Front. Earth Sci. 2020, 14(1): 77–89
https://doi.org/10.1007/s11707-018-0727-7

RESEARCH ARTICLE

Integrating logistic regression with ant colony optimization for smart urban growth modelling

Shifa MA (✉)[1], Feng LIU[2], Chunlei MA[3], Xuemin OUYANG[3]

1 School of Architecture and Urban Planning, Guangdong University of Technology, Guangzhou 510090, China
2 Shenzhen Longhua District Development Research Institute, Shenzhen 518110, China
3 School of Marine Sciences, Sun Yat-Sen University, Guangzhou 510275, China

Abstract Urban growth does not always strictly follow historical trends; the government may reshape urban growth patterns with considerations of ecological conservation or other plans. Both urban dynamic rules and landscape characteristics are the two main factors influencing the spatial patterns of cities, and obtaining an optimized spatial pattern is very important for sustainable urban growth. Therefore, in this study, we integrated logistic regression (LR) with the ant colony optimization (ACO) model to analyze the optimal scenario for smart urban growth. The LR model was used to discuss the relationship between urban patterns and environmental variables such as topography, development centers, and traffic conditions. Then, the urban growth probability was generated using the parameters obtained from LR. The ACO model was further integrated to optimize urban land allocation, which can meet the requirement of high growth probability, and a connected and compacted landscape pattern. This can solve the problem of urban land only being allocated by LR from being distributed fragmentarily in the space. With this integrated model, Guangzhou City, a rapidly developing area in China, was selected as a case study. The urban patterns derived from LR, as well as a simulation scenario using logistic regression-based cellular automata (LR-CA), were used in the comparison. Six landscape metrics were chosen to validate the performance of this proposed model at the pattern level. The results show that the LR-ACO model has a better performance in urban land allocation. This study demonstrated that models that couple dynamic rules and planning objectives can provide plausible scenarios for smart urban growth planning.

Keywords logistic regression, ant colony optimization, smart growth, urban planning

Received January 19, 2018; accepted July 10, 2018

E-mail: whuma@163.com

1 Introduction

Urbanization is one of the most typical phenomena in global land use and cover change (LUCC), especially in fast-developing countries such as China (Lambin and Meyfroidt, 2011). Two large international programs named IGBP and IHDP have considered LUCC, especially urban expansion in the next 50 years, to be the key field to be studied in global change science (Klepeis and Turner, 2001). It has been proven that urban expansion will result in many environmental problems, such as agricultural land loss and a decrease in biodiversity, among others (Lawler et al., 2014). We should not only adapt to this change but also control the unsustainable development trend as much as possible (Godschalk, 2004). Therefore, a core issue of urban planning is to seek an optimal pattern for sustainable urban growth (Persson, 2013). Related studies have proven that efficient decision-making models can provide support for optimal urban land allocation at the regional scale (Rindfuss et al., 2004).

In the early stages, urban land allocation in the geo-space was subjectively estimated from the experience of planners, which is a nonquantitative decision process but has set up the theoretical basis for land allocation in urban planning. Later, with the development of geographic information systems (GIS), land use suitability evaluated from the influence of multiple factors has been used by planners to allocate urban land quantitatively (Cerreta and De Toro, 2012). In addition, the multi criteria evaluation (MCE) method has been widely used to integrate the various spatial factors for suitability evaluation because of its simplicity (Liu et al., 2014). The analytical hierarchy process (AHP) method has been mainly used to define the hierarchical relationship among various spatial variables (Tudes and Yigiter, 2010). However, the weights of spatial variables are greatly dependent on the order defined by the analyzers, and the importance of each spatial factor is determined based on the decision makers' experience. Thus, the spatial patterns of urban land allocation under different weight combinations present obvious disparity, particularly because the dynamic characteristics of the influence from various spatial variables are commonly ignored.

Some studies have found that cities have self-organized characteristics, and the future status of a land cell is also influenced by its neighborhood. Dynamic simulation models such as cellular automata (CA) have been applied in the simulation of urban growth (Santé-Riveira et al., 2010). The core steps of CA modeling are to define the conversion rules, that is, to estimate the contribution of spatial driving forces, which is then used to retrieve the growth probability (Jantz et al., 2010). The higher probability the cell has, the higher the suitability for urban growth will be. In addition, a series of data mining methods, including logistic regression (LR) (Wu, 2002), artificial neural networks (ANN) (Li and Yeh, 2002), and particle swarm optimization (PSO) (Feng et al., 2011), have been applied to derive conversion rules. Those algorithm-based CA models have been widely applied for supporting urban planning to simulate urban growth in many cities (Wu and Webster, 1998; Silva et al., 2008; Aljoufie et al., 2013; Gounaridis et al., 2018).

However, in both suitability evaluation models and geo-simulation methods, only the influence of spatial factors is considered for urban land allocation; landscape characteristics are basically omitted (Santé-Riveira et al., 2010), and then chaotic and fragmentary urban growth may be generated (Poelmans and Van Rompaey, 2010). Meanwhile, the urban growth pattern in the future may not always follow historic trends, and decision makers will reshape the growth directions or poles in the next planning period with the consideration of urban growth requirements (Hess and Sorensen, 2015). Actually, landscape characteristics are the significant factors for urban land allocation in practical planning (Ligmann-Zielinska et al., 2008). Connected and compacted urban patches are preferred by planners for assigning development activities. Adjacent, small urban patches are generally grouped into one large patch in most practical planning processes.

Therefore, from the perspective of land use planning, both land suitability and landscape metrics are generally considered the main aspects for generating an optimal pattern, and they may show conflicting characteristics in geospace (Stewart and Janssen, 2014). Urban land allocation can be viewed as a typical nondeterministic polynomial (NP) problem from the perspective of spatial optimization (Li et al., 2011). Various heuristic algorithms have been developed to achieve the optimal pattern under the given objectives and constraints (Aerts et al., 2003), such as the genetic algorithm (GA) (Holzkämper and Seppelt, 2007), simulated annealing (SA) (Santé-Riveira et al., 2008), ant colony optimization (ACO) (Li et al., 2009), and PSO (Masoomi et al., 2013). Therein, ACO has been widely applied to land use allocation and zoning ecological protection areas (Li et al., 2011; Liu et al., 2012; Li et al., 2018). In those applications, the optimal patterns have been obtained from the evaluation results of land suitability (Cao et al., 2012; Liu et al., 2012), which have ignored the dynamic disciplines of cities.

As discussed above, both the dynamic rules and landscape characteristics should be taken into account for urban land allocation. The integration of simulation methods and optimization models can solve such allocation problems. LR is a classical method to estimate the growth probability; it can avoid the subjective judgment of the contribution of spatial factors and can also reflect dynamic future trends. It can also provide basic input information for performing pattern optimization. Therefore, in this study, an integrated model was proposed to retrieve the optimal pattern of urban growth, in which the growth probability calculated from LR was coupled into ACO to obtain the optimal scenario for urban land allocation. Guangzhou City, a fast-developing area located in southern China, was chosen as a case study to implement the LR-ACO model. The results derived from traditional simulation models such as LR and CA were also used in the comparison to validate the performance of the LR-ACO model.

