Urban Sprawl Modelling.
The Case of Sanandaj City, Iran

Sassan MOHAMMADY¹, Mahmoud Reza DELAVAR¹
¹ University of Tehran, College of Engineering, Department of Surveying and Geomatics Engineering, GIS Division, Tehran, IRAN
E-mail: Sassanmohammady@ut.ac.ir, mdelavar@ut.ac.ir

Keywords: urban sprawl, modelling, Shannon Entropy, Particle Swarm Optimization (PSO)

Abstract

World urban population has dramatically increased from 22.9% in 1985 to 53% in 2013 fact that has caused unprecedented urban environmental destruction and shortage of infrastructure needed to support the population. In terms of the pressure on the built environment, urban sprawl increases the financial and environmental costs associated with infrastructure, waste disposal, energy consumption and the use of natural resources. Urban sprawl causes much damage to the natural environment by creating and furthering the spread of pollution. Thus, it becomes necessary to monitor, analyze and model the city growth. Efforts have been made to predict potential urban development in accordance with the existing and/or planned infrastructure based on smart city development plans. A number of models have been employed to detect urban land use/cover changes considering the various known drivers of land use change. The case study area is the city of Sanandaj in the west of Iran. In this study, we used Particle Swarm Optimization algorithm for modelling the urban sprawl during 1987-2000, and we employed the Landsat imageries acquired in 1987 and 2000 for modelling urban sprawl in this area for the period 1987-2000. We also considered a number of influencing factors to predict the potential urban growth modelling including the following: distance to street, distance to district centre, distance to developed area, distance to green space, slope and the number of urban cell in a 3×3 neighbourhood. We used Kappa statistics for an accurate assessment in order to compare the simulated map in 2000 with the real one.

1. INTRODUCTION

In 2008, half of the world population started living in urban areas [1]. It should be mentioned that most of the urban population growth has occurred in the developing countries [2]. In Iran, in the last decade (1996-2006), urban population growth rate was of about 2.5%, whereas the rural population showed a negative growth rate of 1.8% due to immigration and transferring of rural population to urban areas [3]. Iran’s urban population registered increases from 47% in 1976 [4] to 61% in 1996 [3], [4] and to 69% in 2013 [3]. This volume of urban growth causes so many negative effects such as urban sprawl and unplanned urban growth. Urban sprawl increases occupied spaces with high agricultural, ecological, or landscape value, the cost of supplying public services and car dependency mobility [5], [6]. Unplanned urban growth has caused many problems such as improper waste disposal, poor quality of life, noise and air pollution, polluted drinking water and traffic problems [7]. It also causes reduced open space, the deterioration of old, and unplanned or poorly planned land development [8]. In fact, a major problem occurring in cities is the lack of proper and scientific development. The results of this kind of development are the destruction of agricultural land, urban development in high slope elevations and natural hazards, increased infrastructure and utilities costs and the lack of optimum use of land [9]. Urbanization can be defined as the general transformation of land cover/use categories from being non-developed to developed which consequently cause socio-economic changes in the territorial [10], [11]. In fact, cities as complex dynamic systems are unexpected
and self-organizing, with non-linear processes [12], [13], [14]. Investigating the overall trend of conversion of non-urban land use into urban land use and recognizing the variables that influence this trend have great importance in long-term decision making and planning [15]. In fact, urban sustainable growth requires balance between economic activity, population growth, pollution, infrastructure, waste, and noise [16]. Monitoring and mapping of land use change pattern provides invaluable information for managing land resources and for city managers to project future trends of land productivity [17]. Modelling of land use change can be achieved by integrating the valuable tools of Geospatial Information Systems (GIS) and remotely sensed data [18], [19].

Spatial models are useful tools to give urban managers a better perception of urban development and assist land use policy making and urban management and provide information for evaluating environmental effects [20]. Dynamic spatial urban models enable us to evaluate future development and generate planning scenarios. These models allow for the exploring of the impacts of different urban management and planning policies implementations [7]. In other words, urban expansion models attempt to recognize the land use change drivers thought to determine conversions of land between different categories and project future changes in urban areas based on past trends [21].

