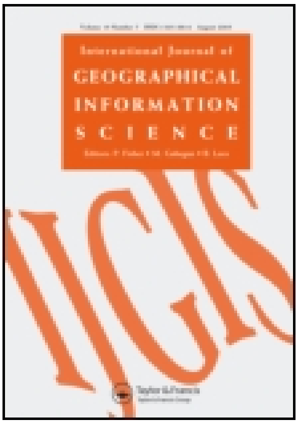

International Journal of Geographical Information Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgis20

Simulating urban growth by integrating landscape expansion index (LEI) and cellular automata

Xiaoping Liu[a], Lei Ma[a], Xia Li[a], Bin Ai[a], Shaoying Li[b] & Zhijian He[a]
[a] Guangdong Key Laboratory for Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, Guangdong, China
[b] School of Geographical Sciences, Guangzhou University, Guangzhou, Guangdong, PR China
Published online: 06 Sep 2013.

To cite this article: Xiaoping Liu, Lei Ma, Xia Li, Bin Ai, Shaoying Li & Zhijian He (2014) Simulating urban growth by integrating landscape expansion index (LEI) and cellular automata, International Journal of Geographical Information Science, 28:1, 148-163, DOI: 10.1080/13658816.2013.831097

To link to this article: http://dx.doi.org/10.1080/13658816.2013.831097

International Journal of Geographical Information Science, 2014
Vol. 28, No. 1, 148–163, http://dx.doi.org/10.1080/13658816.2013.831097

Taylor & Francis
Taylor & Francis Group

Simulating urban growth by integrating landscape expansion index (LEI) and cellular automata

Xiaoping Liu[a]*, Lei Ma[a], Xia Li[a], Bin Ai[a], Shaoying Li[b] and Zhijian He[a]

[a]*Guangdong Key Laboratory for Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, Guangdong, China;* [b]*School of Geographical Sciences, Guangzhou University, Guangzhou, Guangdong, PR China*

(*Received 21 November 2012; accepted 30 July 2013*)

Traditional urban cellular automata (CA) model can effectively simulate infilling and edge-expansion growth patterns. However, most of these models are incapable of simulating the outlying growth. This paper proposed a novel model called LEI-CA which incorporates landscape expansion index (LEI) with CA to simulate urban growth. Urban growth type is identified by calculating the LEI index of each cell. Case-based reasoning technique is used to discover different transition rules for the adjacent growth type and the outlying growth type, respectively. We applied the LEI-CA model to the simulation of urban growth in Dongguan in southern China. The comparison between logistic-based CA and LEI-CA indicates that the latter can yield a better performance. The LEI-CA model can improve urban simulation accuracy over logistic-based CA by 13.8%, 10.8% and 6.9% in 1993, 1999 and 2005, respectively. Moreover, the outlying growth type hardly exists in the simulation by logistic-based CA, while the proposed LEI-CA model performs well in simulating different urban growth patterns. Our experiments illustrate that the LEI-CA model not only overcomes the deficiencies of traditional CA but might also better understand urban evolution process.

Keywords: urban simulation; outlying growth; landscape expansion index (LEI); cellular automata

1. Introduction

Over the past 30 years, urban population increased from 1.5 billion in 1975 to 3.3 billion (50% of the world's population) in 2007. According to the United Nations' prediction, the growth trend will continue, as urban population is anticipated to exceed 60% (5.0 billion) by 2030, and the majority of this growth will occur in developing countries (United Nations 2004). Massive immigration to cities has resulted in rapid expansion of urban areas as well as land-use change. Accelerating urban growth has placed heavy pressure on land resources and has brought about a series of environmental and social problems (Liu et al. 2012). Planning and management in these fast-growing regions has become more complex and difficult (Leao *et al.* 2004). Urban planners are faced with the challenge of urban expansion and are required to seek new planning techniques to solve these problems. Recently, computer-based land-use models have been used to address challenges in the fast urbanization process (Herold *et al.* 2003). These models can provide an improved ability

*Corresponding author. Email: liuxp3@mail.sysu.edu.cn

to assess future growth and to create planning scenarios and have the potential to become important tools in supporting urban planning, policy making and assessing the environmental or ecological effects of urbanization (Verburg et al. 1999, Berling-Wolff and Wu 2004, He et al. 2011).

A variety of computer-based approaches have been presented to model the spatial process of urban growth, such as spatial logistic regression (Cheng and Masser 2004), cellular automata (White and Engelen 1993, Clarke *et al.* 1997, Liu et al. 2008a), artificial neural networks (Pijanowskia et al. 2002), multi-agent systems (Benenson 1998) and land-use spatial optimization models (Liu et al. 2012). Among these models, cellular automata (CA) is probably the most popular urban simulation model because of its simplicity, its ability to reproduce complex emergent dynamics and its affinities to raster geographic information system GIS and remote-sensing data (Torrens and O'Sullivan 2001, Liu et al. 2010a). As a type of bottom–up model, CA has been widely used for simulating various geographical and ecological processes, including plant competition (Matsinos and Troumbis 2002), forest insect propagation (Bone *et al.* 2006), forest fire diffusion (Berjak and Hearne 2002), epidemic propagation (Sirakoulis *et al.* 2000), landscape change (Wang and Zhang 2001), urban growth (Batty and Xie 1994, Wu 2002) and land-use change (Clarke and Gaydos 1998, Li and Yeh 2002). These studies have shown that CA is able to effectively simulate and predict complex geographical processes.

Among these applications of CA, urban simulation has been extensively explored and has most case studies (Liu et al. 2008b). The core of urban CA is defining the transition rules, which is used to calculate development probability as a function of the neighbourhood, a series of contributing spatial factors, an inertia effect and a stochastic disturbance term. Determining transition rules is challenging when CA is used to simulate real-world cases, as it involves many spatial variables and parameters (Liu et al. 2008a). Early method for defining transition rules is based on a visual test, which determines the most optimal combination by visually comparing simulated patterns with actual ones (Clarke *et al.* 1997). A general model of transition potential is calculated as a weighted sum. Multiple criteria evaluation or logistic regression is used to determine the weights (Wu and Webster 1998, Wu 2002). Another type of transition rules are based on fuzzy-set approaches, which consider the uncertainty of human behaviour in the simulation (Liu and Stuart 2003). Recently, artificial intelligence methods have been increasingly incorporated in urban CA models, such as artificial neural networks (Li and Yeh 2002), case-based reasoning (CBR) (Li and Liu 2006), kernel-based learning machines (Liu et al. 2008b), ant colony optimization (Liu et al. 2008a) and artificial immune systems (Liu et al., 2010a). These learning algorithms aim to discover transition rules automatically based on empirical data.