2 Urban land allocation model integrated with LR and ACO

2.1 Definition of land use allocation for urban growth

Urban growth simulation or urban land allocation has generally been performed with a raster surface. In addition, the optimization of an urban land allocation pattern can be regarded as a mathematical problem. The urban land area to be allocated can be expressed as a set of grid cells with the size of R rows $\times$ C columns, and the cell (i, j) $(i = 1, 2, \ldots, R; j = 1, 2, \ldots, C)$ is assigned a certain type of land use. This process is mainly concerned with which nonurban cells will be converted into urban land in a planning period (Stewart et al., 2004). In the optimization model, we use a binary variable x_{ij} to describe the status of cell (i, j): if the cell is allocated with urban land, $x_{ij} = 1$; otherwise $x_{ij} = 0$. Whether the cell is allocated with urban land in the future is determined by the following factors.

2.1.1 Optimization objectives of the urban growth pattern

1) Maximum urban growth probability

The spatial pattern of urban growth is mainly determined by locational conditions, that is, a series of driving factors such as transportation, topography, and the surrounding environment (Godschalk, 2004). Therefore, first, the growth probability must be evaluated. In this study, topography and proximate variables are used to evaluate the urban growth probability (Sandhya Kiran and Joshi, 2013), which is expressed as:

$$f_{\mathrm{Devp}}=\frac{\sum_{i=1}^{R}\sum_{j=1}^{C}x_{ij}\cdot S_{ij}}{\sum_{i=1}^{R}\sum_{j=1}^{C}x_{ij}}, \tag{1}$$

where f_{Devp} is the average growth probability for each land patch, and S_{ij} is the suitability of cell (i, j) for urban growth, which can be estimated using the LR model in terms of the historical dynamic process.

2) Maximum compactness of the urban allocation pattern

In addition to locational conditions, the future status of a certain cell is also influenced by its neighboring cells. The overall pattern formed by all urban land patches is also a significant factor in the urban planning layout. If only the growth suitability of each patch is considered for urban land allocation, urban land will be distributed as a fragmentary pattern in the space. In general, a mostly compact pattern of urban patches at the landscape scale is expected for developers to perform actual urban development activities and maintain ecological processes (Olsen et al., 2007). Therefore, from the perspective of urban planning, landscape metrics are important for allocating urban land. The compactness index has been commonly used to measure the landscape pattern of an individual land patch. It is generally characterized as the ratio of area to the perimeter of the land patch, and a more compact pattern is closer to a circle. In this study, an average compactness index at the landscape scale, which can be regarded as the average compactness of all the urban patches within a defined region, was proposed for directing urban land allocation to make almost all patches more regular (Chen et al., 2010).

$$f_{\mathrm{Comp}}=\sum_{u\in U}\frac{2\pi\sqrt{a_u/\pi}}{l_u}/U, \tag{2}$$

where f_{Comp} is the average compactness of all urban land patches; l_u and a_u are the perimeter and area, respectively, of patch u, which can be calculated within a neighborhood window; and U is the total number of urban land patches.

According to the analysis above, urban land allocation can be viewed as a multi-objective problem. Many studies have discussed the solutions of multi-objective optimization processes (Cao et al., 2011; Huang et al., 2013; Masoomi et al., 2013). The linear combination of different objectives has been adopted in area optimization problems, which has proven to be an efficient method for retrieving optimal patterns (Li et al., 2011). The global objective (i.e., total utility) of urban land allocation can be expressed as the following linear combination:

$$F=\omega_{\mathrm{Devp}}\times f_{\mathrm{Devp}}+\omega_{\mathrm{Comp}}\times f_{\mathrm{Comp}}, \tag{3}$$

where ω_{Devp} and ω_{Comp} are the weights of development probability and landscape compactness, respectively, f_{Devp} and f_{Comp} are normalized within the range of 0–1. ω_{Devp} and ω_{Comp} are used to control the relative importance of the development probability versus landscape compactness. The optimal scenario can be achieved by adjusting the values of ω_{Devp} and ω_{Comp} according to the planners' preference.

2.1.2 Constraints of the optimal pattern for urban growth simulation

Though the status of a cell is mainly determined by its suitability or neighborhood influence, not all cells can be used for urban growth. Subregions, such as different administrative districts of a city, play dissimilar roles in future urban growth. The total amount of urban land in each region is also restrained within a reasonable range when considering the development situation. Therefore, quantitative and spatial constraints should be accordingly incorporated into urban land allocation.

1) Quantity demand of urban land

With rapid socio-economic development and urbanization, massive construction land is demanded for assigning development activities. The total quantity of urban land is determined by economic and social status, which will result in the different rates of urban growth in different planning periods. The quantity demand, that is, the number of urban land cells, is required to be determined and adjusted when allocating urban land. It is described as follows:

$$\sum_{i=1}^{R}\sum_{j=1}^{C}x_{ij}=Q, \tag{4}$$

where Q is the quantity demand of urban land in the whole region during the planning period, which can be forecasted by means of statistical analysis based on socio-economic factors such as the annual urban growth rate.

2) Spatial structure of urban growth

In most simulation results, additional urban land is basically scattered around present urbanized areas, and urban growth presents a typical chaotic and blind pattern, which can break the spatial balance and cause a series of ecological and environmental problems (Wei and Ye, 2014). Hence, planners have proposed the theory of polycentric patterns for urban growth; cities can be designed with multiple centers to keep the spatial balance of resources and the socio-economy (Jantz et al., 2010). During the planning period, the government reshapes the internal structure of a city by selecting some subregions as priority growth areas, which have higher growth rates and are assigned more urban land than others. The quantity of urban land for each subregion should be coordinated and constrained to obtain a practicable urban growth pattern. This can be expressed as follows:

$$\sum_{i=1}^{R}\sum_{j=1}^{C} x_{ij}\cdot H_z = \mu_z; \quad \sum_{z=1}^{Z}\mu_z = Q, \tag{5}$$

where μ_z is the quantity of urban land assigned to the z^{th} subregion z (z = 1, 2, …, Z), H_z is a binary variable that identifies whether cell (i, j) belongs to the z^{th} subregion; if cell (i, j) $\in z$, $H_{ij} = 1$, otherwise $H_{ij} = 0$.

3) Conservation of ecologically sensitive areas

Apart from the quantity constraint, land use preserved for ecological services must be excluded from urban growth. Physical or legal characteristics have been considered as important constraints to prevent a cell from urban expansion (Mitsova et al., 2011). These constraints generally include physical limitations (e.g., elevation and locations of mountains) and government-protected lands (e.g., forests, wetlands, basic farmlands, and lakes). These cells are usually classified sensitive ecological conservation areas. Additional urban land cannot encroach on cells belonging to ecological conservation areas. During the allocation process, the set of cells prevented from urban growth is updated using the following expression:

$$x = x\cdot ESA, \tag{6}$$

where ESA is a binary variable that identifies whether cell (i, j) should be prevented from urban growth: $ESA_{ij} = 0$ if cell (i, j) is classified as an ecologically sensitive area; otherwise, $ESA_{ij} = 1$.