2. THEORY AND METHODOLOGY

2.1. Study area

Population growth and urban sprawl of physical development in the Iranian cities have always played a significant role in environmental degradation with a strong pressure on local, regional and global conditions [22].

![Sanandaj](image)

Fig. 1. The study area.

The study area in this research is the city of Sanandaj in the west of Iran. This city has experienced fast urbanization in the last few decades. Sanandaj population has increased from about 60,000 people in 1965 to 380,000 in 2011. This large urbanization process has occurred due to social and economic attraction of this city and migration from other around cities to this city [9]. Figure 1 shows the location of study area, covering a selected area of about 13 km by 12 km.

2.2. Data preparation

Urban growth can be systematically and efficiently monitored and mapped using time series satellite data [23], [24], [25], [26], [21], [27], [28]. In this research, the two Landsat imageries acquired in 1987 and 2000 have been used. These imageries were obtained from TM sensor and were projected in World Geodetic System (WGS) 1984, Universal Transfer Mercator (UTM) zone 38N coordinate system. These imageries were classified using Maximum Likelihood classification (MLC) method in ENVI 4.7. The implemented dataset includes six parameters: distance to district centre, distance to developed area, distance to green space, slope, distance to street and number of urban cell in a 3*3 neighbourhood (Moore neighbourhood). Following Yeh and Li [29], all of the inputs parameters normalized using Eq. 1. The normalized maps are shown in Figure 2.

\[
dp_j = \frac{dp_j - dp_{\text{min}}}{dp_{\text{max}} - dp_{\text{min}}} \tag{1}
\]

2.3. Shannon entropy

In this study we used Shannon Entropy for analyzing urban sprawl in Sanandaj city during 1987 and 2000. Shannon’s entropy (Eq. 2) is a well-accepted method for determining the sprawled urban pattern [30], [23]. A single policy for the entire city never works with equal degrees of effectiveness for all [27]. Thus, the area should be categorized in different zones. In this study, the zone is defined according to distance to city centre. In other words, the distance between each urban pixel and city centre is calculated and categorized using 1000 meter interval. The implemented Shannon Entropy is defined below:

\[
H_i = -\sum_{j=1}^{m} p_j \log_2 (p_j) \tag{2}
\]

where:

- $H_i$ - Shannon Entropy for each temporal span,
- $p_j$ - proportion of built up area in $j$ zone to total built up area,
- $m$ - is the number of zones ($m=7$),
- $i$ - is temporal span (=1).
According to Bhatta, Saraswati, and Bandyopadhyay (2009), half of the $\log_e (m)$ is a threshold for determining urban sprawl [27]. In this study due to $m=7$, half of the $\log_e (7)$ is 0.9730. It means if Shannon Entropy is larger than this value, it can be safely said that this region is sprawled and if it becomes less than this value, the region is non-sprawling. The obtained Shannon Entropy is equal to 1.5782. Thus, it can be safely said that this city has experienced sprawl. Figure 3 shows urban cells in each zone.

As it is shown in Figure 3 a great part of Sanandaj urban development has occurred in distant area from the city centre.

Fig. 2. The normalized input data.

Fig. 3. Distribution of urban cell in zones.
2.4. Sprawl modeling

Due to the increasing trend of urbanization along with potential environmental consequences, urban growth modelling appears to have an unavoidable role in urban planning to assist in decisions related to sustainable urban development [31], [32], [33]. In this study a PSO based method was used for modelling urban sprawl from 1987 to 2000. After model calibration, we compared the simulated 2000 map of Sanandaj with the real 2000 map.

