthermore, the weighting coefficients of parameters affecting the urban growth were selected using PSO. In order to evaluate the results of the prediction, the percent correct match index is calculated. According to the results, by combining optimization techniques with cellular automata model, the urban growth patterns can be predicted with accuracy up to 75%. It is also concluded that by increasing the width of the neighborhood influence, the results of the forecast can be closer to reality. Therefore, production of extensive maps of the neighborhood using spatial indicators such as enrichment factor could contribute significantly to improving the accuracy of forecasting.

### References

- Rezazade, R., Mirahmadi, M. (2009), "Cellular automaton, a novel method for simulation of urban growth." Journal of Technology Research. Vol.4, No.4, PP. 9.
- Junfeng, J. (2003). "Transition Rule Elicitation for Urban Cellular Automata models (Case Study:Wuhan-China)." MSc.thesis, Internationalinstitute for geo-information science and earthobservation (ITC).The Netherlands.
- Ziaiean Firouz abadi, P., Shakiba, A., Mottakan, A. (2009) "Remote sensing(RS),Geo spatial information system(GIS) and Cellular automata(CA) as a tool for simulating urban land use

- change (case study: city of shahrekord)." *Environmental Science*. seventh year, No.1, PP. 16.
- Batty, M., Xie, Y., Sun, Z. (1999). "Modeling urban dynamics through GIS-based cellular automata." *Computers, environment and urban systems*, PP.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*. PP.247-262.
- White, R., Engelen, G. (1993). "Cellular automata and fractal urban form: A cellular modeling approach to the evolution of urban land-use patterns." *Environment and Planning A*. PP.1175–1199.
- Li, X., Yeh, A.G.O. (2002). "Neural-network based cellular automata for simulating multiple land use changes using GIS." *International Journal of Geographical Information Science*. PP. 323–343.
- Wu, F., Webster, C. J. (1998). "Simulation of land development through the integration of cellular automata and multi-criteria evaluation." *Environment and Planning B-Planning & Design*. PP. 103–126.
- Wolfram, S. (1984). "Cellular automata as models of complexity." *Nature*. PP. 419-424.
- 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*. PP. 188– 196
- Al-kheder, S., Wang, J., Shan, J. (2008). "Fuzzy inference guided cellular automata urban growth modelling using multi-temporal satellite images." *International Journal of Geographical Information Science*. PP.1271–1293
- Li, X., Yang, Q., Liu, X. (2007). "Genetic algorithms for determining the parameters of cellular automata in urban simulation." *Science in China Series D: Earth Sciences*. PP. 1857–1866.
- Andries, P. E. (2006). "Fundamentals of computational swarm intelligence." Hoboken:Wiley.
- Eberhart, R. C., Shi, Y. (1998). "Comparison between genetic algorithms and particle swarm optimization." *Lecture Notes in Computer Science*. PP. 611–616.
- White, R., Engelen, G. (1993). "Cellular automata and fractal urban form: A cellular modeling approach to the evolution of urban land-use patterns." *Environment and Planning A*. PP. 1175–1199.
- Wu, F. (2002). "Calibration of stochastic cellular automata: The application to rural–urban land conversions." *International Journal of Geographical Information Science*. PP. 795–818.
- 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*. PP.239-248.
- Hagen, A. (2003). "Multi-method assessment of map similarity." *International Journal of Geographical Information Science*. PP. 235-249.
- 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*. PP.197-215.