International Journal of Applied Earth Observation and Geoinformation 38 (2015) 164–174

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Deriving urban dynamic evolution rules from self-adaptive cellular automata with multi-temporal remote sensing images

Yingqing He[b], Bin Ai[a,*], Yao Yao[c], Fajun Zhong[c]

[a] School of Marine Sciences, Guangdong Key Laboratory for Urbanization and Geo-simulation, Guangdong Provincial Key Laboratory of Marine Resources and Coastal Engineering, Sun Yat-sen University, Guangzhou 510275, Guangdong, China
[b] Pearl River Hydraulic Research Institute, Guangzhou 510275, Guangdong, China
[c] School of Geography and Planning, Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Guangzhou 510275, Guangdong, China

ARTICLE INFO

Article history:
Received 9 September 2013
Accepted 22 December 2014
Available online 21 January 2015

Keywords:
Artificial immune system
Cellular automata
Urban dynamic simulation
Self-adaptive
Multi-temporal remote sensing images

ABSTRACT

Cellular automata (CA) have proven to be very effective for simulating and predicting the spatio-temporal evolution of complex geographical phenomena. Traditional methods generally pose problems in determining the structure and parameters of CA for a large, complex region or a long-term simulation. This study presents a self-adaptive CA model integrated with an artificial immune system to discover dynamic transition rules automatically. The model's parameters are allowed to be self-modified with the application of multi-temporal remote sensing images: that is, the CA can adapt itself to the changed and complex environment. Therefore, urban dynamic evolution rules over time can be efficiently retrieved by using this integrated model. The proposed AIS-based CA model was then used to simulate the rural-urban land conversion of Guangzhou city, located in the core of China's Pearl River Delta. The initial urban land was directly classified from TM satellite image in the year 1990. Urban land in the years 1995, 2000, 2005, 2009 and 2012 was correspondingly used as the observed data to calibrate the model's parameters. With the quantitative index figure of merit (FoM) and pattern similarity, the comparison was further performed between the AIS-based model and a Logistic CA model. The results indicate that the AIS-based CA model can perform better and with higher precision in simulating urban evolution, and the simulated spatial pattern is closer to the actual development situation.

Introduction

Many geographical phenomena contain both spatial and temporal features, for example, urban growth, disease propagation, fire diffusion, population migration, flood inundation, etc., hence, their spatio-temporal dynamic development processes appear to be more important than the finally resultant spatial patterns (Liu et al., 2008). Additionally, spatial evolution patterns of regional geography phenomena should be considered when simulating and predicting the variation of global resources, environment and atmosphere. However, conventional models generally pose the problem of process analysis due to the lack of dynamic information, which cannot satisfy the growing requirements proposed for the spatio-temporal simulation (Goodchild, 1992; Batty, 1993). As a result, it is useful to introduce dynamic models integrated with spatio-temporal information for simulating complex geographical systems. As a microscopic model following a "bottom-up" approach, cellular automata (CA) have proven efficient for simulating the dynamic evolution processes of complex geographical systems, which can provide spatial information for global models of resources, environment and atmosphere.

The theoretical framework for CA's application in geosciences was perhaps first introduced in detail by Couclelis (1985, 1988, 1989). A variety of experiments were accordingly conducted for simulating urban evolution by using CA models. It has demonstrated that CA can generate complex spatial patterns with only simple conversion rules through the simulation of a virtual city (Couclelis, 1985). Later, White and Engelen (1993) developed a CA model to investigate the fractal properties of cities and urban evolution. Clarke and Gaydos (1998) utilized CA models to simulate the evolution process of real cities, selecting San Francisco

* Corresponding author at: School of Marine Sciences, Guangdong Key Laboratory for Urbanization and Geo-simulation, Guangdong Provincial Key Laboratory of Marine Resources and Coastal Engineering, Sun Yat-sen University, 135 West Xingang Rd., Guangzhou 510275, China. Tel.: +86 20 84115833; fax: +86 20 84115833.
E-mail addresses: heyingqing@foxmail.com (Y. He), abin@mail.sysu.edu.cn (B. Ai), whuyao@foxmail.com (Y. Yao), areszhong@qq.com (F. Zhong).

http://dx.doi.org/10.1016/j.jag.2014.12.014

and Washington as the sample areas, and the model's parameters were further calibrated using historical digital maps. Li and Yeh (2000) applied a constrained CA model to the urban sprawl simulation of Dongguan city in China's Guangdong Province. Most related researches have assumed that the historical trend of urban growth will continue into the future.

It has also shown that the simulation of a virtual city with a CA model does not require detailed and realistic data, and the model's parameters accordingly do not require calibration (White and Engelen, 1993). However, it is very essential to calibrate the CA's parameters and retrieve reasonable transition rules for the urban evolution simulation of real cities. These rules are represented by the parameters which are associated with the spatial variables involved in the simulation (Liu et al., 2008). Usually, historic data are used to derive these parameters according to calibration procedures. Various approaches have been explored to solve this problem, and most of the existing approaches assume that transition rules are static. The same set of transition rules will be applied to any location with the time variation, regardless of the possible changes to the simulation environment. This is not in agreement with the complexity of geographical phenomena.

As for urban growth, it does not follow a fixed model for the reason that spatial patterns and driving forces are dynamic at different stages. That is, not all the urban growth in different periods follows the same historic evolution rule. It is also unreasonable to use the same set of rules when a study area is large. Therefore, it is difficult for conventional approaches with only static rules to describe the probably varied relationships in the spatio-temporal dimension. Some researchers have attempted to introduce auxiliary techniques for discovering dynamic transition rules of CA. For instance, Clarke et al. (1997) once proposed a Monte Carlo method to solve this problem. However, the Monte Carlo model requires a long computation time, as the possible combinations between many parameters are numerous, so this model is only suitable for simulating urban evolution of a medium or small region. Also, Liu et al. (2014) first imposed the urban growth theory into a CA model to retrieve different types of transition rules successfully, which can disclose the main patterns of urban growth and reflect the actual evolution process. The transition rules still cannot adequately describe the dynamic characters of urban growth, especially for a long-term simulation. Techniques such as particle swarm optimization, ant colony optimization, Markov model and gravitational field model have been also applied to the solution of such problems (Feng et al., 2011; Yang et al., 2012; Guan et al., 2011; He et al., 2013). However, for large, complex regions or too long a period of simulation, it is necessary to introduce artificial intelligence methods to derive dynamic transition rules. That is, the transition rules should be self-adaptive, capable of memorizing and vary dynamically with the environment variation, from which urban dynamic evolution rules can be discovered. Liu et al. (2008) first attempted to combine the artificial intelligence techniques with the CA model to retrieve adaptive transition rules, which were generally applied in the urban evolution simulation to provide auxiliary information for decision-making. Other artificial intelligence methods have been increasingly incorporated into urban CA models, including artificial neural networks (ANNs) (Li and Yeh, 2002), kernel-based learning machines (Liu et al., 2007), genetic algorithm (Li et al., 2008) and artificial immune system (AIS) (Liu et al., 2010), among others.