2.4.1. PSO

Particle Swarm Optimization (PSO) as a popular swarm intelligence based method was first introduced by Kennedy & Eberhart (1995) for optimizing continuous nonlinear functions [34]. This popularity is partially due to the fact that in the case of PSO algorithm only a small number of parameters has to be tuned and also due to the easiness of the implementation of algorithms [35]. As an evolutionary algorithm, PSO is based on a population of candidate solutions (swarm of particles), which have improved capability for solving complex problems, high convergence speed and good generality for different global optimization problems [36]. As a population-based algorithm this method performs a parallel search on a space of solutions [37]. In the PSO algorithm, starting with a randomly initialized population and moving in randomly chosen directions, each particle goes through the searching space and remembers the best previous positions of itself and its neighbours [38]. Particles of a swarm communicate good positions to each other as well as dynamically adjust their own position and velocity derived from the best position of all particles (Gbest) [38].

The number of swarms is an important issue in working with PSO. The large number of swarms makes calculation quite time-consuming but on the other hand makes the space to be searched completely. On the other hand, a small number of swarms may cause the space to be searched poorly. Thus, finding the proper number of swarms should be done carefully. In this research, 2000 particles are used as the swarm population. Following White and Engelen (2000), Wu (2002), 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 Eq. 3 [39], [40]. Thus, a probability map produces which probability of transforming each cell from other non-urban land use to urban land use is presented.

\[ P^r_{ij} = (P_j)_y \times (P_\Omega)_y \times \text{Con}() \times P_r \]  

where \((P_j)_y\) shows the value of the cell for conversion from non-urban to urban (Eq. 4).

Following Feng et al. (2011), logistic regression is the proposed method for \((P_j)_y\) (Eq. 4) [41]. \((P_\Omega)_y\) shows number of urban cell in its neighbourhood.

In this research, Moore neighbourhood was used. Thus, \((P_\Omega)_y\) is equal to ratio of number of urban cell in a 3x3 neighbourhood to number of cell in the predefined neighbourhood.

\[ \text{Con}(.) \] is a limitation factor. In this research, slope is the limitation factor. It means the value for cells with slope greater than 20 degree is 0 and else 1. Factor \(P_y\) is used in order to control the effect of the stochastic factor (Eq. 5), where \(\gamma\) is a random number in the range of (0, 1), and \(\beta\) ranges from 0 to 10 which.

This factor is represented as a stochastic disturbance in the model [41]. Eq. 6 is the fitness function in this research.

\[ f^0_0 \] is the situation of cell (urban/ non-urban).

\[ (P_j)_y = \frac{1}{1 + \exp[(-a_y + \sum_{k=1}^{m} d_k)]} \]

\[ P_y = (1 + (-\ln \gamma)^\beta) \]

\[ F(a) = \sum_{i=1}^{N} \sum_{j=1}^{M} (P_y(a) - f^0_0)^2 \]

As mentioned before, 2000 swarms (solutions) were proposed. During the program running, a list of Gbest and iteration numbers was produced.

Figure 5 presents the Gbest cost values versus iteration during 450 iterations. The Gbest cost started at 2225.3 and reached 251.9144 in 450 iterations.

The coefficients obtained from logistic regression are shown in Table 1. Distance to street had the highest coefficient among all of the parameters. On the other hand, distance to city centre had the lowest coefficient and consequently, the least importance to the growth of the city from 1987 to 2000.

Table 1. The importance of each variable.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance to city centre</td>
<td>-0.0533</td>
</tr>
<tr>
<td>Distance to developed area</td>
<td>0.2460</td>
</tr>
<tr>
<td>Distance to green space</td>
<td>-0.4474</td>
</tr>
<tr>
<td>Distance to street</td>
<td>6.3534</td>
</tr>
<tr>
<td>Constant</td>
<td>-6.096</td>
</tr>
</tbody>
</table>

Figure 4 presents the reference and the simulated 2000 map of Sanandaj.
2.5. Accuracy assessment

2.5.1. Kappa statistics

This method was used for the comparison of the two maps. Kappa statistics is much used to assess the similarity between the observed and predicted results [42]. This parameter is calculated from Table 2 [43]. According to Monserud and Leemans (1992) 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 [44].