Most urban CA models deem that urban development takes place in peripheral areas, which means that only regions adjacent to urban development zones can be converted into urban land. These models may well have exaggerated the role of neighbourhood in urban development. In fact, the spatial pattern of urban growth consists of three categories: infilling, edge-expansion and outlying (Liu et al., 2010b). Infilling growth is characterized by a non-developed pixel being converted to urban and surrounded by urban. Edge-expansion refers to the newly developed urban area spreading out from the fringe of existing urban patches. Outlying growth refers to the newly developed urban area occurring beyond existing developed areas. Traditional CA model considers neighbourhood as a very important factor of urban growth, simulating the infilling and edge-expansion growth patterns quite well but not having the same success with simulating growth in the outlying category. New settlements emerge in areas with high urban development suitability (flat terrain and good transportation, among others) without respect to existing urban infrastructure. These settlements tend to develop independently in the outlying growth pattern, which is quite prevalent and significant in the early expansion of cities (Heimlich and Anderson 2001, Wilson et al. 2003). Traditional CA model takes into consideration adjacent expansion (infilling and edge-expansion) only; it assumes that the newly developed growth patch is adjacent to the existing urban areas. The simulation results often suffer from the cluster effect and deviate greatly from the actual urban layout (Liu et al., 2010a). Therefore, the outlying growth pattern needs to be incorporated into the CA model to simulate real urban expansion. Attempts have been made to take outlying expansion into account in the CA model by using the SLEUTH urban growth model, which randomly selects potential newly grown cells to generate outlying expansion (Clarke and Gaydos 1998). However, random selection may not be able to catch actual areas occurred during outlying expansion.

Uniform rules cannot describe different urban expansion types due to varying driving factors. Thus, it is necessary to discover the rules of the adjacent growth (including infilling and edge-expansion) and the outlying growth patterns in the CA model. Urban expansion patterns must be classified prior to urban simulation. Liu et al. (2010b) conducted a quantitative analysis of the urban expansion process by using landscape expansion index (LEI), which allows one to quantify the dynamic changes of landscape at two or more time points. This index can examine the way in which urban landscape evolves and reveals the relationships between the spatial distribution of urban landscape as well as its evolution process. The proposed LEI divides urban landscape expansion into three categories: infilling, edge-expansion and outlying. To compensate for the CA model's inability to simulate outlying growth, this article proposes a novel model, called LEI-CA, which integrates LEI and CA for simulating different urban expansion patterns. The LEI index is used to classify urban growth types with multi-temporal remote-sensing data. Then, transition rules of different urban growth types are discovered subsequently by using CBR technique, which expresses the principles of urban evolution implicitly according to discrete case analysis rather than equation rules (Li and Liu 2006). Lastly, discriminated transition rules for different urban growth types are used to simulate urban growth. The LEI-CA model was then applied to the simulation of urban growth in Dongguan. Simulated results demonstrate that the proposed model has the ability to simulate outlying growth, and it also can achieve better simulation accuracy than logistic-based CA.

This article is organized as follows. Section 2 reviews the LEI index and defines the urban growth pattern. Section 3 describes the proposed LEI-CA model in detail. In Section 4, we present the experimental results and discussion. Then, some important conclusions are drawn from the LEI-CA model validation experiment in the end.

2. Landscape expansion index (LEI) and urban growth pattern

Landscape expansion index (LEI) was proposed by Liu et al. (2010b). It can be used to identify the expansion types of a certain landscape and its distribution patterns from multi-temporal remote-sensing data. In contrast with traditional landscape indices which only reflect the spatial characteristics at a given time, LEI can capture the information on formation processes of a landscape pattern (Liu et al. 2010b). The LEI index divides the spatial pattern of urban growth into three types, i.e., infilling, edge-expansion and outlying, while other patterns can be regarded as variants or hybrids of these three basic forms.

The LEI index is defined by using the buffer analysis, which can be used in queries to determine which entities occur either within or outside the defined buffer zone. Figure 1

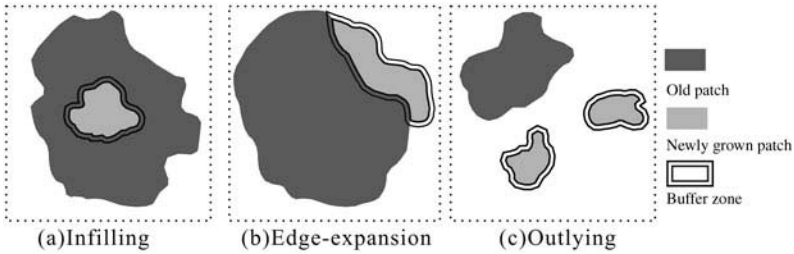

Figure 1. Three types of landscape expansion.

illustrates the buffer zone of new patches with respect to three typical urban growth patterns. A set of rules are defined to identify these growth patterns: (1) if a newly grown patch belongs to the infilling growth, the buffer zone is mostly occupied by the old patch (Figure 1(a)); (2) the area in buffer zone is mixed by vacant land and the original landscape if the newly grown patch is the edge-expansion type (Figure 1(b)) and (3) if the newly grown patch is classified as outlying growth, its buffer zone is composed exclusively of vacant land (Figure 1(c)).

Therefore, the LEI index can be calculated by examining the characteristics of its buffer zone, as shown in Equation (1):

$$\mathrm{LEI} = 100\frac{A_o}{A_o + A_v} \tag{1}$$

where LEI is the landscape expansion index for a newly grown patch, A_o is the intersection between the buffer zone and the old patches and A_v is the intersection between the buffer zone and the vacant category. According to this definition, the value of LEI varies from 0 to 100. Urban growth pattern is identified as infilling when LEI $>$ 50, edge-expansion when LEI ranges from 0 to 50 and outlying growth when LEI $=$ 0.