Inspired by natural biological systems, AIS can be defined as intelligent computation systems. Currently, AIS has been widely applied to the problem solving such as pattern recognition (Carter, 2000), intelligent optimization (Chun et al., 1997; Liu et al., 2011), machine learning (Timmis, 2000; Timmis and Neal, 2001), adaptive control (Kumark and Neidhoefer, 1997), fault detection (Dasgupta and Forrest, 1995), remote sensing classification (Zhong et al., 2006; Zhong and Zhang, 2012), and land use allocation (Huang et al., 2013). AIS usually mimic the form and function of biological antibodies in learning and memorizing new information. It can recall previously learned information and perform pattern recognition in a highly decentralized fashion. Liu et al. (2010) first applied an AIS-based CA model in the analysis of urban planning for a large area, and the results indicate that it is promising for an AIS-based CA model to solve complex geographical problems with the features of self-adaptation, self-learning, and memorizing. In this study, an urban CA model coupled with an artificial immune system (AIS) was introduced to derive dynamic transition rules automatically. The self-adaptive CA model was then applied to the urban growth simulation of Guangzhou city, located in the core area of China's Pearl River Delta, from the year 1990–2012. The classical Logistic CA was also implemented with the same set of data to make a comparison. Preliminary results suggest the proposed model can perform better than the Logistic CA when simulating the urban evolution of the large region.

AIS-based geographical cellular automata

Basic principles of artificial immune system

The AIS algorithm is mainly based on the idea of biological immune systems. Biological immune systems can be viewed as parallel, self-adaptive, self-learning, self-organizing and distributed systems with the capability to control complex systems through variation over time (Kim and Bentley, 1999). Inspired by theoretical immunology and observed immune system functions, AIS emerged as a computational intelligence technique for solving complex problems (Tarakanov and Dasgupta, 2000). An AIS implements an information memorizing and processing mechanism similar to that of a biological immune system. It has been widely applied since the late 1980s.

Generally, a biological immune system consists of cells, molecules, and organs that aim to protect the biological body against infection. The defence mechanism is based on the adaptive immune responses, in which antibodies evolve to obtain stronger capabilities of dealing with certain antigens (De Castro and Timmis, 2002; Liu et al., 2010). The basic algorithm of immune systems can be regarded as clonal selection, which is commonly used for describing the properties of an adaptive immune response to an antigenic stimulus (De Casro and Von Zuben, 2000). The principle of clonal selection was first proposed by Jerne (1973), and it is deemed that only those cells, that is, antibodies with high affinities to certain antigens are selected to proliferate; otherwise, the cells will be eliminated. These selected cells can easily recognize antigens and are subject to an affinity maturation process. The process of improving the affinity with the selected antigens is also called the maturing process of antibodies.

Based on the clonal selection, an AIS imitates a biological system by adapting a set of 'antibodies' through responding to a set of known 'antigens', and then by using 'mature antibodies' to process unknown 'antigens' (Liu et al., 2010). The known 'antigens' are usually used to build the training set, in which each 'antigen' contains both the problem description and its state. The 'antibodies' form a problem solver (e.g., a classifier) and their adaptations or evolutions are essentially the training process. Therefore, the number of 'antibodies' can be much smaller than the number of known 'antigens'. Specifically in an urban CA model, 'antigens' are the cells to be classified in the simulation and 'antibodies' are the classifiers that will assign urban land to cells based on their features. Once the 'antibodies' are 'mature', that is, the problem solver is trained, they can be used to solve new problems represented by unknown 'antigens'. So basically this is a way to derive generalized 'rules'.

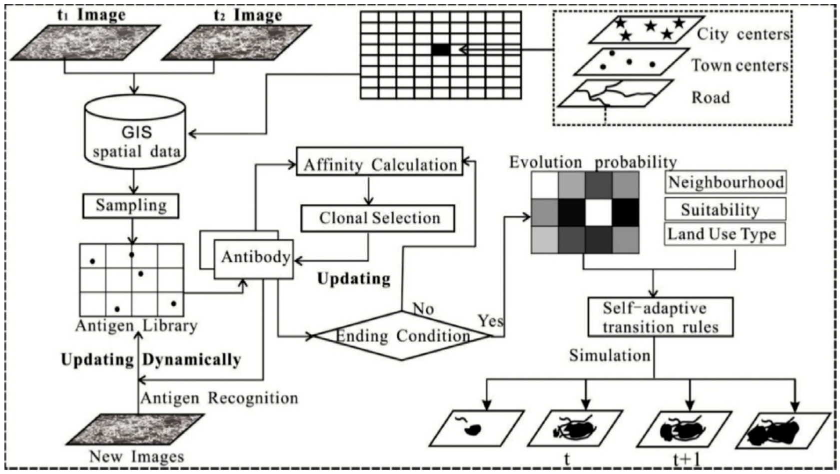

Fig. 1. Discovery of self-adaptive transition rules with the AIS-based CA.

Artificial immune system and geographical cellular automata

When an AIS is used to solve actual geography problems, it is necessary to correlate problem descriptions with the theory of the immune system, define mathematical representations of immune elements, and design corresponding algorithms. It can be taken as a typical recycling process, including antigen definition, initial antibody generation, affinity calculation, clonal selection, antibody mutation and antibody updating. This process will be used to calculate the evolution probability of an urban CA model.

Considering the key processing of AIS, the main procedure of the AIS-based CA model can be correspondingly described as Fig. 1. First, the initial antigen library is generated with the remotely sensed imageries collected at the times of t_1 and t_2, respectively. The antigen library can be repeatedly updated when additional remotely sensed imageries are available for retrieving the possible change in urban evolution. Then new antigens will be automatically recognized when the simulation environment is varied. The updated process can reflect the self-adaptive and self-learning capability of the AIS-based CA to a certain degree. Detailed procedures of AIS-based CA are provided in the following sections.