2.2 Urban growth modeling with the coupling of LR and ACO

2.2.1 Urban growth probability estimation using LR

The urban growth pattern is influenced by driving factors from the perspective of a self-organized process. The quantitative relation between growth probability and driving factors can be estimated using LR. Whether the cell is likely to experience urban growth in the future is accordingly determined by the probability derived from the quantitative relation (Jokar Arsanjani et al., 2013), which can be described as follows:

$$S_{ij} = \frac{1}{1+\exp[-(\beta_0+\beta_1 d_1+\cdots+\beta_k d_k+\cdots+\beta_K d_K)]}, \tag{7}$$

where, S_{ij} is the growth probability of cell (i, j), β_k is the regression coefficient for the k^{th} spatial factor d_k, and K is the number of main spatial factors to urban growth. The value of d_k is normalized within the range of 0–1. In general, the regression coefficient can be obtained by training the selected sample points.

2.2.2 Optimization of urban growth pattern using the ant colony algorithm

With the optimization objectives and constraints described above, ACO is used to optimize the urban growth pattern. The ACO algorithm, which was first proposed by the Italian scholar Dorigo in the 1990s (Dorigo et al., 1996), can solve various optimization problems by simulating the behavior of ants in seeking food. During the searching process, each ant releases a chemical substance called a pheromone trail that evaporates over time. The mechanism of classical ACO can be explained by solving the travelling salesman problem (TSP), that is, to find the shortest route connecting N given cities. The probability of a city being selected by the ants is thus determined by two components: (a) the amount of pheromone trail on the path; and (b) heuristic information. Although ACO has been proven to efficiently solve problems implemented on raster surfaces, such as the zoning of land use (Ma et al., 2017), heuristic information, such as the pheromone trail and the probability calculation, should still be modified for adapting the actual objectives and constraints to optimize the urban growth pattern.

1) Global pheromone accumulation

Ants can visit any cell allowed for urban growth and release a pheromone trail on the visited cells. The urban growth pattern can be formed from the pheromone feedback of artificial ants. The amount of deposited pheromone is determined by the total utility of the urban growth pattern, that is, a solution. The larger the amount of pheromone the cell has, the more the ants will be attracted to select it. A larger amount of pheromone is in turn deposited on this cell. The communication between ants based on the pheromone feedback can correspondingly generate the maximum utility. Hence, the accumulated pheromone trail can be calculated by the total utility of the urban growth pattern selected by the ants. It is expressed as follows:

$$\tau_{ij}(t+1) = (1-\rho)\times\tau_{ij}(t) + \rho\times\frac{\sum_{n=1}^{N} x_{ij}^{n}(t)\cdot\Delta\tau_{ij}^{n}(t)}{\sum_{n=1}^{N} x_{ij}^{n}(t)},$$

$$\Delta\tau^{n} = F^{n}, \tag{8}$$

where τ_{ij} is the pheromone on cell (i, j), ρ is the evaporation rate of the pheromone, $\Delta\tau_{ij}^{n}$ is the pheromone released by the n^{th} ant on cell (i, j), and F^{n} is the total utility of the urban growth pattern selected by the n^{th} ant. Obviously, the more ants that travel on the cell (i, j), the more pheromone that will be accumulated.

2) Local heuristic information

Heuristic information is generally designed to guide the travel of ants. It is commonly calculated according to the land use suitability for an area optimization problem (Liu et al., 2014). However, optimal urban growth not only requires maximizing the suitability but also maximizing the probable connected and compacted pattern. Meanwhile, ants can adapt to the changed environment; that is, the ants will adjust their traveling path based on the local environment. Using only suitability as heuristic information will cause dispersive urban patches scattered in the geospace. Hence, both the urban suitability and local landscape metric are incorporated into the heuristic information to update the ants' status in this study. This is expressed as follows:

$$\eta_{ij}(t+1) = \omega_{\mathrm{Devp}} \cdot S_{ij} + \omega_{\mathrm{Comp}} \cdot C_{ij}^{\Omega}(t), \tag{9}$$

where η_{ij} is the heuristic function for directing the urban land allocation on cell (i, j), S_{ij} is the growth probability (i.e., urban suitability) estimated from LR, and $C_{ij}^{\Omega}(t)$ is the local landscape compactness of the urban patch around cell (i, j) at time t, which is calculated from the density of urban land to group the small urban patches together within the neighborhood Ω of cell (i, j). ω_{Devp} and ω_{Comp} represent the weight given for the growth probability and local landscape compactness, respectively. The size of the connected patches is determined based on the neighborhood window Ω. The larger the Ω is, the larger the connected patch is. Ω is designed as a Moore window, the size of which is set to decrease from the maximum value to the minimum value during the iterative process, and the decrementing value is determined based on the total number of iterations. Generally, the initial value of Ω is subjectively set based on the decision-makers' experience. In this study, the initial maximum value of Ω is assigned the average size of the urban patches in 2015, and the minimum size is set as 3×3.

In this study, assuming that there are N ants traveling on the raster surface with the size of R rows and C columns, the growth pattern selected by an ant represents an allocation scenario, and each cell can be taken as a node on the ant's traveling path. Whether the grid cell is allocated with urban land is determined by the probability, which can be described as:

$$p_{ij}^{n}(t) = \frac{\omega \cdot \tau_{ij}(t) + (1-\omega) \cdot \eta_{ij}(t)}{\sum\limits_{\forall ij \in \mathrm{Allowed}} (\omega \cdot \tau_{ij}(t) + (1-\omega) \cdot \eta_{ij}(t))}, \tag{10}$$

where p_{ijk}^{n} is the probability of the n^{th} ant selecting the cell (i, j), and ω is the weight given for measuring the contribution of the pheromone trail and heuristic information. $\forall ij \in$Allowed is the set of grid cells allowed for urban growth.

3 Case study in Guangzhou City

3.1 Study area

Guangzhou City (longitude: 112°57′–114°3′ E, latitude: 22°26′–23°56′ N) is a fast-developing area and the core city of the Pearl River Delta, which is located in South China. It includes 11 county districts and covers an area of approximately 7434.4 km^2 (Fig. 1). There are mountains in the northeast region, with hills and basins in the central section. Its southern part is adjacent to the Pearl River, which connects China's Hong Kong and China's Macao. Guangzhou's superior geographical position has made it one of the top reformers in China since 1978. With its rapid development, the GDP and permanent resident population reached approximately 2000 billion and 14.5 million, respectively, by the year 2017. However, rapid and chaotic urbanization has caused a series of ecological and environmental problems (Li et al., 2011). In addition, this urbanization will continue in the future. The optimal pattern of urban land allocation can provide a reference for sustainable urban growth planning. In this study, the optimal urban growth pattern in 2035 will be discussed in terms of the current urban planning of Guangzhou.