3. Integrating LEI with CA for urban simulation based on CBR

Most of the existing CA models assume that transition rules are static in the spatio-temporal dimension. It is unreasonable to apply the same set of rules to any location and time. CBR is the process of solving new problems by retrieving stored records of prior problem-solving cases. The use of CBR techniques can avoid knowledge-soliciting problems in CA simulation (Li and Liu 2006). Discrete cases can be obtained by multi-temporal historical data to represent spatio-temporal variations of transition rules for CA model. New temporal data can be regarded as new cases and can be added dynamically into case library with time. By this way, case library would be updated dynamically, representing the self-adaptive and self-learning capabilities of the case-based CA model. With increased cases in case library, the system would accumulate more and more experience and knowledge. This paper attempts to integrate LEI with geographical CA based on CBR, which is used to discover transition rules of different types of cell. Figure 2 shows the methodology of using CBR and LEI analysis for establishing CA model, which can simulate outlying urban growth. First, the LEI value of each newly grown patch is calculated for classifying urban growth type. Then, case libraries of the outlying growth and the adjacent growth (including infilling and edge-expansion) are established through stratified sampling. Next, for

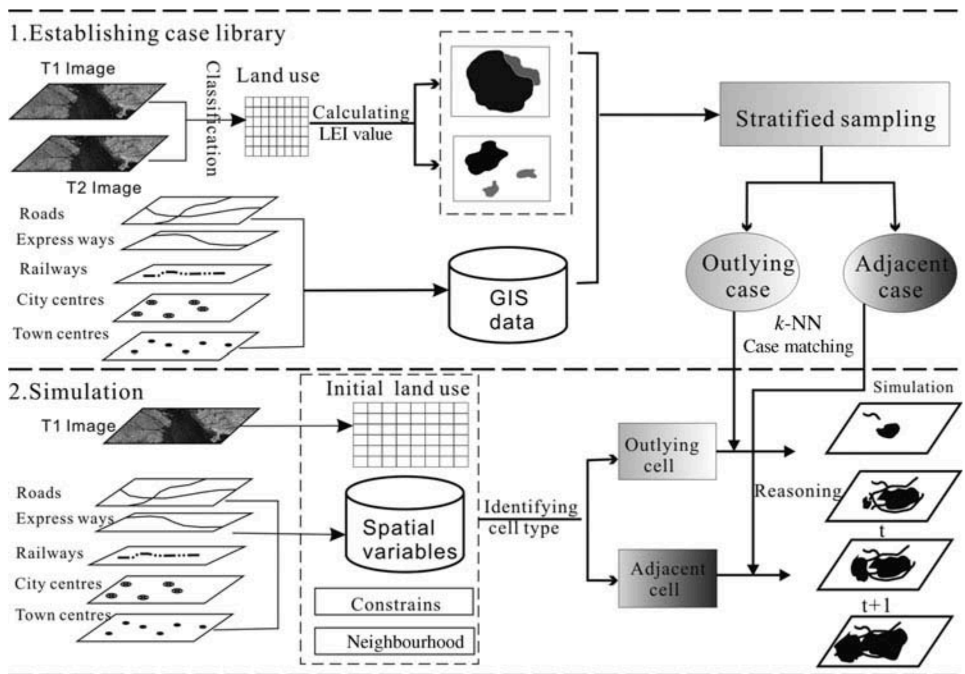

Figure 2. The procedure of LEI-CA model.

each cell, we assume that this cell will be converted into urban. Then, buffer zone of cell is created to calculated LEI value according to Equation (1). Growth type of each cell is classified as outlying growth or adjacent growth based on LEI value. Lastly, according to different types of case libraries, case matching is carried out to retrieve the experience for urban simulation by using *k*-NN algorithm. Detailed procedures of LEI-CA model are provided in the following sections.

3.1. Establishing different types of case libraries based on LEI

The first step in the LEI-CA model is to identify the pattern of the newly grown patch by using LEI, which divides urban landscape expansion into three categories: infilling, edge-expansion and outlying. Neighbourhood influence is a very important factor for urban growth in infilling and edge-expansion patterns, which can be merged into the adjacent growth. On the contrary, the outlying growth means that new settlements emerge in areas irrespective of the existing urban infrastructure. It is therefore necessary to discover transition rules of adjacent growth and outlying growth . The CBR technique is then used to establish the CA model. Cases are the basic units for the reasoning process. The first step in CBR is to construct the case library which will replace explicit transition rules or equations. In the proposed LEI-CA model, each case is represented as a vector which consists of three parts: attributes (features) of each cell, case type (adjacent or outlying) and state conversion (solution). Attributes include proximity (distance) variables and land-use types. The proximity variables may involve distance to road, distance to town centre, distance to city centre, and distance to railway. Several land-use types are considered in the simulations, including water, forest and protected farm lands, which are incapable of being converted into urban land. Orchard, grass, shrubland and farm lands can be converted into urban land. The second part is the type of each case which is identified by the calculated LEI.The state conversion part is the solution, which determines if the state will be converted into an urban area or not. A case can then be represented formally as follows:

$$I = (a_1(i), a_2(i), \cdots, a_N(i); C; S) \tag{2}$$

where $a_1(i), a_2(i), \cdots, a_N(i)$are the attributes of case i, C is represented as the type of case and S is set as a Boolean variable to represent the state of the case; if the case is urbanized, $S = 1$, otherwise, $S = 0$. N is the total number of the attributes. Stratified random sampling method is used to retrieve only a portion of original data for establishing the case libraries, which can be divided into outlying and adjacent cases.

3.2. Case matching based on the k-NN algorithm

Case matching is the process of finding queried cases that are close to the current one from the case library. The matching is usually based on the similarity between a queried case (i) and an existing case (j) of the samples in the case library. The similarity can be measured by the distance between the two vectors in the attribute space. In this study, the similarity is calculated as follows:

$$\mathrm{Sim}(i,j) = 1/(1 + \sqrt{\sum_{l=1}^{n} (a_l(i) - a_l(j))^2}) \tag{3}$$

where $a_l(i)$ is the lth feature of a case.

Then, the forecasting result is generated by using the k-NN algorithm, which is a method for classifying objects based on closest training examples in the feature space. The k-NN algorithm is a basic approach to CBR, which calculates the similarity between each queried case and the training cases in the case library to determine its nearest neighbour list. Intuitively, the k-NN algorithm assigns to each new queried case the majority class (state) among its k-nearest neighbours (Li and Liu 2006). A closer neighbour indicates that the neighbour is more similar to the new queried case and should therefore carry a high weight in making the decision. The outcome of the queried case is estimated by using the following expression:

$$\hat{f}(i) \longleftarrow \arg\max \sum_{j=1}^{k} w_j \cdot \delta(C, S, f(j)) \quad \begin{cases} \delta(C, S, f(j)) = 1, \text{if } S = f(j) \\ \delta(C, S, f(j)) = 0, \text{if } S \neq f(j) \end{cases} \tag{4}$$

where k is the total number of the nearest neighbours, C is the type of case library and S is the finite set of target class value. In this study, it represents the state of a cell (e.g., 1 for urbanized and 0 for non-urbanized); w_j is proportional to the inversed distance function:

$$w_j = 1/\sqrt{\sum_{l=1}^{n} (a_l(i) - a_l(j))^2} \tag{5}$$

3.3. Estimating development probability by using CBR

The state conversion of a cell can be determined according to its k-nearest neighbour by using CBR. However, this can only yield a Boolean value – converted or not. The conversion probability is usually used to produce more plausible simulation results.The conversion probability of a case can thus be estimated by using the following equation (Li and Liu 2006):

$$P^{C}_{\text{proximity}}(i) = \frac{\sum_{j=1}^{k} w_j \delta(C, 1, f(j))}{\sum_{j=1}^{k} w_j \delta(C, 1, f(j)) + \sum_{j=1}^{k} w_j \delta(C, 0, f(j))} \tag{6}$$