Definition of the antigen in geographical cellular automata

As for an AIS-based CA model, the first step is to establish a reasonable antigen library, which will be used to replace explicit transition functions or rules derived from traditional CA. In an AIS-based CA model, each antigen is of course represented as a vector consisting of two parts: attributes (problem description) and the state conversion (solution). The attributes are composed of land use type and proximity variables such as distance to road, distance to city center, distance to railway, etc. An antigen can then be represented as follows:

$$\mathrm{Ag} = (a_1(D), a_2(D), \ldots, a_N(D); S) \tag{1}$$

where $a_1(D), a_2(D), \ldots, a_N(D)$ are a series of variables corresponding to the antigen Ag, including proximity distances and land use type. S is a Boolean variable that uses 1 to represent the state of being urbanized and 0 for otherwise. To make the attributes comparable, all the variables are normalized within the range of 0–1.

Antibody initialization and affinity calculation

As mentioned before, attributes of an antibody can be commonly initialized with a series of random numbers:

$$\mathrm{Ab} = (a_1(R), a_2(R), \ldots, a_N(R); S) \tag{2}$$

where $a_1(R), a_2(R), \ldots, a_N(R)$ represent the features of an antibody Ab corresponding to spatial variables and proximity variables. Each attribute in Eq. (2) is initially set to be a random number ranging from 0 to 1.

In an artificial immune system, the affinity is used to represent the binding degree between an antibody and an antigen. Any antibody with higher affinity to an antigen is more likely to be selected, retained and cloned. The affinity between an antibody and an antigen can be calculated using the following equations:

$$\mathrm{Af}(\mathrm{Ag}, \mathrm{Ab}) = \frac{1}{1 + d(\mathrm{Ag}, \mathrm{Ab})} \tag{3}$$

$$d(\mathrm{Ag}, \mathrm{Ab}) = \sqrt{\sum_{n=1}^{N} (a_n(D) - a_n(R))^2} \tag{4}$$

where Af(Ag, Ab) is described as affinity between the antigen Ag and the antibody Ab varying from 0 to 1, and d(Ag,Ab) represents their similarity. In this study, the similarity is calculated using a Euclidean distance, expressed as Eq. (4).

Clonal selection and mutation

Antibodies with higher affinity to antigens are selected to constitute an antibody library. The cloning probability of an antibody mainly depends on its affinity and concentration. The concentration of the antibody $g(L_g)$ can be calculated using the following equations:

$$L_g = \sum_{h=1}^{H} C_h \tag{5}$$

$$C_h = \begin{cases} 1, \mathrm{Af}_{gh} \geq T \\ 0, \mathrm{Af}_{gh} < T \end{cases} \tag{6}$$

where Af_{gh} is defined as the affinity between the antibody g and the antigen h, T is set as the threshold value for affinity. Then the cloning rate of the antibody g can be represented as follows:

$$P_g = \frac{a \times \mathrm{Af}_g}{L_g} \tag{7}$$

where P_g is described as the cloning rate, Af_g refers to the affinity of the antibody g, and a is a constant. For each antibody, the higher its affinity, the higher cloning probability it will have. Meanwhile, antibodies with lower concentration will have higher cloning probability, while antibodies with higher concentration will have lower cloning probability. This cloning mechanism not only protects antibodies with higher quality, but also facilitates antibodies with lower concentration development.

After the cloning process, the antibodies are mutated to increase their diversity. The mutation degree for an antibody is inversely proportional to its affinity, that is, the higher the affinity is, the lower the mutation degree will be. The mutation process can be represented and realized in multiple forms. In this study, the mutation of an antibody is to be performed using the following equation:

$$\mathrm{Ab}_v = \mathrm{Ab} - (1 - e^{-d(\mathrm{Ag},\mathrm{Ab})})(\mathrm{Ab} - \mathrm{Ag}) \tag{8}$$

where Ab_v refers to the mutation antibody. As is shown in Eq. (8), an antibody with higher affinity will have a lower mutation rate. After mutation, some antibodies show increased affinity while other antibodies show decreased affinity. Those with the latter will tend to be eliminated. As a result, the affinity of any antibody would be gradually increased and an antibody would mature through the process of clonal selection and mutation.

Affinity attenuation

Meanwhile, new available urban land data would be regarded as new antigens and supplemented into the antigen library. In this study, new urban land data will be generated with available remote sensing images. Then, corresponding new antibodies will be automatically derived from AIS to recognize these new antigens. The affinity of old antibodies will gradually decrease over time. When affinity attenuation reaches to the given threshold value, old antibodies would be possibly eliminated. Affinity attenuation is defined as follows:

$$\mathrm{Af}_{t+1} = \beta \times \mathrm{Af}_t \tag{9}$$

where Af_t and Af_{t+1} are the affinity of an antibody at time t and $t+1$, respectively, and β represents the attenuation coefficient to control the increasing rate to affinity, which is significant to the simulation of dynamic systems. Depending on the adjustment of the affinity for the dynamic CA model, new antibodies will be efficiently protected in one way while old antibodies will be restrained and even eliminated in the other way. All the old antibodies would be totally eliminated in terms of the attenuation mechanism with increasing time. Therefore, the antibody library would not result in data explosion, owing to gradual increasing in the set of antibodies over time. And this can also adapt the evolution mechanism to the changed environment and retrieve the dynamic evolution rules of urban growth.

Using AIS to discover transition rules of CA

Features of an antigen can be directly derived from spatial data and remotely sensed imageries to create an antigen library. The initial antibodies are generated with the values of random numbers. These antibodies are inclined to be 'matured' by means of clonal selection and mutation. This can be taken as the training stage in many other machine-learning methods. The mature antibodies are matched with some antigens, which can identify the antigens' structure through memorization. These antibodies can then be used to recognize the queried cells composed of various variables, such as spatial distance and land use type. The affinity between an antibody and the queried cells can be calculated using Eq. (3), and the antibodies with best affinity are used to determine the class (state) of the cells. Following the conventional method, i.e., elitist selection, the cell is assigned with the majority class (state) of its k best antibodies. However, this method can yield only a Boolean value – converted or not. As urban evolution cannot be exactly forecasted, it would be necessary to introduce fuzzy concepts into the recognition of urban growth. So the conversion probability is used to produce more plausible simulation results. In other words, a roulette wheel algorithm (Lipowski and Lipowska, 2012) is used to calculate the probability of urban growth in the following way:

$$P_{\mathrm{dev}}(\mathrm{ij}) = \frac{\sum_{u=1}^{k} \mathrm{Af}(\mathrm{Ab}_u, x_t) \times \delta(f(\mathrm{Ab}_u), 1)}{\sum_{u=1}^{k} \mathrm{Af}(\mathrm{Ab}_u, x_t) \times \delta(f(\mathrm{Ab}_u), 1) + \sum_{u=1}^{k} \mathrm{Af}(\mathrm{Ab}_u, x_t) \times \delta(f(\mathrm{Ab}_u), 0)} \tag{10}$$