3.2 Data materials and data processing

3.2.1 Spatial factors for urban growth probability estimation

Remote sensing images and GIS spatial data were used to assist the urban growth pattern optimization. The Landsat TM image (122-44) collected in 2005 was used to retrieve the urban land. Other spatial data were applied to the acquisition of factors used to calculate the urban growth probability. A total of 9 spatial variables were used to discuss the driving forces of urban dynamics. These variables were: 1) proximate distance to administrative centers and the traffic network including municipal centers, district centers, large town centers, small town centers, subways, expressways, and roads; and 2) topographical conditions, including digital elevation model (DEM) and slope. All of these spatial factors were normalized within the range of 0–1 in ArcGIS software. Figure 2 shows the spatial patterns of those factors used in the LR-ACO model.

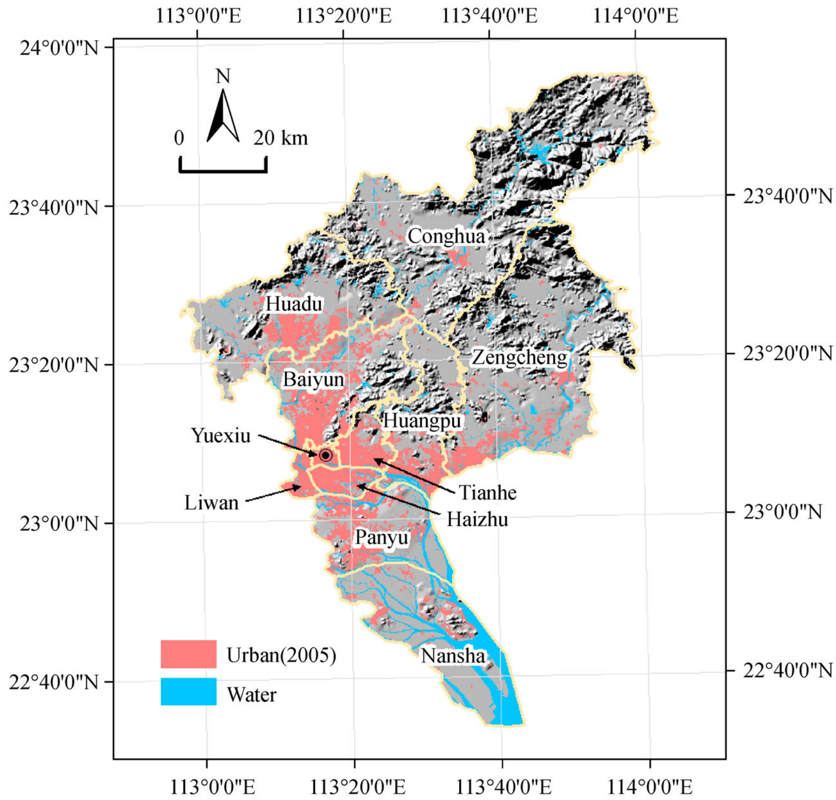

Fig. 1 Location and administrative districts of the study area.

3.2.2 Identifying the ecologically sensitive areas prevented from urban growth

The urban growth probability is estimated based on locational and topographical conditions. However, not all nonurban land can be used for actual urban growth; there are also ecologically sensitive areas (e.g., river, wetland, mountains) that are prevented from urban growth. These areas should be excluded from the set of cells for urban growth. In this study, forest-parks, rivers, and large reservoirs were identified as ecologically sensitive areas, and Fig. 3 shows the identification result.

3.2.3 Defining the size of urban growth for each subregion

During the allocation process, not only ecological conservation but also structural regulation should be considered. The government usually sets the urban growth directions to optimize the internal structure of the city. For example, Guangzhou City expanded by approximately 1045 km^2 of urban encroachment until 2005, and new built-up areas are still increasing around current urban patches, thus showing a chaotic expansion pattern and causing many city issues. From the perspective of smart growth, the policy was proposed to slow down the speed and adjust the expansion direction of urban growth in China, and the key planning concepts can be described as follows: the central city is oriented for smart growth; appropriate urban growth will be directed in the northern part, while new urban cores are expected to be located in the eastern and southern parts, especially in Nansha district. Furthermore, China's sixth state high-tech developmental zone was planned in 2009. This strategy will result in rapid urban growth in the eastern and southern parts. In particular, the urban growth in northeast Guangzhou will be seriously restrained because of ecological preservation. Considering the above planning concepts, the study area was divided into seven subregions based on administrative boundaries, geographical boundaries, topographic conditions, and the planned growth directions. In addition, the future quantity of urban land for each subregion was assigned according to local demand prediction, which was retrieved from the concept planning in 2035 of Guangzhou City (Table 1). The total quantity of urban growth for Guangzhou City in 2035 is correspondingly assigned to be approximately 1900 km^2. The distribution of subregions is also shown in Fig. 3.

3.3 Results and discussion

3.3.1 Growth probability estimated with LR

The relationship between urban growth and spatial factors was first built by LR. To build this quantitative relation, a

