A Swarm Optimization Based Method for Urban Growth Modelling

Sassan Mohammady

GIS Division, Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran

http://dx.doi.org/10.5755/j01.erem.69.3.6653

(Received in March, 2014; accepted in September, 2014)

Land use activity is a major issue and challenge for town and country planners. Urban planners must be able to allocate urban land area to different applications with a special focus on the role and function of the city, its economy, and the ability to simulate the effect of user interaction with each other. Continuing migration of rural population to cities and population increases has caused many problems of today's cities including the expansion of urban areas, lack of infrastructure and urban services as well as environmental pollution. Local governments that implement urban growth boundaries need to estimate the amount of urban land required in the future given anticipated growth of housing, business, recreation and other urban activities. Urban growth is a complex process that encounters a number of sophisticated parameters that interact to produce the urban growth pattern. Urban growth modelling aims to understand the dynamic processes. Therefore, interpretability of models is becoming increasingly important. Different approaches have been applied in spatial modelling. In this study, Particle Swarm Optimization (PSO) has been used for modelling of urban growth in Qazvin city area (Iran) during 2005 to 2011. Landsat imageries, taken in 2005 and 2011 have been used in the study. Main parameters in this study are distance to residential area, distance to industrial area, slope, accessibility, land price and number of urban cell in a 3×3 neighbourhood. Figure of Merit and Kappa statistics have been used for estimating accuracy of the proposed model.

Keywords: PSO, urban growth, GIS, urban planning

1. Introduction

Urban population worldwide has increased from 22.9% in 1985 to 47% in 2010 (Jiang et al, 2012). Studies predict that in 2025, 56.6% of the entire world’s population will live in the cities (UN, 2009). Therefore, the high density population and the public needs to absorb basic resources leads to uneven utilisation of resources (Bhatta, 2010). In some cases, the growth rate of urban areas was faster than population growth rate (Pijanowski et al, 2010). In recent years, rapid population growth and urbanisation has doubled importance of the land use. It could be stated that natural forces and human activities have been the most important factors in land use and land cover changes in local to global scale (Haub, 2007). However, the effect of human activities is more significant than natural forces. Negative effects of urban growth are the following: degradation of the environment, decreased quality of urban space and social degradation (Jenks, 2000).

Effort to better understand and analyse the relevant components of urban development can certainly help to solve urban problems (Rakodi, 2001). Understanding the urban characteristics and their interactions with other components of the environment, and use the knowledge/ability to manage the environment for effective planning requires a better understanding of the environment and the dynamic pattern of different relations in space-time.

Understanding and recognition of the spatial and temporal dynamics of complex urban systems and assessing the impact of urban growth on the environment lies in the domain of modelling and simulation, which requires innovative approaches and
techniques (Yong and Lo, 2003). Management and planning of urban spaces needs a careful monitoring of changes in patterns of land use. In this situation, urban growth models have been used by city planners and city managers. Urban modelling process consists of two phases. The first phase covers understanding of the dominant mechanisms for creating and making changes in the areas, while the second phase is prediction of the future (Foot, 1981).

Monitoring of the changes reveals the status of changes and pattern information for city managers and urban planners. A new field of soft computing in urban science is urban growth modelling. The soft computing can provide suitable and low cost solutions for environmental problems.

Today’s urban construction and environmental changes are associated with various factors. Some urban policies can lead to undesirable consequences such as urbanised areas that have led to environmental damage. Based on the discussion of sustainable cities (Jenks et al., 2000), recognised factors can be a useful in evaluating urban policies. New patterns of urban sprawl development, decentralisation, urban policy outcomes could lead to new problems. Spatial distribution of urban settlements may result in increased threats to social, physical and natural hazards. Therefore, a large number of urban growth models have been developed to be used by urban planners. The main aim is to use these models in supporting smart decisions towards sustainable development in urban areas.

2. Methods

2.1. Study area

The study area in this paper is Qazvin region (Figure 1) that includes Qazvin city and 5 small cities around it. The study area is 450 km by 360 km. Qazvin city is the capital of the province and has had the population 552928 in 1996 and this population reached 777975 in 2006. The main reason for population growth during 1996 - 2006 was increase of industrial and agricultural activities in this area. A number of new factories in this area made this city one of the most industrialised cities in Iran.