$$\delta(f(\mathrm{Ab}_u), 1) \begin{cases} = 1, \text{ iff}(\mathrm{Ab}_u) = 1 \\ = 0, \text{ iff}(\mathrm{Ab}_u) = 0 \end{cases} \tag{11}$$

$$\delta(f(\mathrm{Ab}_u), 0) \begin{cases} = 1, \text{ iff}(\mathrm{Ab}_u) = 0 \\ = 0, \text{ iff}(\mathrm{Ab}_u) = 1 \end{cases} \tag{12}$$

where $\mathrm{Af}(\mathrm{Ab}_u, x_t)$ denotes the affinity between the antibody Ab_u and the queried cell x_t, the contribution of an antibody is determined by its affinity to the queried cell, and $f(\mathrm{Ab}_u)$ is the class (state) of the antibody Ab_u.

The final conversion probability is mainly determined by the growth probability, neighborhood state and other constraint factors. This can be described as follows:

$$P_t(\mathrm{ij}) = A \times P_{\mathrm{dev}}(\mathrm{ij}) \times \mathrm{con}(\mathrm{suit}(\mathrm{ij})) \times \Omega_{t-1}(\mathrm{ij}) \tag{13}$$

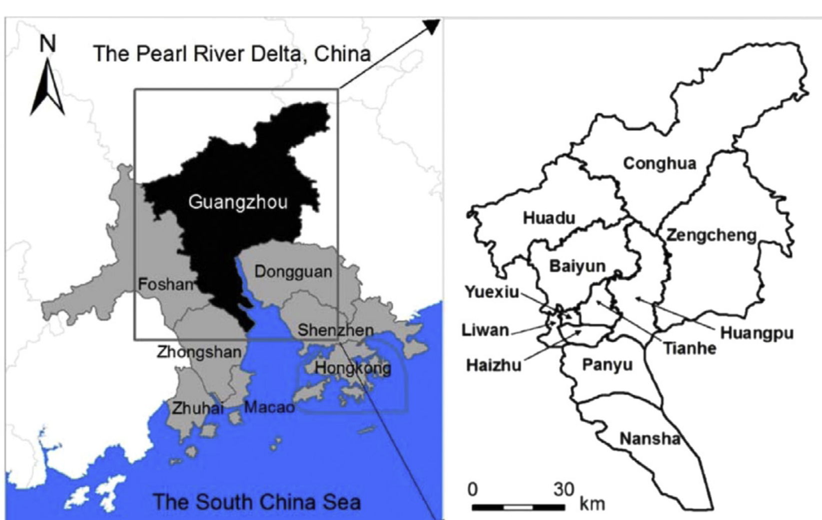

Fig. 2. Location and administrative districts of the study area Guangzhou city.

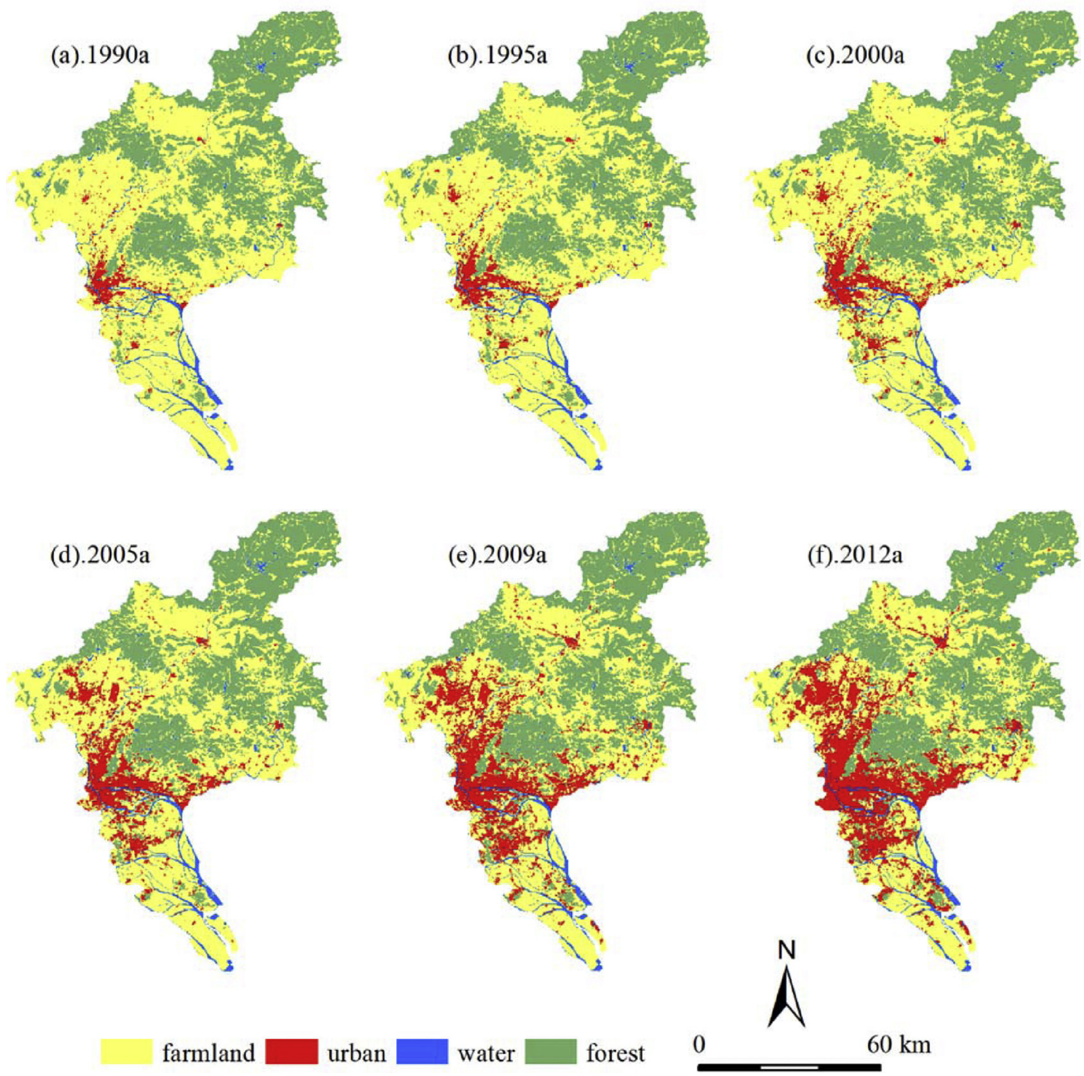

Fig. 3. Urban evolution process of Guangzhou city during the period 1990–2012.