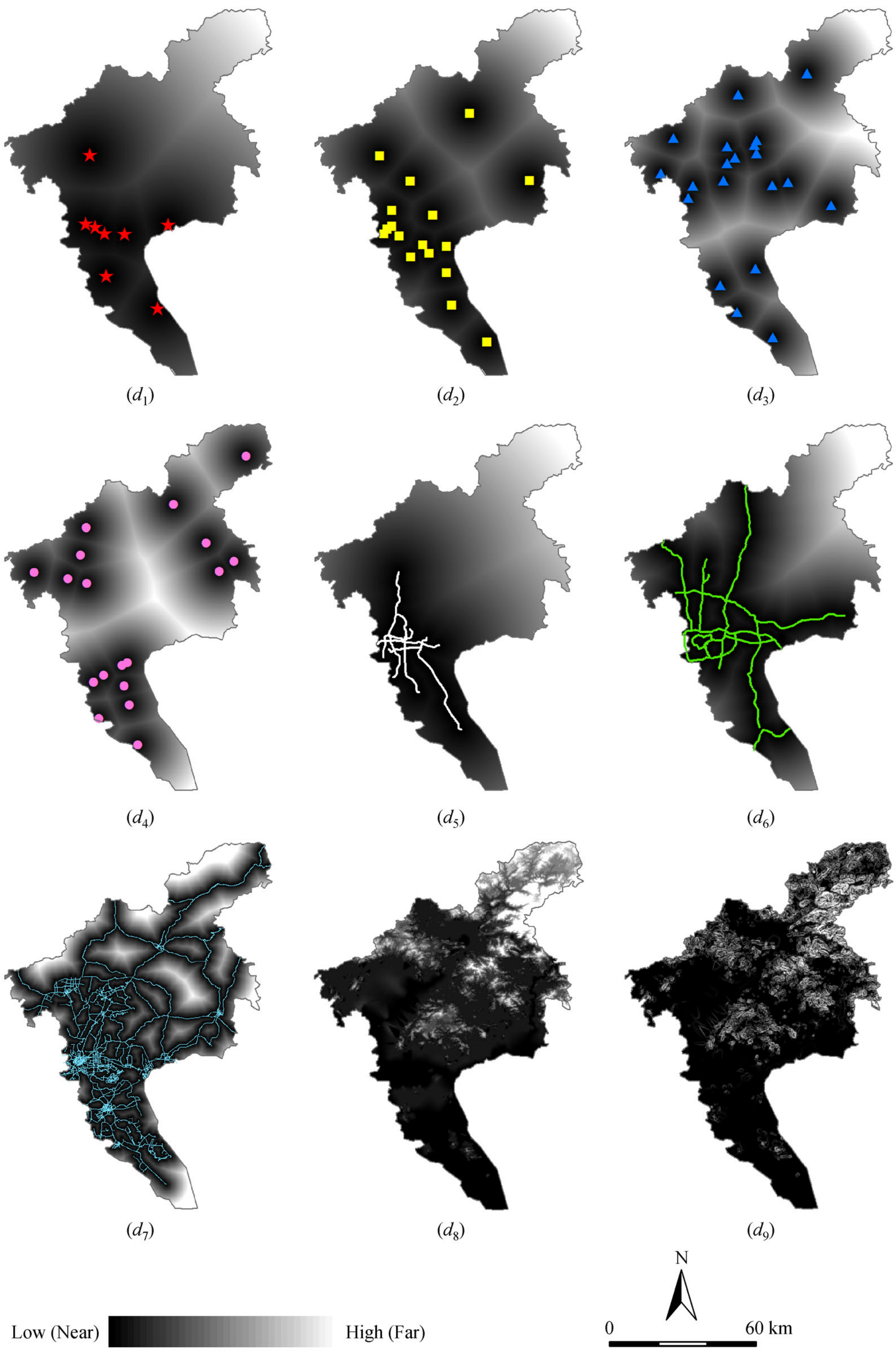

Fig. 2 Spatial variables for estimating urban growth probability: d_1) distance to the municipal centers; d_2) distance to the district centers; d_3) distance to the large town centers; d_4) distance to the small town centers; d_5) distance to the subways; d_6) distance to the expressways; d_7) distance to the roads; d_8) DEM and d_9) slope.

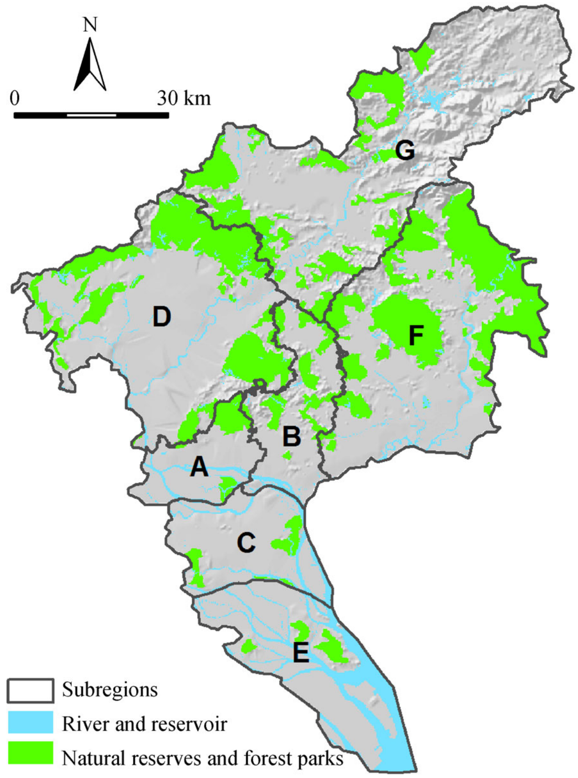

Fig. 3 Ecological sensitive areas and macro zones for urban sprawling: A) Central zone; B) Eastern zone; C) Southern zone; D) Northern zone; E) Nansha district; F) Zengcheng district and G) Conghua district.

Table 1 Urban land demand in very zones (200 m × 200 m cells, 761 rows × 555 columns)

Sub-zone	2005a	2035a
A (Central zone)	5088	6453
B (Eastern zone)	3144	5475
C (Southern zone)	3929	6894
D (Northern zone)	12172	15808
E (Nansha district)	2188	4428
F (Zengcheng district)	4975	6215
G (Conghua district)	1786	2227

total number of 20000 sample points were randomly selected. The coefficient of each factor was estimated using logistic regression models in SPSS software. Table 2 lists the estimation results. Then, we determined the growth probability by using spatial analysis tools in ArcGIS. Fig. 4 shows the spatial pattern of growth probability.

3.3.2 Generating the optimal scenario for urban growth

Based on the estimated growth probability, the parameters of the LR-ACO model should be determined first. The evaporation rate of the pheromone has usually been set as 0.7 in many previous studies (Liu et al., 2012). In this study, the LR-ACO for urban growth optimization was performed with the following given parameters: a total number of 20 ants were designed, the evaporation rate ρ of the pheromone was set as 0.7, and the weight ω for measuring the relative importance of η and τ was set as 0.5. Another important parameter was used to regulate the preference of the growth probability and landscape pattern. If the weight for landscape compactness was set as zero, then the derived pattern was the same as the segmentation result of LR based on the given threshold. In contrast, if a

Table 2 Coefficients of spatial driven factors with logistic regression

Spatial variables	B	Sig.	Exp(B)
(d_1). distance to the municipal centers	−4.305	0.000	0.014
(d_2). distance to the district centers	−4.555	0.000	0.011
(d_3). distance to the large town centers	1.974	0.000	7.198
(d_4). distance to the small town centers	1.610	0.000	5.004
(d_5). distance to the subways	−2.567	0.000	0.077
(d_6). distance to the expressways	−1.266	0.102	0.282
(d_7). distance to the roads	−13.822	0.000	0.000
(d_8). DEM	−4.045	0.030	0.018
(d_9). slope	−7.900	0.000	0.000
constant	0.352	0.011	1.422