Integration of algorithms and expert systems with reasoning and inference mechanisms enables development of knowledge-based models for describing and modelling of urban processes.

In this research, PSO approach has been used for modelling the physical development of Qazvin area from 2005 to 2011. Satellite imageries used in this study come from 2005 and 2011 Landsat ETM+ sensor (Figure 2). ENVI 4.7 has been used in processing of satellite imageries. The land use map and accessibility network have been also used in shape file format in ArcGIS 9.3. In this study, attraction, economic and physical parameters have been considered. The dataset includes distance to residential area, distance to industrial area, accessibility, slope, land price and number of urban cell in a 3 by 3 neighbourhoods. All of input data is normalised according to Pijanowski et al, 2002 (Figure 3).
Particle swarm optimization (PSO) method as an evolutionary computation technique has been developed by Kennedy and Eberhart in 1995 and is based on the social behaviour simulation (Kennedy & Eberhart, 1995; Maurice, 2007). In this method, particles try to reach a better position (optimum fitness). In fact, the particles are moving towards promising regions of the search space by using their own experience during the search as well as experience of other particles (Parsopoulos and Vrahatis, 2008). Each particle stores the best position among all of the particles and the best position it has ever visited in the search space (Parsopoulos and Vrahatis, 2008). Equations 1 and 2 present the update of the velocities and positions for each particle (Clerc and Kennedy 2002).

\[v_{ij}^{t+1} = \chi \left( v_{ij} + c_1 r_1 [P_{best,i} - x_{ij}] + c_2 r_2 [G_{best} - x_{ij}] \right)\]  
\[x_{ij}^{t+1} = x_{ij}^t + v_{ij}^{t+1}\]  

Where:  
- \(i = 1, 2, ..., N;\)  
- \(j = 1, 2, ..., n;\)  
- \(c_1, c_2\) - positive constants called cognitive and social parameters;  
- \(r_1, r_2\) - independent random parameters between 0 to 1;  
- \(\chi\) - constriction factor;  
- \(G_{best}\) - the best position in all of iterations between all of particles;  
- \(P_{best}\) - the best position for each particle.

This method has been used in urban growth modelling by Pinto et al, 2011; Feng et al, 2011; Rabbani et al, 2012. According to Wu, 2002, the probability for a cell in the position of \((i, j)\) to convert from non-urban to urban state can be computed from the following equation (White & Engelen, 1993; Wu, 2002):

\[P_{ij} = (P_1)_{ij} \times (P_2)_{ij} \times Con(\cdot) \times P_r\]  

Where \((P_1)_{ij}\) is a logistic regression function which is used to obtain the value of each cell to convert from non-urban to urban. Distance to residential area, distance to industrial area, accessibility and land price are in the logistic regression as the inputs (Equation 4). During the training process, the parameters coefficient is determined.

\[(P_1)_{ij} = \frac{1}{1 + \exp[-(a_0 + \sum a_k d_k)\Sigma a_k d_k]}\]
Where: $d_k$ - effective parameters (Distance to residential area, distance to industrial area, accessibility and land price);
$a_0$ - constant parameter;
$a_k$ - coefficients of each of the independent parameters that must be calculated.

$P_{ij}$ is the number of urban cell in the cell’s neighbourhood (Equation 5). In this research, Moore neighbourhood is used as the neighbourhood definition.

$$P_{ij} = \frac{\text{Count}(S_j=u_{\text{Urban}})}{n^2}$$  

(5)

Con (.) relates to the natural limitations and constraints such as slope and elevation or proximity of specific areas in the research area. It also depends on areas that are prohibited for future development, e.g. military areas, protected areas like national parks, etc. Physical conditions like slope and elevation are the most important ones. The urban development in this area is expensive and time consuming (development of urban infrastructure, transportation and construction of facilities). Thus, existence of this factor is unavoidable. In this research, we considered slope factor as the constraint. The 20 degree slope is the threshold value. Thus, cell with slope value higher than the threshold value is 0 and lower value is considered 1. $P_i$ factor is used to model the effect of random errors. According to Wu, 2002, fitness function in this study is proposed in equation 6.