where A is defined as an adjusting factor, con(suit(ij)) refers to the combined constrain ranging from 0 to 1, and $\Omega_{t-1}(\mathrm{ij})$ is the percentage of urbanized cells in a 3×3 neighborhood window:

$$\Omega_{t-1}(\mathrm{ij}) = \frac{\sum\limits_{3\times 3} N(\mathrm{urban(ij)})}{3 \times 3 - 1} \tag{14}$$

Implementation and results

The AIS-based CA model was then applied to the urban evolution simulation of Guangzhou city (longitudes 112°57′–114°03′E, latitudes 22°26′–23°56′N), a fast-growing region located in the core area of China's Pearl River Delta shown as Fig. 2. This city includes the following eleven districts, with a total area of 7434 square kilometers and a total population of 1292.68 million in 2013: Tianhe, Yuexiu, Liwan, Haizhu, Baiyun, Nansha, Huangpu, Panyu, Huadu, Zengcheng and Conghua. It is expected that the AIS-based CA model can provide a better understanding of the historical urban dynamics and aid urban planners in exploring future development alternatives for Guangzhou city.

Actual urban land was directly derived from the classification of Thematic Mapper satellite images (TM 122-44) in the years 1990, 1995, 2000, 2005, 2009 and 2012 with the pixel size of 30 m × 30 m. The land use types include natural water, fishpond, forest, farmland, built-up area, and bare land. The classification was implemented using the technique of object-based classification (Definiens Developer 7.0, 2003). The land use types of fishpond, farmland, and bare land were aggregated as 'non-urban' types, while natural water was reclassified as 'restricted area' since no growth was allowed during simulation. In addition, the land use type of forest was also considered as a 'restricted area' because most of the forests are located on mountains and protected by the planning bureau of Guangzhou city. Then the actual evolution process of Guangzhou in the period 1990–2012 was obtained, as shown in Fig. 3. The actual land use data reveals a quite rapid urbanization process in Guangzhou city during the period 1990–2012. In this period, the amount of new urbanized area is about 1121.32 km^2, accounting for 86.19% of the total built-up area in the year 2012. Further, the urbanization rate and spatial urbanization distribution show great differences among all the simulation periods.

Initialization and correction of the AIS-based CA model

To constitute the antigen library, partial urban areas derived from the reclassification maps of the years 1990 and 1995 were selected as initial antigens using the sampling function of remote sensing software ERDAS 8.7. As the initial antigen library is very significant to the formation of an AIS-based model, a total of 20% of sample points (including both urban and non-urban cells) were randomly selected from the classification images for building the training set, and spatial coordinates of those sample points were also retrieved. Then a series of spatial variables related to urban growth were chosen to assign the attributes for these antigens. Spatial variables mainly include land use type, topographic data and proximate indices (e.g., distance to roads, distance to city centers, distance to railways, distance to expressways, distance to district centers, distance to state highways, distance to provincial highways, etc.). All these spatial variables were prepared using a raster database with a 30-m resolution. And the corresponding training set required for the CA model and the initial antigen library were generated by using the sample function in ArcGIS. During the computation process, it was also necessary to set some parameters for AIS algorithm, so a and β in Eqs. (7) and (9) were assigned to be 2 and 0.9, respectively, in this study based on the testing experiments, which were used to control the rate of clonal selection and attenuation coefficient. Thereafter, initial antibodies would be generated randomly with the TM data collected in the years 1990 and 1995 using AIS, and its affinity would be inclined to mature gradually after cloning and mutation. Mature antibodies were applied to the identification of the urban evolution, and growth probability was obtained according to Eq. (13).

On the other hand, multi-temporal historical data were used to calibrate the CA model for deriving dynamic transition rules. Based on the evolution probability initialized with the historical data in the years 1990 and 1995, the antigen library was gradually updated to reflect the possible change with new temporal TM images respectively collected in the years 1995, 2000, 2005, 2009 and 2012. New antigens were accordingly derived from those images, also based on the classification, and corresponding new antibodies were automatically generated to recognize those new antigens. Transition rules for urban growth were derived from the corresponding antibody library during the simulation period. In this way, the antibody library would be updated dynamically, reflecting the self-adaptive and self-learning capabilities of an AIS-based CA model. With the increasing antibodies in the antibody library, the system would accumulate more and more experiences and knowledge on urban evolution. In addition, the AIS-based CA is capable of memorizing and identifying antigen structure, so when the same antigen resurfaces, the antibody would react in a more violent manner.

Results of the AIS-based CA model

The urban evolution probability of Guangzhou city was derived with the AIS-based CA in the periods 1990–1995, 1995–2000, 2000–2005, 2005–2009 and 2009–2012. It illustrated that the simulation model for Guangzhou city is heterogeneous at different stages of urban growth. Additionally, urban evolution generally shows the dynamic pattern at different stages. For example, the urban land consumption shows dissimilar patterns in different periods. It has proven that static transition rules cannot adequately explain this. However, the AIS-based CA model can be used to extract self-adaptive transition rules owing to the self-adaptive, self-learning and memorizing capabilities. With the application of the AIS-based CA model, it also illustrated that urban land consumption in the study area varied dynamically during the period of 1995–2012 (Fig. 4). Urban land consumption showed relative equal between the periods 1990–1995 and 1995–2000 and showed an increasing trend in the two periods 2000–2005 and 2005–2009, but showed a slight decreasing trend in the period 2009–2012. This conforms to the actual urban land consumption conditions. The self-adaptive transition rules thus can better reveal the heterogeneity of urban growth patterns in different periods and better reveal the dynamic changes of urban land consumption at different stages, from which the dynamic urban evolution rules can be efficiently discovered.

With the dynamic transition rules obtained above, the urban expansion of Guangzhou city in the period 1990–2012 was then simulated. At the first stage, the initial urban land was acquired from classification of the TM image in the year 1990. Because CA are stochastic models with uncertainties, the AIS-based CA was run about 100 times for the period 1990–1995 to validate its simulation accuracies. The urban land in the years 2000, 2005, 2009, and 2012 was then simulated with 200 iterations, 300 iterations, 400 iterations and 500 iterations, respectively. Fig. 5 shows the simulation results. The same training set was also used and a similar procedure was also implemented for the Logistic CA. Fig. 6 shows the corresponding simulation results for the comparison analysis.