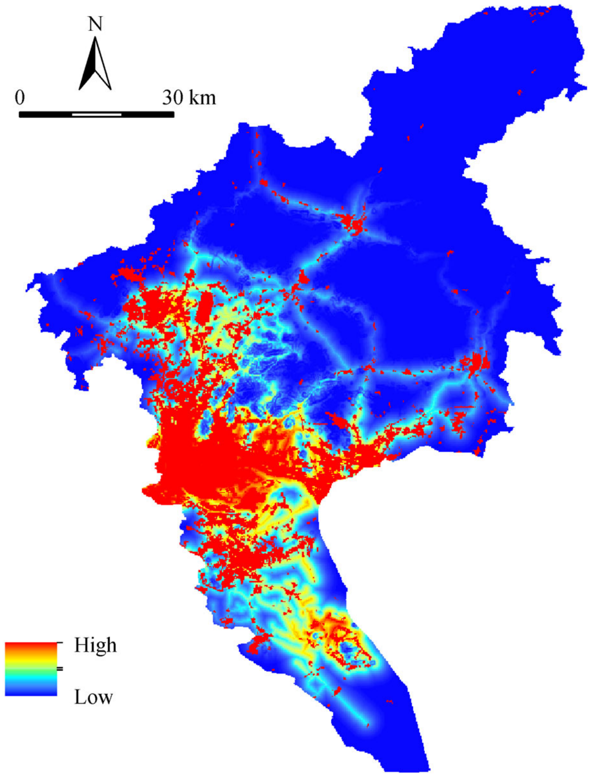

Fig. 4 Land use development probability estimated by LR

higher weight was given for landscape compactness, then the urban patches were closer to regular shapes. Which factor is given the higher weight usually depends on the preferences of the planners. The values of the weight for each factor ranging from 0–1 were set, and various combinations were defined to analyze the sensitivity of the LR-ACO. The weights for the two factors were divided with intervals of 0.2 in this study, and the corresponding feasible weights were 0, 0.2, 0.4, 0.6, 0.8, and 1. Different allocation scenarios were derived under different weight combinations. In addition, the visual comparison indicated that the weight for growth probability ranged from 0.40 to 0.8, and the weight for compactness approximated at 0.40 was rational for application. Therefore, ω_{Devp} and ω_{Comp} were respectively set as 0.6 and 0.4 to generate the optimal growth pattern from LR-ACO. The result is shown in Fig. 5(a).

3.3.3 Result validation with the LR-CA

Traditional simulation models such as LR and CA have also been applied to allocate urban land (Clarke and Gaydos, 1998; Verburg et al., 1999). These two typical models simulate the urban growth pattern in the perspective of a historical dynamic mechanism. This is different from the modeling principle of LR-ACO. To validate the performance of the LR-ACO model, the two models, LR and LR-CA, were performed with the same dataset to make the comparison in this study. Figure 5 shows the simulation results of LR, LR-CA and LR-ACO.

The results were quantitatively validated from the aspects of growth probability and landscape metrics. Growth probability can be used to describe the rationality of urban growth under locational conditions. The total growth probability values of the three results were calculated and are shown in Table 3. In addition, landscape metrics have generally been applied to measure the characteristics of the spatial pattern. A total of 6 landscape indices were selected, including: 1) the number of patches (NP), which is used to measure the fragmentation of urban patches. The lower the value of NP is, the lower the influence of urban growth on the ecological systems is; 2) perimeter-area fractal dimension (PAFRAC), which is used to describe the complexity of urban patches. If the value of PAFRAC is closer to 1, then the shape of the urban patches is more regular. Those patches are commonly preferred by developers and planners; 3) the mean proximity index (PROX_MN), which reflects the fragmentation and proximity of similar patches. The higher the value of PROX_MN is, the greater the expectation for land use planning; 4) the contagion index (CONTIG), which is used to measure the degree of aggregation of urban patches. A lower value reflects that the landscape is more fragmentary; 5) the connectivity index (CONNECT), which

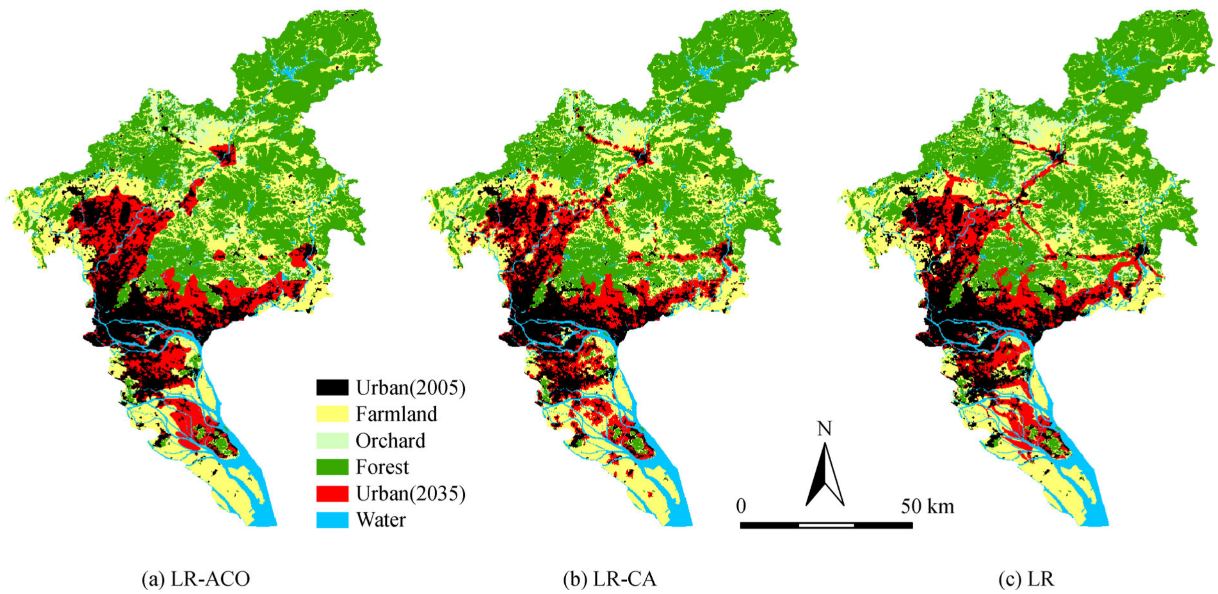

Fig. 5 Urban growth patterns generated from different models.

Table 3 Sprawling probability and compactness analysis

	Total development probability	Landscape metrics					
		NP	PAFRAC	PROX_MN	CONTAG	CONNECT	AI
LR-ACO	43067.052	5135	1.3769	821.9494	41.0035	0.1139	84.7435
LR-CA	41345.3082	5834	1.3795	892.2633	39.7508	0.0987	83.5858
LR	45121.8852	5612	1.3807	804.0474	39.7722	0.1025	83.6004

measures the connectivity among the patches; and 6) the aggregation index (AI). The meaning of this index is similar to that of CONTIG.

The landscape metrics were calculated using FRAGSTATS 4.1 (University of Massachusetts, Amherst). Table 3 lists the values of the landscape metrics for different simulation patterns. From Table 3**,** it can be seen that the urban growth pattern generated from LR is of the highest probability, while the landscape metrics show poorer effects. Although the growth probability of the growth pattern derived from LR-ACO is ranked in the middle, almost all the landscape metrics of this pattern have the best performance except for PROX_MN. This comparison illustrates that the LR-ACO model can generate an optimal urban growth pattern, which presents a balance between the locational probability and the landscape pattern.