The rules of the adjacent growth and the outlying growth are discovered in the LEI-CA model. If the queried case is outlying growth, the effect of neighbourhood is ignored. Then the development probability is estimated by the combination of$P_{\text{proximity}}(i)$, some constraints and a random factor:

$$P(i) = AP^{O}_{\text{proximity}}(i) RA \text{con}(i) \tag{7}$$

where $P^{O}_{\text{proximity}}(i)$ is the conversion probability of outlying growth, con(i) is constraint factor, which is used to adjust the conversion probability. For example, it is unlikely that urban development takes place in locations with rivers, steep slopes or protected farm lands. A is an adjusting factor and RA is an error item, as shown below

$$RA = (1 + (-\ln \gamma)^{\alpha}) \tag{8}$$

where γ is a uniform random variable within the range of 0–1 and α is the parameter to control the size of the stochastic perturbation.

If the queried case belongs to adjacent growth, the number of urbanized cells in the neighbourhood can significantly influence urban development. Thus, the development probability can be estimated by using the following equation:

$$P(i) = B \cdot P^{A}_{\text{proximity}}(i) \cdot RA \cdot \text{con}(i) \cdot \Omega(i) \tag{9}$$

where $P^{A}_{\text{proximity}}(i)$ is the conversion probability of adjacent growth, B is an adjusting factor and $\Omega(i)$ is the percentage of urbanized cells in the neighbourhood. In general, Moore neighbourhood (eight cells) is adopted in urban CA:

$$\Omega(i) = \frac{\sum_{3 \times 3} N(\text{urban}(i))}{3 \times 3 - 1} \tag{10}$$

where $\sum_{3\times3} N(\text{urban}(i))$ refers to the total number of urbanized cells in a 3×3 neighbourhood window around the cell under concern. Then, the final development probability can be estimated by combining Equations (7) and (9):

$$P(i) = \begin{cases} A \cdot P^{O}_{\text{proximity}}(i) \cdot RA \cdot \text{con}(i) & \text{if the queried case is outlying growth} \\ B \cdot P^{A}_{\text{proximity}}(i) \cdot RA \cdot \text{con}(i) \cdot \Omega(i) & \text{else} \end{cases} \tag{11}$$

4. Model implementation and results

The proposed LEI-CA model was applied to the simulation of urban expansion in Dongguan, a fast-growing city in the east of Pearl River Delta in China (Figure 3). Dongguan has an area of about 2465 km^2 and is located at the corridor between Guangzhou and Shenzhen. Rapid urban expansion occurred in this area due to fast economic development in the past 30 years. In the initial period of development, the outlying expansion was the dominant growth type (Liu et al. 2010b). Thus, the LEI-CA model is appropriate for simulating the urban expansion in this area.

The actual urban areas in the years 1988, 1993, 1999 and 2005 were obtained by the classification of the Thematic Mapper (TM) satellite images. The LEI value of each newly grown patch is calculated to classify urban growth type for three periods, i.e., 1998–1993,

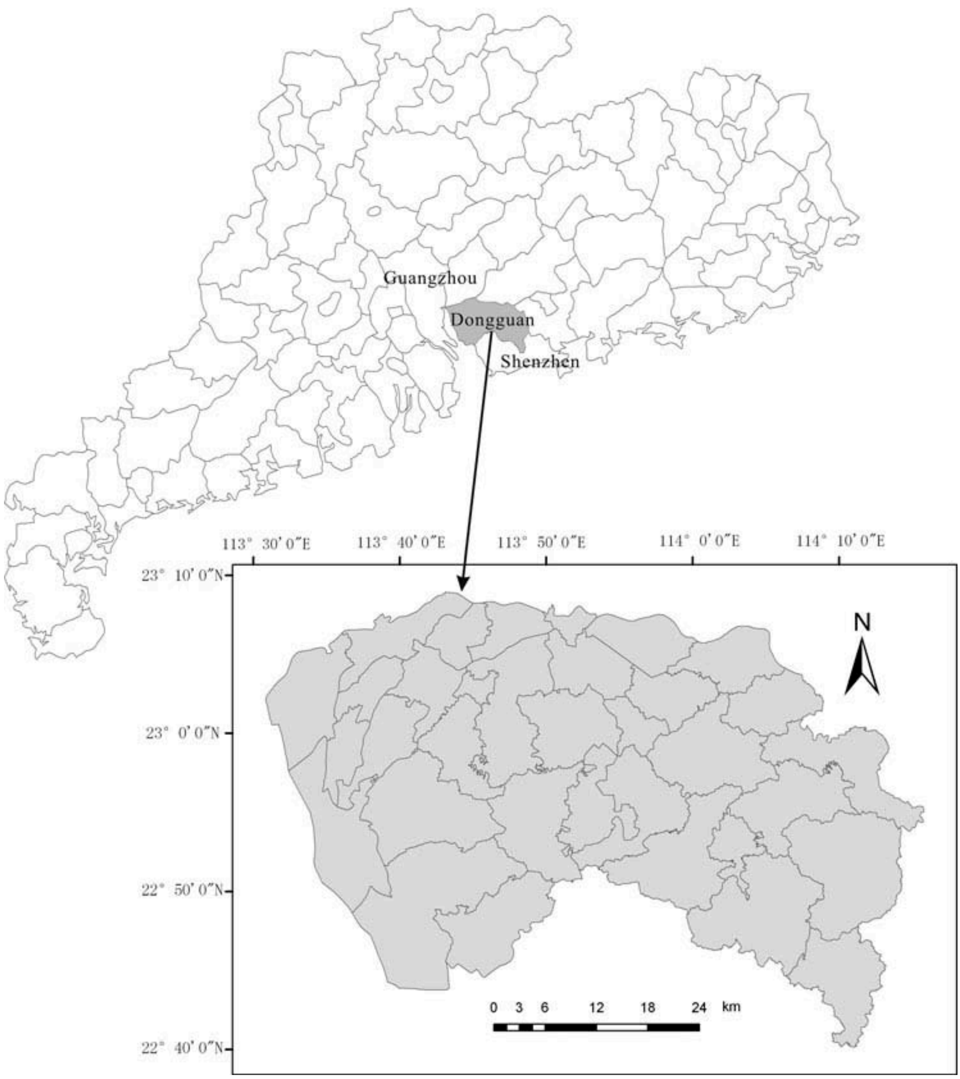

Figure 3. Location of the study area (Dongguan).