Model validation

A city can be viewed as a complex system for the interference of numerous uncertain factors. It is difficult to simulate its dynamic evolution accurately, and only the spatial pattern close to realistic urban growth can be achieved with the simulation techniques.

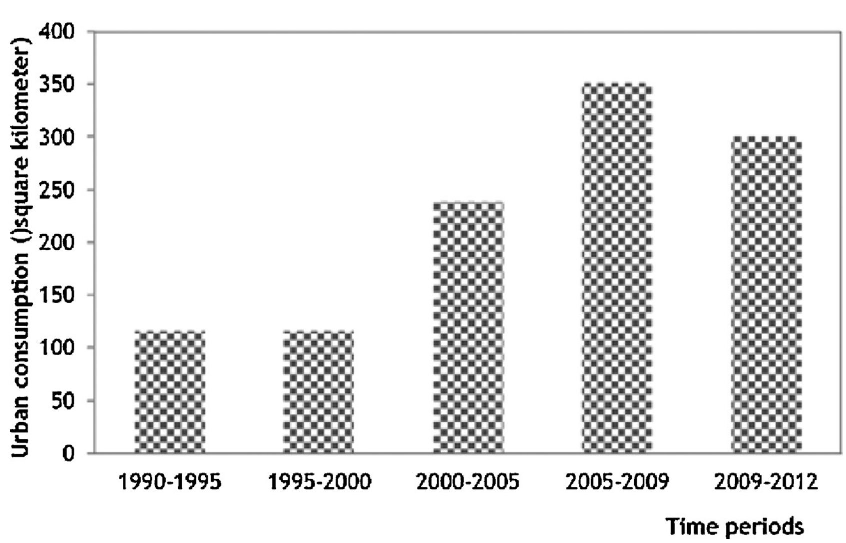

Fig. 4. Urban land consumption of Guangzhou city during different periods.

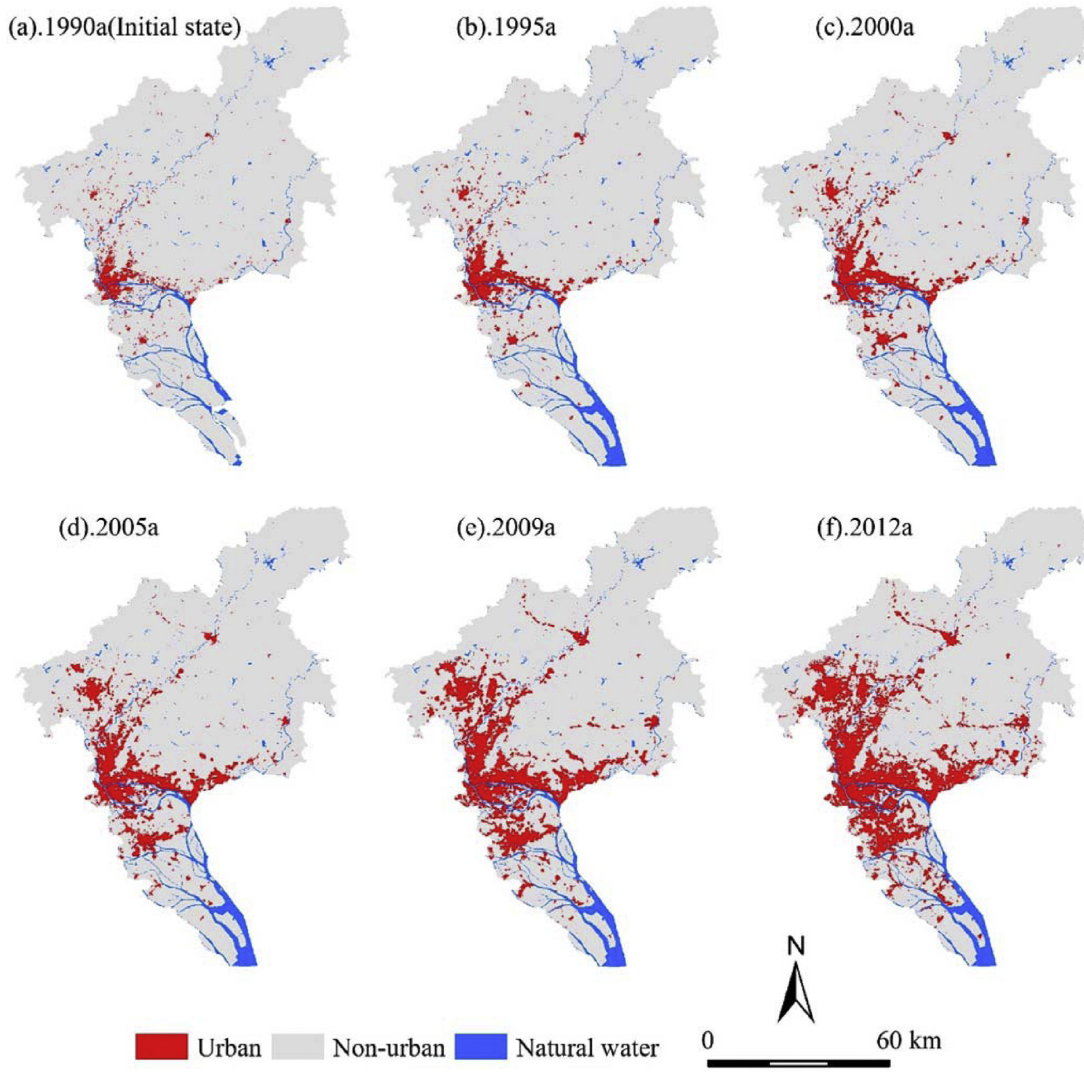

Fig. 5. Simulation results of Guangzhou city during the period 1990–2012 with the AIS-based CA.

From the visual comparison between Figs. 3 and 5, it shows that the overall spatial patterns of different periods derived from the AIS-CA model are closer to the actual situations than those derived from the Logistic CA. When a CA model is applied to the evolution simulation of a real city, it is necessary to validate the conformity between the simulation results and the actual situations. The performance of the AIS-based CA model was further evaluated quantitatively in this study from two aspects: cell-level agreement (point to point analysis) and pattern-level similarity (landscape aspect).