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

(6)

Our aim is to minimise the fitness function.

2.3. Accuracy assessment

Kappa. Kappa coefficient, as a statistical method is used for evaluating model performance in modelling urban development and similarity between two maps (imageries). This method is a well-known method in measuring similarity between two maps and have been used by (Monserud and Leemans, 1992, Hagen, 2003, Pijanowski et al, 2005, 2009, Foroutan and Delavar, 2012, Mohammady, 2013, Bhagyanzar et al, 2012, Fonji and Taff, 2014, Yuan et al, 2005, Erener and Düzgün, 2009, Satiprasad, 2013, Reis, 2008, Torahi and Rai, 2011, Pontius et al, 2001, Fuglsang et al, 2012, Samardzic-Petrovic, 2013). This index can help in measuring the similarity of two maps of the spatial distribution (Congalton and Green, 2009). This procedure is more appropriate because it uses all elements of comparison matrix for assessing the compatibility of simulation and reality maps. Table 1 illustrates the general shape of the probability. Table 1 is used to calculate kappa from the comparison matrix and its elements are obtained. Kappa coefficient is calculated using equation (9) (Pijanowski et al., 2005).

<table>
<thead>
<tr>
<th>Class</th>
<th>1</th>
<th>2</th>
<th>…</th>
<th>C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
<td>$P_{11}$</td>
<td>$P_{12}$</td>
<td>…</td>
<td>$P_{1C}$</td>
<td>$P_{1T}$</td>
</tr>
<tr>
<td>$P_{21}$</td>
<td>$P_{22}$</td>
<td>…</td>
<td>$P_{2C}$</td>
<td>$P_{2T}$</td>
<td></td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>$P_{C1}$</td>
<td>$P_{C2}$</td>
<td>…</td>
<td>$P_{CC}$</td>
<td>$P_{CT}$</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$P_{T1}$</td>
<td>$P_{T2}$</td>
<td>…</td>
<td>$P_{TC}$</td>
<td></td>
</tr>
</tbody>
</table>

$$P(A) = \sum_{i=1}^{C} P_{ii}$$  

(7)

$$P(E) = \sum_{i=1}^{C} P_{IT} \cdot P_{Ti}$$  

(8)

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

(9)

It is generally considered that Kappa values for map agreement are the following: >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 is very poor (Pijanowski et al., 2005, Foroutan and Delavar, 2012).

Figure of Merit. Figure of Merit (Equation 10) is a method to evaluate resemblance between actual and simulated map. This method has been developed in 2008 (Pontius et al, 2008). When a simulated map has a high goodness of fit to actual map, Figure of Merit will be high and vice versa. This factor ranges between 0 and 1. The value 1 indicates that all of the target class simulated correctly and 0 indicates that all of the target class cells have been simulated wrongly.

$$\text{Figure of Merit} = \frac{b}{a+b+c+d}$$  

(10)

Where: $a$ - error due to observed change predicted as a persistence; $b$ - correct due to observed change predicted as change; $c$ - error due to observed change predicted as wrong gaining category; $d$ - error due to observed persistence predicted as change.

Percent Correct Match (PCM). Percent Correct Match (PCM) is another evaluation method which is used for measuring similarity between two maps. Percent Correct Match is calculated using the Confusion matrix (Table 2) (Pontius and Schneider, 2001) (Equation 11). This factor ranges from 0 to 1. The value 1 means perfect agreement between two maps (simulated map and the real one). The 0.5 value means random disturbance of land use classes on the map. The 0 value mean perfect disagreement between two maps. Thus, the value closer to 1 means better similarity between two maps.
Table 2. Confusion matrix

<table>
<thead>
<tr>
<th>Model</th>
<th>Real</th>
<th>Urban</th>
<th>Non-urban</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>A</td>
<td>B</td>
<td>A+B</td>
<td></td>
</tr>
<tr>
<td>Non-urban</td>
<td>C</td>
<td>D</td>
<td>C+D</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>A+C</td>
<td>B+D</td>
<td>A+B+C+D</td>
<td></td>
</tr>
</tbody>
</table>

\[ PCM = \frac{A+D}{A+B+C+D} \] (11)

3. Results and discussion

In this study, the number of swarms proposed is 2500. Number of swarms and iteration are the parameters which have effect on PSO performance (Carlisle and Dozier, 2001). Large number of swarms makes the program time-consuming. It also requires more data on spaces. On the other hand, little number of swarms may lead to finding local optimum instead of global optimum and it is considered to be an issue. Thus, finding the proper number of swarms is an important task in working with PSO. In this study, 2500 swarms searched the space in 500 epochs.