Although the quantitative comparison reflects that the global pattern of the simulation result from the LR-ACO model is the most suitable for land use planning, the local pattern and spatial structure of the subregions are also significant aspects of concern by the planners. Five local regions were selected for comparing the differences among the three simulation patterns of urban growth. Figure 6 exhibits that there are large differences among the selected local regions. For the simulation result of the LR-ACO model, the urban patches within regions I and II, which are located in the central city, had more compact and regular distributions than those generated by LR and the LR-CA model. Meanwhile, the selected areas III, IV, and V have been adjusted as the main development directions in the new planning strategy of Guangzhou City. In addition, area III is an important neighborhood in the Pearl River Delta (Fig. 1), and it has also been selected as one of the growth poles. The urban patches in this area are well allocated in the simulation result from LR-ACO and present a connected and compacted pattern, while urban patches exhibited fragmentary and chaotic patterns in the results of LR and LR-CA. In areas IV and V, urban growth is greatly restrained by topographic attributes; the urban patches allocated by LR and LR-CA are mainly scattered along roads. However, a long, striped distribution of urban patches is not feasible for the effective assignment of public services such as water and fitness equipment. Regular and connected patterns of urban patches for these two areas are more likely expected from the view of smart urban growth (Turner, 2007; Abbott and Margheim, 2008). The simulation derived from LR-ACO for these two areas presents this characteristic.

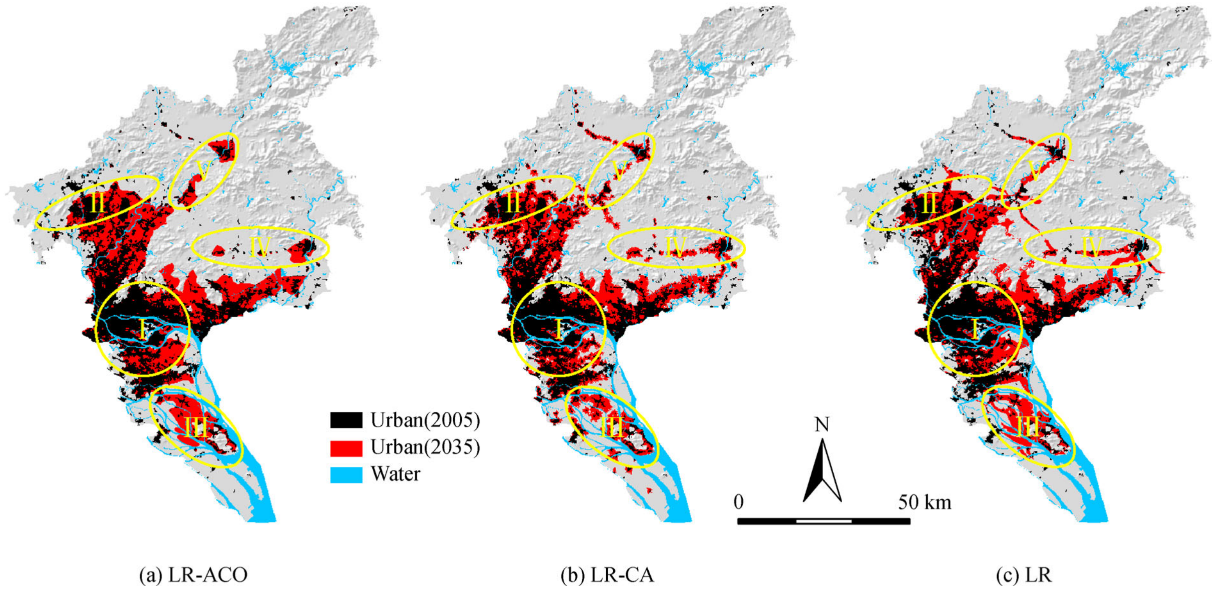

Fig. 6 Comparison of local structures among LR-ACO, LR-CA and LR.

In summary, this visual and quantitative comparison indicated that although this result, especially the local pattern derived from LR-ACO, presented some divergence from the actual planning layout, it was still more efficient in obtaining plausible patterns of urban growth than LR and LR-CA for land use planning.

4 Conclusions

Simulation models, such as LR and CA, have become increasingly popular for allocating land use for urban planning. However, most traditional simulation models assume that urban growth will follow historical trends in the future, which ignores the conflict between urban suitability and landscape metrics. This study integrated LR and the ACO model with the incorporation of historical trends and landscape characteristics to obtain the optimal pattern for urban growth. Therein, urban growth probability is first estimated by using LR. Urban growth probability and landscape metrics are further considered in ACO to generate the optimal scenario for urban planning.

The integrated model is not only discussed in theory but also implemented in the case study area of Guangzhou City, a fast-developing area located in the Pearl River Delta of China. In terms of the planning strategy for the year 2035, the model is performed based on the quantity allocation of each subregion. The two typical simulation models LR and CA are used to generate urban growth patterns with the same dataset to make a comparison. The total urban growth probability and landscape metrics are used to measure the characteristics of the global spatial pattern. In addition, the local visual validation between the planning layout and simulation results is also performed. The visual and quantitative comparisons demonstrate that LR-ACO can generate more plausible urban allocation patterns than LR and CA for land use planning. This coupling method is simple, easy to implement and can provide auxiliary tools for land use planning. However, realistic urban growth is very complex. Other planning factors and constraints should be further incorporated and quantified to improve the availability of the spatial optimization models in urban planning.

Acknowledgements The authors would like to thank the anonymous reviewers for their suggestions and comments. This research was supported by the National Natural Science Foundation of China (Grant No. 41901311).

References

Abbott C, Margheim J (2008). Imagining Portland's urban growth boundary: planning regulation as cultural icon. J Am Plann Assoc, 74 (2): 196–208

Aerts J C J H, Eisinger E, Heuvelink G B M, Stewart T J (2003). Using linear integer programming for multi-site land-use allocation. Geogr Anal, 35(2): 148–169

Aljoufie M, Zuidgeest M, Brussel M, van Vliet J, van Maarseveen M (2013). A cellular automata-based land use and transport interaction model applied to Jeddah, Saudi Arabia. Landsc Urban Plan, 112: 89–99

Cao K, Batty M, Huang B, Liu Y, Yu L, Chen J (2011). Spatial multi-objective land use optimization: extensions to the non-dominated sorting genetic algorithm-II. Int J Geogr Inf Sci, 25(12): 1949–1969

Cao K, Huang B, Wang S, Lin H (2012). Sustainable land use optimization using boundary-based fast genetic algorithm. Comput