For cell-level aspect, the indicator of 'figure of merit' (FoM) (Pontius et al., 2007) has been commonly used by many researchers. The indicator of 'FoM' is actually a ratio, where the numerator is the number of instances that are observed developed and correctly simulated as developed, while the denominator is the total number of instances excluding persistently non-changed instances. It is calculated using the following equation (Pontius et al., 2008):

$$F = \frac{B}{(A + B + C + D)} \times 100\% \qquad (15)$$

where F is the 'FoM', A represents the error due to observed developed and simulated as persistence, B is the agreement due to observed developed and simulated as developed, C is defined as the error due to observed developed and simulated as incorrect gaining category, and D is the error due to observed persistence and simulated as developed. As the CA models only simulate the change of states from non-urban to urban, the value of C should be equal to 0 (Chen et al., 2014). A higher value of 'FoM' indicates a higher cell-level agreement. The observed development patterns were overlaid with the respective results of the AIS-based CA and the Logistic-based CA to identify four groups of cells (i.e., persistent non-change, observed non-change simulated change, observed change simulated non-change, and observed change simulated change) for calculating 'FoM'. Figs. 7 and 8 show the overlay pattern of each simulation period. Table 1 lists the comparison results.

According to the quantitative accuracy evaluation by using this index for land use development evaluation, the values are basically within the range of about 1.00–25.00% (Pontius et al., 2008). From Table 1, it can be found that the 'FoM' shows the value of 22.20%

Table 1
Comparison of 'FoM' between the AIS-CA and the Logistic CA.

	1995	2000	2005	2009	2012
AIS-CA	0.287725	0.222157	0.286999	0.26029	0.247653
Logistic CA	0.222011	0.173942	0.226134	0.221867	0.179603

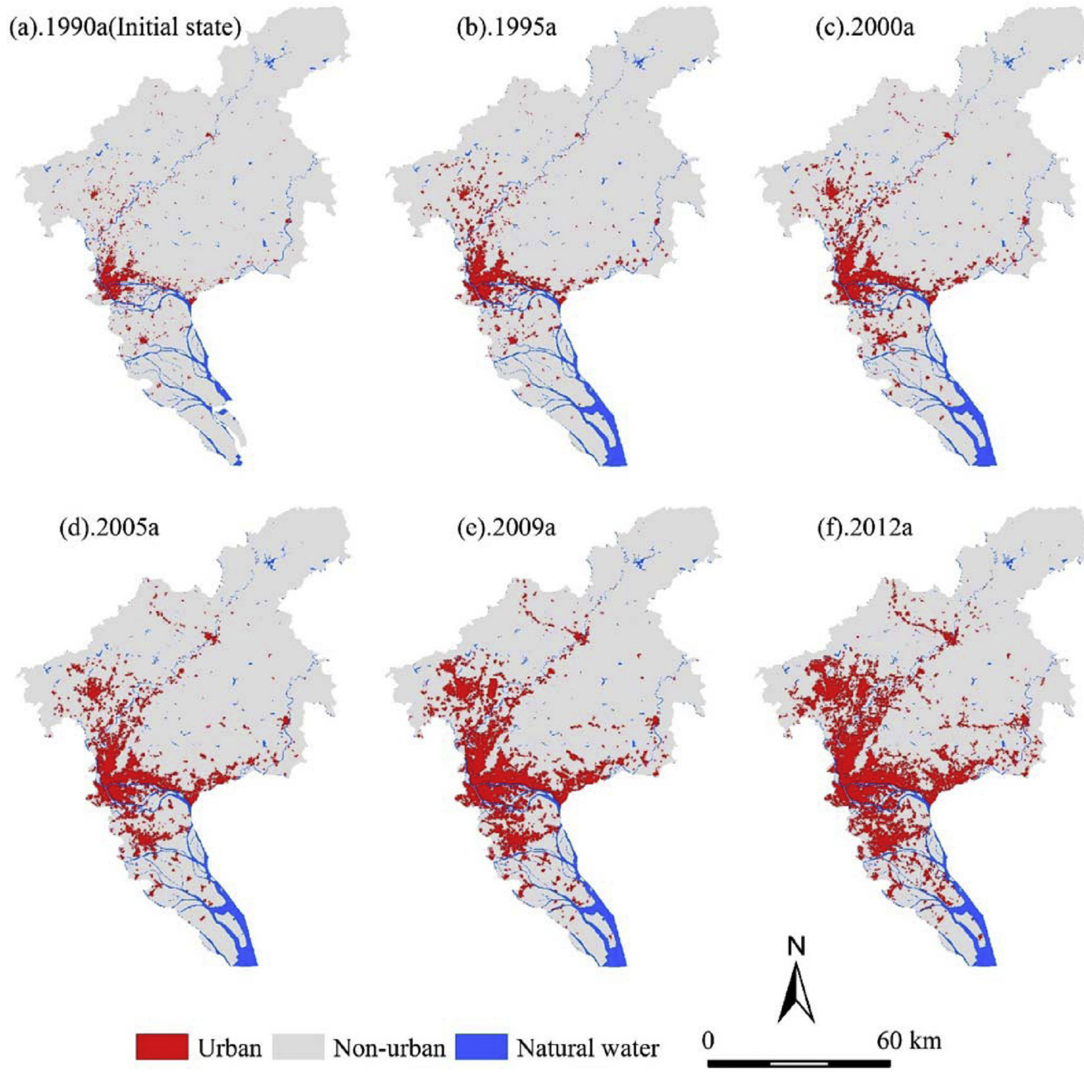

Fig. 6. Simulation results of Guangzhou city during the period 1990–2012 with the Logistic CA.

(1995), 17.39% (2000), 22.61% (2005), 22.19% (2009) and 17.96% (2012), respectively, with a mean value of 20.47% for the simulation results of the Logistic CA. For the results of the AIS-based CA, the 'FoM' shows the value of 28.77% (1995), 22.22% (2000), 28.70% (2005), 26.03% (2009) and 24.77% (2012), respectively, with a mean value of 26.10% (5.63% higher than that of the Logistic CA).