Table 3. The logistic regression parameter

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>-12.7082</td>
</tr>
<tr>
<td>Land price</td>
<td>0.3482</td>
</tr>
<tr>
<td>Distance to industrial area</td>
<td>0.5453</td>
</tr>
<tr>
<td>Distance to residential area</td>
<td>-5.9294</td>
</tr>
<tr>
<td>Constant</td>
<td>12.3199</td>
</tr>
</tbody>
</table>

The result of modelling using the PSO structure is presented in the Figure 5. The Kappa statistics (Table 4), Figure of Merit and calculated Percent Correct Match are 0.6614, 0.5386 and 0.9752 respectively. According to Pijanowski et al, 2005 and Sousa et al, 2002, the obtained Kappa statistics from the simulation of 2011 map is considered as very good model.

![Simulated 2011 map using PSO algorithm](image)

Fig. 5. The simulated 2011 map using PSO algorithm
4. Conclusions

Integration of swarm based algorithms, Geospatial Information System (GIS) tools and remote sensing data can be considered as a powerful scientific method in monitoring, analysing and modelling urban growth. In this study, PSO algorithm is used for modelling of the development of the Qazvin city during 2005 - 2011.

A big swarm number causes larger parts of the search space to be searched and it may reduce the number of iterations needed to obtain a good optimisation result. Needless to say, a big swarm size is time-consuming.

The simulated 2011 map using PSO algorithm has had an acceptable agreement to the real 2011 map. The Kappa statistics, PCM and Figure of Merit obtained from this simulation are equal to 0.6614, 0.9752 and 0.5386 respectively. According to recent researches (Pijanowski et al., 2005, Sousa et al., 2002), the kappa statistics in this research could be considered as very good. According to Pontius et al., 2008, the Figure of Merit obtained from this research is acceptable.

According to Table 3, accessibility has obtained the highest value in the logistic regression coefficients and thus, accessibility is the most important factor in development of Qazvin region during 2005 - 2011 (Table 2). On the other hand, land price due to the least value in the logistic regression coefficients has the least effect on developing of Qazvin region.

According to Figure 2, a large volume of development in 2005 - 2011 has happened in the industrial area. This city has attracted a large population during recent years. In the context of huge population growth and increasing rate of agricultural land transformation to the urban area, the city planners and urban manager should prepare policies for future development of the city, especially in agricultural areas.

Two satellite imageries have been used to analyse urban expansion for Qazvin city. The greater number of satellite imageries and imageries with smaller ground pixel size like Quickbird and Ikonos satellite imageries data definitely increase accuracy and reliability of research results. However, it should be mentioned that these data are not free and more important than that, for a developing country like Iran, which historical urban data and land use map is not stored properly or not even existed, the research using free and reliable satellite imageries data is a practical scientific method for analysing urban growth.

In this study, the modelling and simulation procedure is done using economic, physical and attraction data. The physical data (slope), economic data (land price) and attraction data include distance to residential area, distance to industrial area, accessibility and number of urban cell in a 3 by 3 neighbourhoods. It should be mentioned that greater number of input parameters can be examined by the model. Needless to say, the model has no limitation concerning the number and type of input data. In spite of so many other models like SLEUTH, CLUE-S which unable to cover social data, this model covers all sort of data. The dataset can be improved using more input data. Attraction data like distance to green space and distance to region centers, social data like population density, population income and physical data like elevation and distance to fault, can also be examined in the model. However, the greater number of input data needs more time for training the model. In other words, the greater number of input parameters make the computation very time consuming. Thus, finding the proper number of input data is an important task.

This method is a reliable tool for urban planners and city managers to identify the future urban expansion trends and to examine their policies. Urban planners can predict urban expansion for the future periods and analyse the population growth rate to prevent occurring urban sprawl. It could also be used for area allocation for different facilities in the urban areas in the future plans.