1993–1999 and 1999–2005. A series of spatial variables that are related to urban conversion were prepared using GIS function. The spatial variables include distance to city proper, distance to town centres and proximity to roads. It should be noted that the proximity to roads, used in the simulation, was dynamic for the modelling period. The dynamic changes of data on roads (1988–1999) with vector format were obtained from the digitization of the transportation maps and remote-sensing images (Figure 4). Then, stratified sampling method was employed to retrieve only a portion of the original data as a case library which consists of three major groups of cases: outlying growth, adjacent growth and non-urbanized. A total of 3000 cases, including 1000 outlying growth cases and 2000 adjacent growth cases, were obtained to represent the complex relationships between urban development and spatial variables. Previous studies indicated that the spatial expansion patterns of the study area have changed significantly in different periods (Liu et al. 2010b). In this study, the case library was updated to reflect the possible change in relationships by using additional satellite images of the 1999 and 2005 urban areas.

The proposed LEI-CA model was implemented using Visual C#.NET and ArcEngine of ArcGIS. ArcEngine is used for accessing spatial data as well as a tool for distance calculation and focal operations. Visual C#.NET is used to implement the CBR algorithm and to calculate the LEI value.

The CBR algorithm was then used to obtain transition probabilities, which were applied to the simulation of urban development in Dongguan, with the land-use of 1988 as the initial state. The buffer distance used to calculate LEI value is set to be 30 m. The simulation was conducted in discrete temporal steps. Land use in 1993, 1999 and 2005 was simulated by running LEI-CA model with 100, 200 and 300 iterations, respectively. The growth

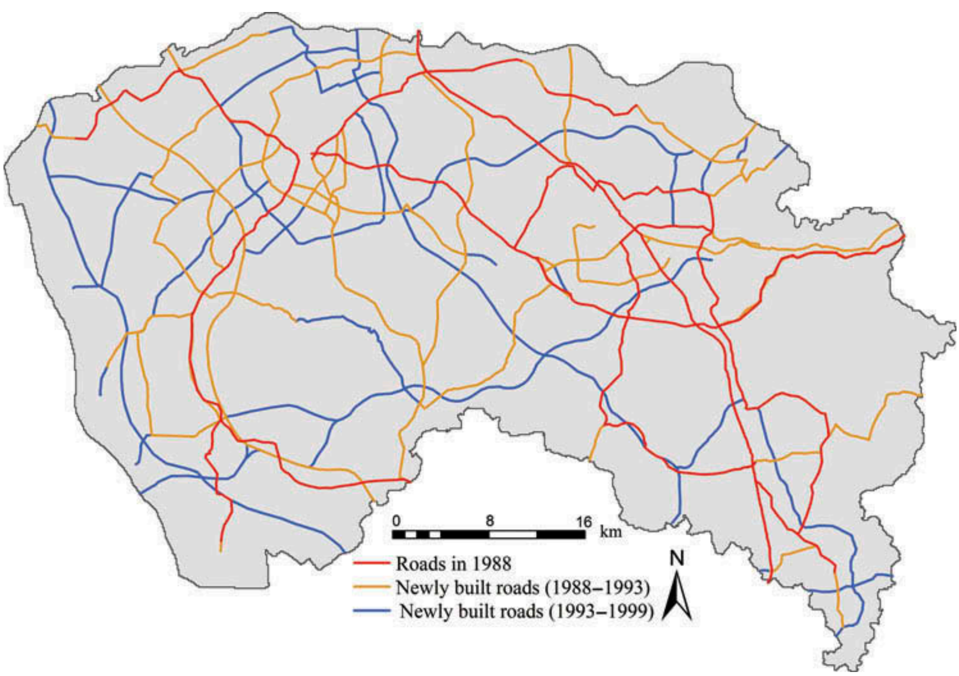

Figure 4. The change of transportation network in Dongguan for 1988–1999.

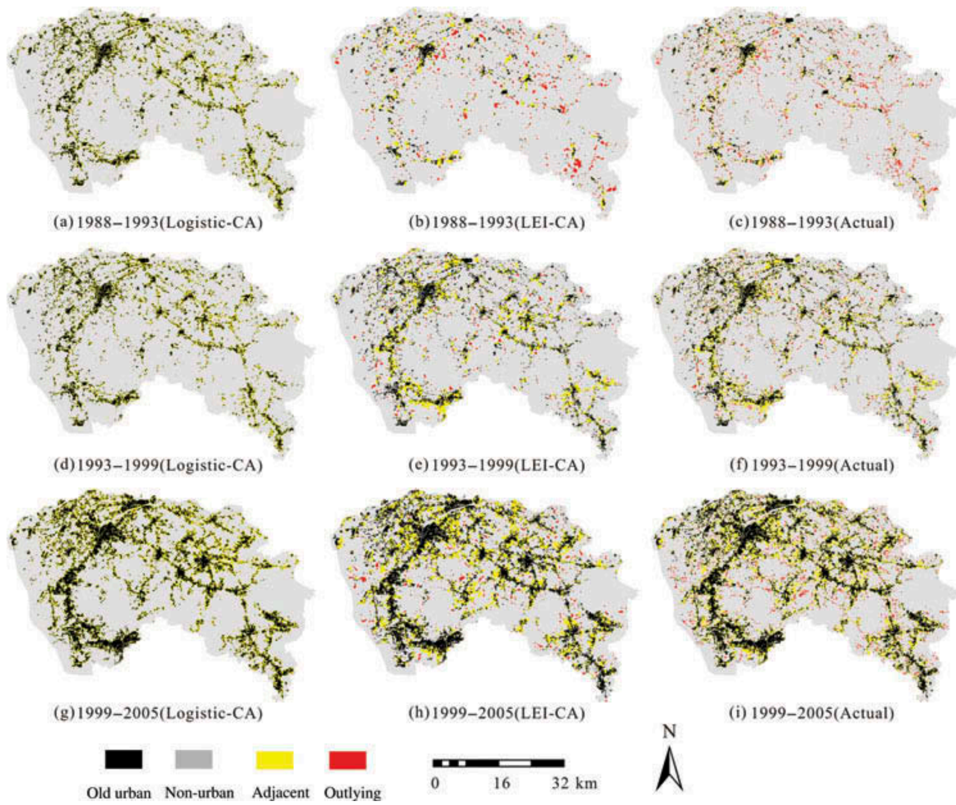

Figure 5. Simulated and actual urban development of Dongguan from 1988–2005.

types of urban patches were then classified by calculating LEI. Figure 5 shows the simulated spatial distribution of different urban growth types during the period of 1988–2005. The results demonstrate that the LEI-CA model can simulate outlying expansion for this fast-growing region.