Table 2. Contingency matrix.

<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>1</td>
<td>(P_{11})</td>
<td>(P_{12})</td>
<td>...</td>
<td>(P_{1C})</td>
<td>(P_{1T})</td>
</tr>
<tr>
<td>2</td>
<td>(P_{21})</td>
<td>(P_{22})</td>
<td>...</td>
<td>(P_{2C})</td>
<td>(P_{2T})</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>C</td>
<td>(P_{C1})</td>
<td>(P_{C2})</td>
<td>...</td>
<td>(P_{CC})</td>
<td>(P_{CT})</td>
</tr>
<tr>
<td>Total</td>
<td>(P_{T1})</td>
<td>(P_{T2})</td>
<td>...</td>
<td>(P_{TC})</td>
<td>1</td>
</tr>
</tbody>
</table>

Kappa statistics could be calculated as following [45]:

\[
P(E) = \sum_{i=1}^{C} P_{iT} \cdot P_{iT}
\]

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

The obtained Kappa statistics was 0.7293.

3. RESULTS AND DISCUSSION

Distance to street had the highest coefficient of all the input parameters and thus, it had the greatest importance in the development of Sanandaj city during 1987-2000. On the other hand, distance to city centre had the lowest coefficient and consequently, the least importance in the growth of the city from 1987 to 2000.

We proposed a number of 2000 swarms. As mentioned, the number of swarms must be determined carefully. And this is because a great number of swarms makes the computation time-consuming and besides, it causes a larger space to be searched in any iteration. On the other hand, the small number of swarms, cause a smaller part of the solution space to be searched in any
iteration. The obtained Shannon entropy value for Sanandaj growth during 1987-2000 was of 1.5782 which is greater than half of the $\log_e(7)$ (0.9730). According to the recent researches on Shannon Entropy [27], it can be safely said that the city of Sanandaj is sprawled.

4. CONCLUSION

PSO as a well-known swarm based method has been used in so many researches. In this research, PSO algorithm has been used for modelling urban growth pattern in Sanandaj city from 1987 to 2000. After calibration, the 2000 simulated map was produced and compared with the real one. The Kappa statistics obtained from the simulation of Sanandaj 2000 map is 0.7293. According to recent researches, this Kappa value consider as very good model [44].

In this study, six predictor variables include distance to district centre, distance to street, distance to developed area, distance to green space; slope and the number of urban cell in a 3*3 neighbourhood have been used. PSO method enables us to model the urban sprawl process using less input parameters, too. Also, PSO algorithm has no limitation for input data and socio-economic parameters like population, density and income can be used in the modelling procedure.

The great number of the used satellite imageries for analyzing sprawl makes the results more reliable. In fact, using more satellite imageries makes it possible to assess urban sprawl multiple times. Thus, urban sprawl trend can be examined continuously. In this study, two Landsat imageries were used for analyzing and modelling urban sprawl in Sanandaj city during 1987-2000.

According to Figure 3, from 1987 to 2000, a large amount of settlement has occurred in distant area from the city centre, especially in zones 2, 3 and 4. It should be mentioned that this kind of growth is responsible of many consequent problems such as: increasing the time travel, increasing the car dependency transportation, increasing the costs of infrastructures, living and transportation and pollutions.

Thus, a suitable policy for better management of the city and decrease the urban sprawl problems is the vertical development of the city and determine a set of certain and accurate policies about the growth of city boundaries.

The integration of cost effective remote sensing data, Geospatial Information Systems (GIS) tools and Artificial Intelligence (AI) algorithms like PSO are a powerful and scientific method in monitoring, analyzing and modelling of urban sprawl. This integration can be significantly used by city planners and land resources mangers to perceive a true perception of urban growth dimension.

REFERENCES


http://dx.doi.org/10.3390/su5020724