References


Clerc M. and Kennedy J. (2002), The particle swarm-explosion, stability, and convergence in multidimensional complex space a multidimensional complex space. IEEE


Maurice C., Particle swarm optimization. 2007, London: ISTE Ltd.


Pinto N. Antunes, and Cladera J., Calibration of a cellular automata model with the particle swarm algorithm, GeoComputation 2011, Session 1B: Genetic Algorithms & Cellular Automata Modelling.

Parsopoulos and Vrahatis, IGI Global, Multi-Objective Particle Swarm Optimization Approaches, 2008, p. 20-4.2.


Satiprasad S., Monitoring urban Land use land cover change by Multi-Temporal remote sensing information in Howrah city, India, International Research Journal of Earth Sciences, 1(5), 1-6, October (2013)


**Sassan Mohammady** - GIS division, Department of Surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran. Main research area: Urban growth modelling.

Tel.: +98 9181731932

E-mail: Sassanmohammady@yahoo.com
A Swarm Optimization Based Method for Urban Growth Modeling

Miesto plėtros modeliavimas naudojantis dalelių spiečiaus optimizavimo metodą

Sassan Mohammady
GIS skyrius, Tyrimų ir geomatikos inžinerijos fakultetas, Inžinerijos kolegija, Teherano universitetas, Teheranas, Iranas

(gauta 2014 m. kovo mėn., priimtas spaudai 2014 m. rugėjo mėn.)

Žemės sklypų išnaudojimas yra pagrindinė problema ir iššūkis, su kuriuo susiduria kiekvienas miestų ir kaimų teritorijų planuotojas. Miestų planų rengėjai turi gebėti urbanistines teritorijas paskirstyti pagal paskirtį, ypatingą dėmesį skirti miesto savitoms savybėms, jo funkcionalumui, atkreipti dėmesį į miesto ekonominių lygių, numatyti ir atkurti sąveikos ryšius tarp vartotojų. Besišaisianti migracija į miestus bei populiacijos augimas sukėlė daug problemų šiandieniniams miestams, įskaitant miesto teritorijų plėtimąsi, infrastruktūros bei miesto paslaugų prieinamumo trūkumą ir aplinkosauginę taršą. Vietos valdžios institucijos, nustatančios miestų augimo ribas, turi įvertinti miesto plėtrai reikalingą žemės teritorijos dydį, numatyti naujų būstų, verslo plėtrą, poilso zonų ir kitų miestui būdingų bei reikalingų teritorijų poreikio augimą, neperžiandant nustatytų naujų miesto ribų. Miesto plėtros modeliavimo tikslas yra suprasti dinamikos procesą, dėl to pačio modelio interpretacija tampa esminė.

İrdviniam modeliavimui buvo pritaikyti keli skirtingi būdai. Šiame tyrime Qazvin miesto nuo 2005 iki 2011 metų plėtros modeliavimui buvo naudojamas dalelių spiečiaus optimizavimo (PSO) metodas. PSO yra gerai žinomas algoritmas, sudarytas remiantis dalelių spiečiaus optimizavimu, taikomu daugelyje tyrimų, ieškant optimalios parametro reikšmės. Studijų sritis apima Irano miestą Qazvin, kurio gyventojų skaicius per keletą pastarųjų dešimtmečių itin sparčiai augo. Tyrimui naudoti satelitiniai vaizdai buvo užfiksuoti 2005–2011 metų laikotarpiu. Miestų augimas yra sudėtingas struktūrinis procesas, kurio metu tarpusavyje sąveikauja gausybė sudėtingų parametrų, kurdamis augančio miesto modelį. Šiame tyrimo pagrindiniai parametrai yra atstumas iki gyvenamosios teritorijos, atstumas iki pramoninių teritorijų, šlaitas, prieinamumas, žemės kaina bei skaičius 3*3 miestinės sąvokos plotų kaimynystėje. Kad analizė būtų tiksliau, buvo naudojamas kokybės rodiklis, procentalus tikslumo apskaičiavimas ir Kappa statistika, nustatyti jų dydžiai buvo atitinkamai 0.5386, 0.9752 ir 0.6614.