Urban CA needs to be validated when it is applied to the simulation of real cities (Li and Liu 2006). The simplest validation method is to visually compare the simulated patterns with the actual ones. Therefore, visual inspections were conducted to compare the simulated urban areas for years 1993, 1999 and 2005 and the actual situations derived from the TM data. It was found that the simulated outlying growth and actual ones are similar (Figure 5). A confusion matrix of the concordance between the simulated and the actual development was further obtained to conduct a quantitative analysis. This matrix was calculated based on a cell-on-cell spatial overlay of these two patterns. Table 1 shows the comparison of these two patterns, which reveals that the total accuracies are 93.8%, 91.3% and 86.8%, while simulation accuracy for the urban land-use category are 64.6%, 76.7% and 81.3% in 1993, 1999 and 2005, respectively.

Another experiment was carried out to compare the performances of the LEI-CA model with those of traditional CA: logistic-based CA, which was proposed by Wu (2002) to simulate urban development in Guangzhou. In this article, logistic-based CA was applied to simulate urban dynamic development in Dongguan for comparison with LEI-CA (Figure 5). As illustrated in Figure 5, logistic-based CA has the ability to simulate

Table 1. Simulation accuracies of the LEI-CA model for Dongguan.

	Simulated 1993 non-urban	Simulated 1993 urban	Accuracy
	1988–1993 (cells)		
Actual 1993 non-urban	218,183	7688	96.5%
Actual 1993 urban	7678	14,018	64.6%
Total accuracy			93.8%
	1993–1999 (cells)		
	Simulated 1999 non-urban	Simulated 1999 urban	Accuracy
Actual 1999 non-urban	190,530	10,776	94.6%
Actual 1999 urban	10,765	35,499	76.7%
Total accuracy			91.3%
	1999–2005 (cells)		
	Simulated 2005 non-urban	Simulated 2005 urban	Accuracy
Actual 2005 non-urban	143,392	16,401	89.7%
Actual 2005 urban	16,395	71,388	81.3%
Total accuracy			86.8%

adjacent urban growth, but it does not perform well in simulating outlying growth. Table 2 shows the simulation accuracies of the logistic-based CA. The total accuracies are 91.5%, 87.4% and 82.2%, while simulation accuracy for the urban land-use category are 50.7%, 65.9% and 74.4% in 1993, 1999 and 2005, respectively. Results indicate that logistic-based CA has much lower accuracies than the LEI-CA model.

The buffer distance may have an effect on the simulated results for LEI-CA. It remains a question whether the simulation accuracy would be significantly changed if a different buffer distance is used. In order to provide a sensitivity analysis on buffer distance, five different buffer distances are used in urban simulation. The buffer distance is varied from 30 to 150 m. The LEI values calculated with different buffer distances are used to classify the growth type of cell. Then, simulation accuracies of LEI-CA under different buffer distances are obtained for implementing sensitivity analysis. As shown in Figure 6, the increase in buffer distance results in a slight decrease in simulation accuracies.

Further in-depth analysis involves comparing the actual urban growth patterns with the simulated ones in different periods. The detailed information about urban growth types was obtained by using LEI. Figure 7 shows the contribution of three urban patches in different periods. As illustrated in Figure 7(a), from 1988 to 1993, urban growth was dominated by the outlying and the edge-expansion types, whereas the infilling growth type occupied only 5.4%. In the periods of 1993–1999 and 1999–2005, the outlying type dramatically decreased. However, the proportion of the infilling type increased tremendously (5.42%–17.78%–27.11%).Throughout all periods, edge-expansion was the dominant growth type. Figure 7(b) illustrates the proportions of different growth types of the simulated patterns based on the LEI-CA model. A comparison between Figure 7(a) and (b) shows that the actual growth patterns were similar to the simulated ones using the LEI-CA model. The outlying type and the infilling type changed significantly in opposite directions. The outlying type dramatically decreased, but the percentage of the infilling type exhibited a

Table 2. Simulation accuracies of the logistic-based CA for Dongguan.

	1988–1993 (cells)		
	Simulated 1993 non-urban	Simulated 1993 urban	Accuracy
Actual 1993 non-urban	215,413	10,451	95.4%
Actual 1993 urban	10,696	11,009	50.7%
Total accuracy			91.5%
	1993–1999 (cells)		
	Simulated 1999 non-urban	Simulated 1999 urban	Accuracy
Actual 1999 non-urban	185,811	15,494	92.3%
Actual 1999 urban	15,775	30,488	65.9%
Total accuracy			87.4%
	1999–2005 (cells)		
	Simulated 2005 non-urban	Simulated 2005 urban	Accuracy
Actual 2005 non-urban	138,064	21,697	86.4%
Actual 2005 urban	22,451	65,332	74.4%
Total accuracy			82.2%

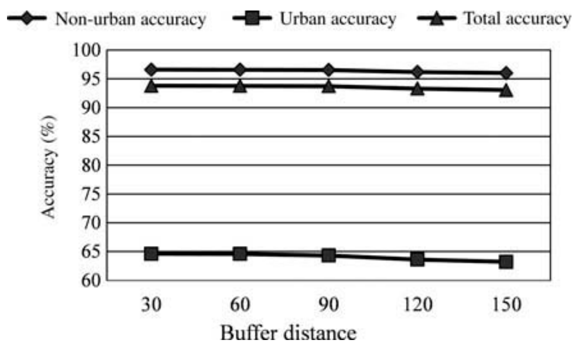

Figure 6. The changes of simulated accuracies based on LEI-CA under different buffer distances.