Environ Urban Syst, 36(3): 257–269
Cerreta M, De Toro P (2012). Urbanization suitability maps: a dynamic spatial decision support system for sustainable land use. Earth Syst Dynam, 3(2): 157–171
Chen Y M, Li X, Liu X P, Liu Y L (2010). An agent based model for optimal land allocation (Agent LA) with a contiguity constraint. Int J Geogr Inf Sci, 24(8): 1269–1288
Clarke K C, Gaydos L J (1998). Loose-coupling a cellular automata model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. Int J Geogr Inf Sci, 12(7): 699–714
Dorigo M, Maniezzo V, Colorni A (1996). The ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern A Syst Hum, 26(1): 1–13
Feng Y, Liu Y, Tong X, Liu M, Deng S (2011). Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. Landsc Urban Plan, 102(3): 188–196
Godschalk D R (2004). Land use planning challenges: coping with conflicts in visions of sustainable development and livable communities. J Am Plann Assoc, 70(1): 5–13
Gounaridis D, Chorianopoulos I, Koukoulas S (2018). Exploring prospective urban growth trends under different economic outlooks and land-use planning scenarios: the case of Athens. Appl Geogr, 90: 134–144
Hess P M, Sorensen A (2015). Compact, concurrent, and contiguous: smart growth and 50 years of residential planning in the Toronto region. Urban Geogr, 36(1): 127–151
Holzkämper A, Seppelt R (2007). A generic tool for optimizing land-use patterns and landscape structures. Environ Model Softw, 22(12): 1801–1804
Huang K, Liu X, Li X, Liang J, He S (2013). An improved artificial immune system for seeking the Pareto front of land-use allocation problem in large areas. Int J Geogr Inf Sci, 27(5): 922–946
Jantz C A, Goetz S J, Donato D, Claggett P (2010). Designing and implementing a regional urban modeling system using the SLEUTH cellular urban model. Comput Environ Urban Syst, 34(1): 1–16
Jokar Arsanjani J, Helbich M, Kainz W, Darvishi Boloorani A (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. Int J Appl Earth Obs Geoinf, 21: 265–275
Lambin E F, Meyfroidt P (2011). Global land use change, economic globalization, and the looming land scarcity. Proc Natl Acad Sci USA, 108(9): 3465–3472
Lawler J J, Lewis D J, Nelson E, Plantinga A J, Polasky S, Withey J C, Helmers D P, Martinuzzi S, Pennington D, Radeloff V C (2014). Projected land-use change impacts on ecosystem services in the United States. Proc Natl Acad Sci USA, 111(20): 7492–7497
Li F, Gong Y, Cai L, Sun C, Chen Y, Liu Y, Jiang P (2018). Sustainable land-use allocation: a multiobjective particle swarm optimization model and application in Changzhou, China. J Urban Plann Dev, 144 (2): 04018010
Li X, Chen Y, Liu X, Li D, He J (2011). Concepts, methodologies, and tools of an integrated geographical simulation and optimization system. Int J Geogr Inf Sci, 25(4): 633–655
Li X, He J, Liu X (2009). Intelligent GIS for solving high-dimensional site selection problems using ant colony optimization techniques. Int J Geogr Inf Sci, 23(4): 399–416
Li X, Lao C, Liu X, Chen Y (2011). Coupling urban cellular automata with ant colony optimization for zoning protected natural areas under a changing landscape. Int J Geogr Inf Sci, 25(4): 575–593
Li X, Yeh A G (2002). Neural-network-based cellular automata for simulating multiple land use changes using GIS. Int J Geogr Inf Sci, 16(4): 323–343
Ligmann-Zielinska A, Church R, Jankowski P (2008). Spatial optimization as a generative technique for sustainable multi-objective land-use allocation. Int J Geogr Inf Sci, 22(6): 601–622
Liu R, Zhang K, Zhang Z, Borthwick A G L (2014). Land-use suitability analysis for urban development in Beijing. J Environ Manage, 145: 170–179
Liu Y, Wang H, Ji Y, Liu Z, Zhao X (2012). Land use zoning at the county level based on a multi-objective particle swarm optimization algorithm. Int J Environ Res Public Health, 9(8): 2801–2826
Ma S, Li X, Cai Y (2017). Delimiting the urban growth boundaries with a modified ant colony optimization model. Comput Environ Urban Syst, 62: 146–155
Masoomi Z, Mesgari M S, Hamrah M (2013). Allocation of urban land uses by multi-objective particle swarm optimization algorithm. Int J Geogr Inf Sci, 27(3): 542–566
Mitsova D, Shuster W, Wang X (2011). A cellular automata model of land cover change to integrate urban growth with open space conservation. Landsc Urban Plan, 99(2): 141–153
Olsen L M, Dale V H, Foster T (2007). Landscape patterns as indicators of ecological change at Fort Benning, Georgia, USA. Landsc Urban Plan, 79(2): 137–149
Persson C (2013). Deliberation or doctrine? Land use and spatial planning for sustainable development in Sweden. Land Use Policy, 34: 301–313
Klepeis P, Turner II B L (2001). Integrated land history and global change science: the example of the Southern Yucatán Peninsular Region project. Land Use Policy, 18(1): 27–39
Poelmans L, Van Rompaey A (2010). Complexity and performance of urban expansion models. Comput Environ Urban Syst, 34(1): 17–27
Rindfuss R R, Walsh S J, Turner B L, Fox J, Mishra V (2004). Developing a science of land change: challenges and methodological issues. Proc Natl Acad Sci USA, 101(39): 13976–13981
Santé I, García A M, Miranda D, Crecente R (2010). Cellular automata models for the simulation of real-world urban processes: a review and analysis. Landsc Urban Plan, 96(2): 108–122
Santé-Riveira I, Boullón-Magán M, Crecente-Maseda R, Miranda-Barrós D (2008). Algorithm based on simulated annealing for land-use allocation. Comput Geosci, 34(3): 259–268
Silva E A, Ahern J, Wileden J (2008). Strategies for landscape ecology: an application using cellular automata models. Prog Plann, 70(4): 133–177
Stewart T J, Janssen R (2014). A multiobjective GIS-based land use planning algorithm. Comput Environ Urban Syst, 46: 25–34
Stewart T J, Janssen R, van Herwijnen M (2004). A genetic algorithm approach to multiobjective land use planning. Comput Oper Res, 31 (14): 2293–2313
Tudes S, Yigiter N D (2010). Preparation of land use planning model using GIS based on AHP: case study Adana-Turkey. Bull Eng Geol Environ, 69(2): 235–245
Turner M A (2007). A simple theory of smart growth and sprawl. J

Urban Econ, 61(1): 21–44
Verburg P H, de Koning G H J, Kok K, Veldkamp A, Bouma J (1999). A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. Ecol Modell, 116(1): 45–61
Wei Y, Ye X (2014). Urbanization, urban land expansion and environmental change in China. Stochastic Environ Res Risk Assess, 28(4): 757–765
Wu F (2002). Calibration of stochastic cellular automata: the application to rural-urban land conversions. Int J Geogr Inf Sci, 16(8): 795–818
Wu F, Webster C J (1998). Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environ Plann B Plann Des, 25(1): 103–126