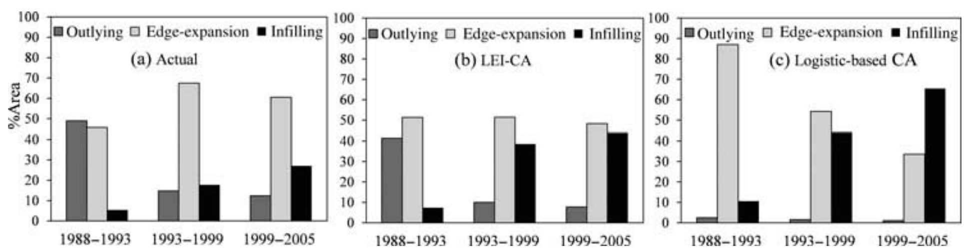

Figure 7. Percentages of growth area for the three urban growth types.

tendency to increase rapidly. This means that the rapid urbanization process in Dongguan from 1988 to 2005 has two distinct phases, namely, diffusion and coalescence. Before 1988, urban areas mainly occurred in the city centre. As the core area grows (1988–1993), it disperses growth to new cores, causing a peak in the number of the outlying type growth patches (Figures 5(b) and 7 (b)). The process of urban development during this period was called diffusion. Thereafter, urban development was seen in the periphery of existing urban land. As the simulation proceeded, the gaps between neighbourhood patches were filled by new settlements (Figure 7 (b)). Urban landscape evolved from an initial state, with the core area, to a later state where urban pixels were more likely to coalesce. This process resulted in urban growth (1993–2005) and is regarded as coalescence. Simulations by using LEI-CA model demonstrate that the urban development process is consistent with the theory of urban growth phases.

However, significant diversions exist between the actual growth patterns and the simulated results of logistic-based CA (Figure 7). The percentage of the outlying type was very low through all periods (Figure 7(c)). This indicates that logistic-based CA might not be able to simulate the outlying growth type. For the period 1993–1999, the LEI-CA and logistic-based CA models both presented a higher percentage of area for the infilling type than the actual patterns. Maybe it is because that both LEI-CA and logistic-based CA exaggerate the role of neighbourhood in urban simulation. This leads to an increase in the percentage of the infilling type and a decrease in the percentage of the edge-expansion type. From 1999 to 2005, the proportion of the infilling type increased up to 65.4%, whereas the proportion of the infilling type was only 27.11% in the actual patterns.

The histograms of LEI for three periods were produced to perform a quantitative comparison between the actual patterns, the simulated results of LEI-CA and the simulations of logistic-based CA (Figure 8). The change trends of LEI in the actual patterns are similar to the simulated results of LEI-CA. However, the LEI histograms of the simulation based on logistic-based CA is different from the actual LEI histograms (Figure 8). LEI peaks

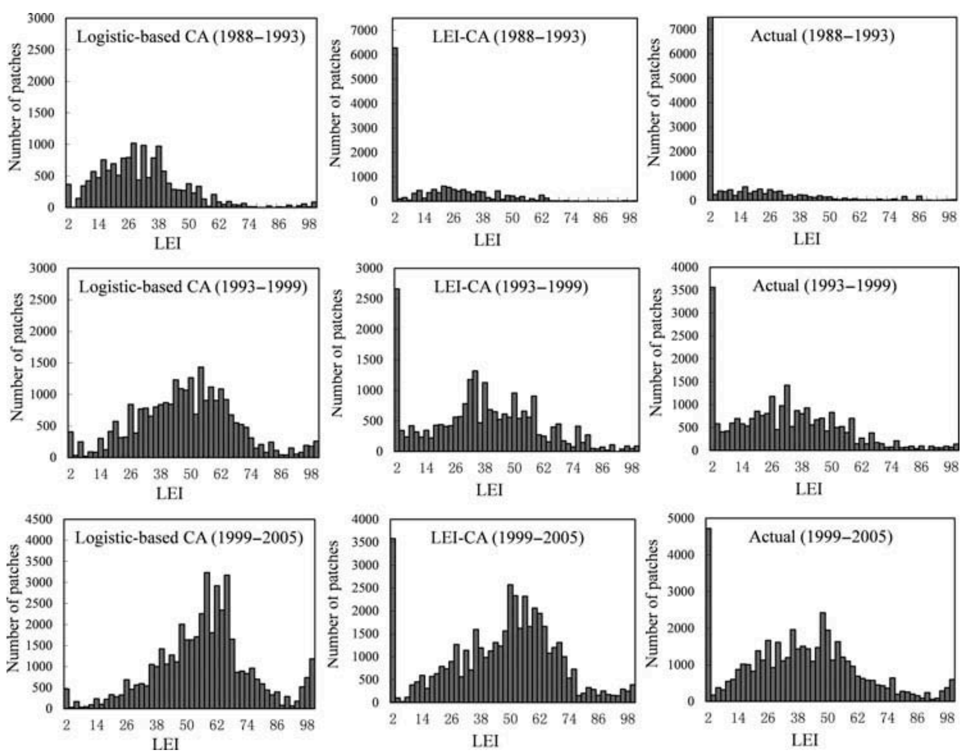

Figure 8. The histograms of LEI based on actual patterns, LEI-CA and logistic-CA.

[0] remained fairly stable for all periods in the actual LEI histograms. This means that the outlying type always exists at some level of abundance. On the other hand, the number of patches for LEI value with [0] is very low in the simulated results of logistic-based CA. Previous comparisons, including Figures 5, 7 and 8 and Tables 1 and 2, indicate that the LEI-CA model performs better than logistic-based CA in terms of simulation accuracy and structural conformity. In particular, the LEI-CA model is very suitable for simulating the outlying growth type, which can overcome the limitation of traditional CA.

5. Conclusion

Since CA was firstly introduced to solve complex geographical problems by Tobler (1970), CA models have become more popular and sophisticated and have been widely used for simulating and predicting the spatial process of urban expansion (Li and Liu 2006; Liu *et al*., 2010a). Many urban CA models are used to produce the best fit between the simulated patterns and actual ones. A major problem with most of these CA models is that they only focus on the role of neighbourhoods in urban development, which leads to traditional CA model being unable to simulate the outlying growth type. This paper proposes a new model (LEI-CA) to incorporate LEI into CA for simulating urban growth. The advantage of LEI-CA is that the outlying urban growth type can be simulated by using this model, which overcomes the deficiencies of traditional CA. Moreover, the LEI-CA model can interpret spatio-temporal processes of urban landscapes.

The proposed LEI-CA model uses different rules to simulate the adjacent growth type (including infilling and edge-expansion) and the outlying growth type. Urban growth type is identified by calculating LEI, which can capture the information on the formation processes of a landscape pattern. The CBR approach is used to discover different transition rules for the adjacent growth and the outlying growth. The self-learning ability and adaptability of CBR is potentially useful for addressing complexities in urban simulation (Li and Liu 2006).

The LEI-CA model has been successfully applied to the simulation of urban growth in Dongguan in southern China. TM satellite images in 1988, 1993, 1999 and 2005 were used to provide the actual urban areas of Dongguan. The LEI index is calculated to identify the three urban growth types mentioned earlier. Then, we establish case libraries of the outlying growth and the adjacent growth through stratified sampling. According to a different case library, case matching is carried out to discover distinct transition rules for urban simulation by using the k-NN algorithm. The urban growth of Dongguan in the period of 1988–2005 is simulated using the LEI-CA model. The comparison between the proposed model and the logistic-based CA model indicates that the LEI-CA model performs better. Our experimental results show that LEI-CA can improve urban simulation accuracy over logistic-based CA by 13.8%, 10.8% and 6.9% in 1993, 1999 and 2005, respectively. Validation of urban growth patterns also improved significantly. The change trends of LEI in the actual patterns are similar to the simulated results of LEI-CA. However, the outlying growth type hardly exists in the simulation by logistic-based CA. This comparison reveals that traditional CA, which considers only the adjacent expansion, has a natural defect, given its inability to simulate outlying growth. It also illustrates that the LEI-CA model has great potential in simulating urban growth patterns, especially for the outlying growth type. More importantly, the proposed LEI-CA model may be used to better understand urban evolution process because it is consistent with the theory of urban growth phases.

The sensitivity analysis experiment indicates that the increase in buffer distance results in a slight decrease in the simulated accuracies. The change size (e.g., size of individual urban development projects) may also have an effect on the simulated results. But, in this article, LEI-CA is a raster-based model, which assumes that each change occurred independently within 30 × 30 m pixel area. In future research, it is hoped that we can build a LEI-CA model based on irregular vector objects. The vector-based CA model will be convenient to analyse the effect of the change size on the simulated accuracies.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant No. 41171308) and a Foundation for the Author of National Excellent Doctoral Dissertation of PR China (Grant No. 3149001).

References

Batty, M. and Xie, Y., 1994. From cells to cities. *Environment and Planning B: Planning and Design*, 21, 531–548.

Benenson, I., 1998. Multi-agent simulations of residential dynamics in the city. *Computers, Environment and Urban Systems*, 22 (1), 25–42.

Berjak, S.G. and Hearne, J.W., 2002. An improved cellular automaton model for simulating fire in a spatially heterogeneous Savanna system. *Ecological Modelling*, 148, 133–151.

Berling-wolff, S. and Wu, J., 2004. Modeling urban landscape dynamics: a review. *Ecology Research*, 19, 119–129.

Bone, C., Dragicevic, S., and Roberts, A., 2006. A fuzzy-constrained cellular automata model of forest insect infestations. *Ecological Modelling*, 192, 107–125.

Cheng, J. and Masser, I., 2004. Understanding spatial and temporal processes of urban growth: cellular automata modelling. *Environment and Planning B*, 31, 167–194.

Clarke, K.C. and Gaydos, L.J., 1998. Loose-coupling a cellular automata model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, 12, 699–714.

Clarke, K.C., Hoppen, S., and Gaydos, L.J., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B*, 24, 247–261.

He, C., *et al*., 2011. Simulation of the spatial stress due to urban expansion on the wetlands in Beijing, China using a GIS-based assessment model. *Landscape and Urban Planning*, 101 (3), 269–277.

Heimlich, R.E. and Anderson, W.D., 2001. *Development at the urban fringe and beyond: impacts on agricultural and rural land*. Washington, D.C: US Department of Agriculture, Economic Research Service, AER–803.

Herold, M., Goldstein, N.C., and Clarke, K.C., 2003. The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote Sensing of Environment*, 86 (3), 286–302.

Leao, S., Bishop, I., and Evans, D., 2004. Simulating urban growth in a developing nation's region using a CA-based model. *Journal of Urban Planning and Development*, 130 (3), 145–158.

Li, X. and Liu, X.P., 2006. An extended cellular automaton using case-based reasoning for simulating urban. *International Journal of Geographical Information Science*, 20, 1109–1136.

Li, X. and Yeh, A.G.O., 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16, 323–343.

Liu, X.P., *et al*., 2008a. A bottom-up approach to discover transition rules of cellular automata using ant intelligence. *International Journal of Geographical Information Science*, 22, 1247–1269.

Liu, X.P., *et al*, 2008b. Simulating complex urban development using kernel-based non-linear cellular automata. *Ecological Modeling*, 211, 169–181.

Liu, X.P., *et al*., 2010a. Simulating land use dynamics under planning policies by integrating artificial immune systems with cellular automata. *International Journal of Geographical Information Science*, 24 (5), 783–802.

Liu, X.P., *et al*., 2010b. A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. *Landscape Ecology*, 25, 671–682.

Liu, X.P., *et al*., 2012. A multi-type ant colony optimization (MACO) method for optimal land use allocation in large areas. *International Journal of Geographical Information Science*, 26 (7), 1325–1343.

Liu, Y. and Stuart, R.P., 2003. Modelling urban development with cellular automata incorporating fuzzy-set approaches. *Computers, Environment and Urban Systems*, 27, 637–658.

Matsinos, Y.G. and Troumbis, A.Y., 2002. Modeling competition, dispersal and effects of disturbance in the dynamics of a grassland community using a cellular automaton model. *Ecological Modelling*, 149, 71–83.

Pijanowskia, B.C., *et al*., 2002. Using neural networks and GIS to forecast land use changes: a land transformation model. *Computers, Environment and Urban Systems*, 26 (6), 553–575.

Sirakoulis, G.C., Karafyllidis, I., and Thanailakis, A., 2000. A cellular automaton model for the effects of population movement andvaccination on epidemic propagation. *Ecological Modelling*, 133, 209–223.

Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region, *Economic Geography*, 46, 234–240.

Torrens, P.M. and O'Sullivan, D., 2001. Cellular automata and urban simulation: where do we go from here? *Environment and Planning B*, 28, 163–168.

United Nations, 2004. *World urbanization prospects: the 2003 revision*. New York, NY: United Nations Economic and Social Affairs, 335 pp.

Verburg, P.H., *et al*., 1999. A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. *Ecological Modelling*, 116, 45–61.

Wang, Y. and Zhang, X., 2001. A dynamic modelling approach to simulating socioeconomic effects on landscape changes. *Ecological Modelling*, 140, 141–162.

White, R. and Engelen, G., 1993. Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environment and Planning A*, 25, 1175–1199.

Wilson, E.H., *et al*., 2003. Development of a geospatial model to quantify, describe and map urban growth. *Remote Sensing of Environment*, 86, 275–285.

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

Wu, F. and Webster, C.J., 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B*, 25, 103–126.