In addition, the spatial pattern simulated by using the AIS-CA model was also compared to determine whether it was consistent with the actual condition. Measures such as landscape metrics have been adopted to validate simulation models in the view of aggregate pattern similarity (Sui and Zeng, 2001; Parker and Meretsky, 2004; Liu et al., 2010). Four landscape metrics selected to delineate the development patterns from different aspects (Dietzel et al., 2005; Seto and Fragkias, 2005) are as follows: (1) number of urban patches (NP) and largest-patch index (LPI), which are usually used to measure the patch size; (2) mean Euclidean nearest-neighbor distance (ENN), which measures the distribution of patches; (3) mean perimeter-area ratio (PARA), which can reflect the shape complexity of the patches. The landscape metrics were calculated using FRAGSTATS 4.1 (University of Massachusetts, Amherst) (McGarigal et al., 2012). Then the pattern-level similarity was estimated through the comparison between the simulated and observed landscape metrics for different models, which was calculated using the following equations (Chen et al., 2014):

$$a_l = 1 - \frac{1}{8}\sum_i \Delta l_i \tag{16}$$

$$\Delta l_i = \begin{cases} |l_{i,s} - l_{i,o}|/l_{i,o} \times 100\%, l = \mathrm{NP, ENN, PARA} \\ |l_{i,s} - l_{i,o}|, l = \mathrm{LPI} \end{cases} \tag{17}$$

where $l_{i,s}$ and $l_{i,o}$ are the values of the ith landscape metrics derived from the simulated pattern and the observed pattern, respectively; l_i is the normalized difference of the ith pair of simulated and observed landscape metrics. The l_i for LPI is calculated as the absolute differences because the original units of LPI are already percentages. Similarities of the AIS-based CA and the Logistic CA for different simulation periods are all listed in Table 2.

As shown in Table 2, the pattern-level similarities of the AIS-based CA vary from 78.92% to 86.59%, with a mean value of 83.02% to actual growth patterns, while those of the Logistic CA basically have the values within the range of 50.43–61.96%. The mean value of those is only 56.96%. For these two models, the simulated ENN shows the largest disagreement. However, the Logistic CA also has very large errors in the simulated NP and PARA. These results indi-

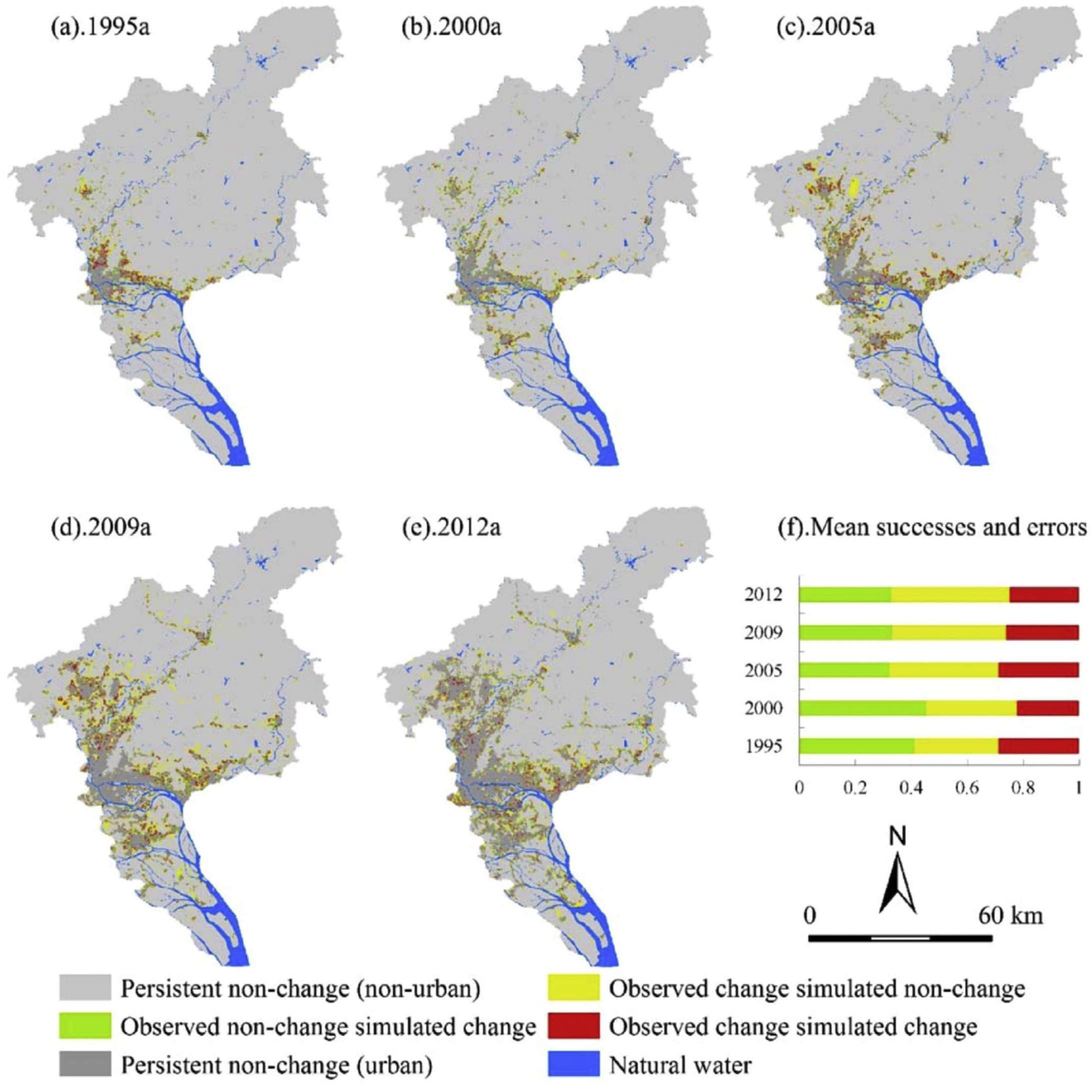

Fig. 7. Successes and errors distribution of the simulations during the period 1995–2012 produced by the AIS-based CA.

Table 2
Comparison of pattern-level similarity between the AIS-CA and the Logistic CA.

Year	Dataset	NP	LPI	PARA	ENN	Similarity (%)
1995	Observed data	1867	89.82	560.22	297.48	
	AIS-based CA	1351	88.86	582.23	385.02	80.31
	Logistic CA	4630	89.12	1157.0	109.33	51.56
2000	Observed	2038	74.66	540.68	284.21	
	AIS-based CA	1442	74.13	574.88	372.56	85.10
	Logistic CA	2054	76.95	1113.97	113.34	50.43
2005	Observed	2077	71.90	563.29	293.384	
	AIS-based CA	1551	72.48	552.09	357.92	86.59
	Logistic CA	4516	71.76	1154.74	92.99	61.96
2009	Observed	2787	68.12	580.67	242.47	
	AIS-based CA	1591	68.46	588.67	360.18	84.19
	Logistic CA	5004	68.73	1170.28	86.84	61.79
2012	Observed	3285	64.28	664.58	213.77	
	AIS-based CA	2350	65.28	630.73	288.13	78.92
	Logistic CA	8437	64.72	1119.33	90.20	59.05

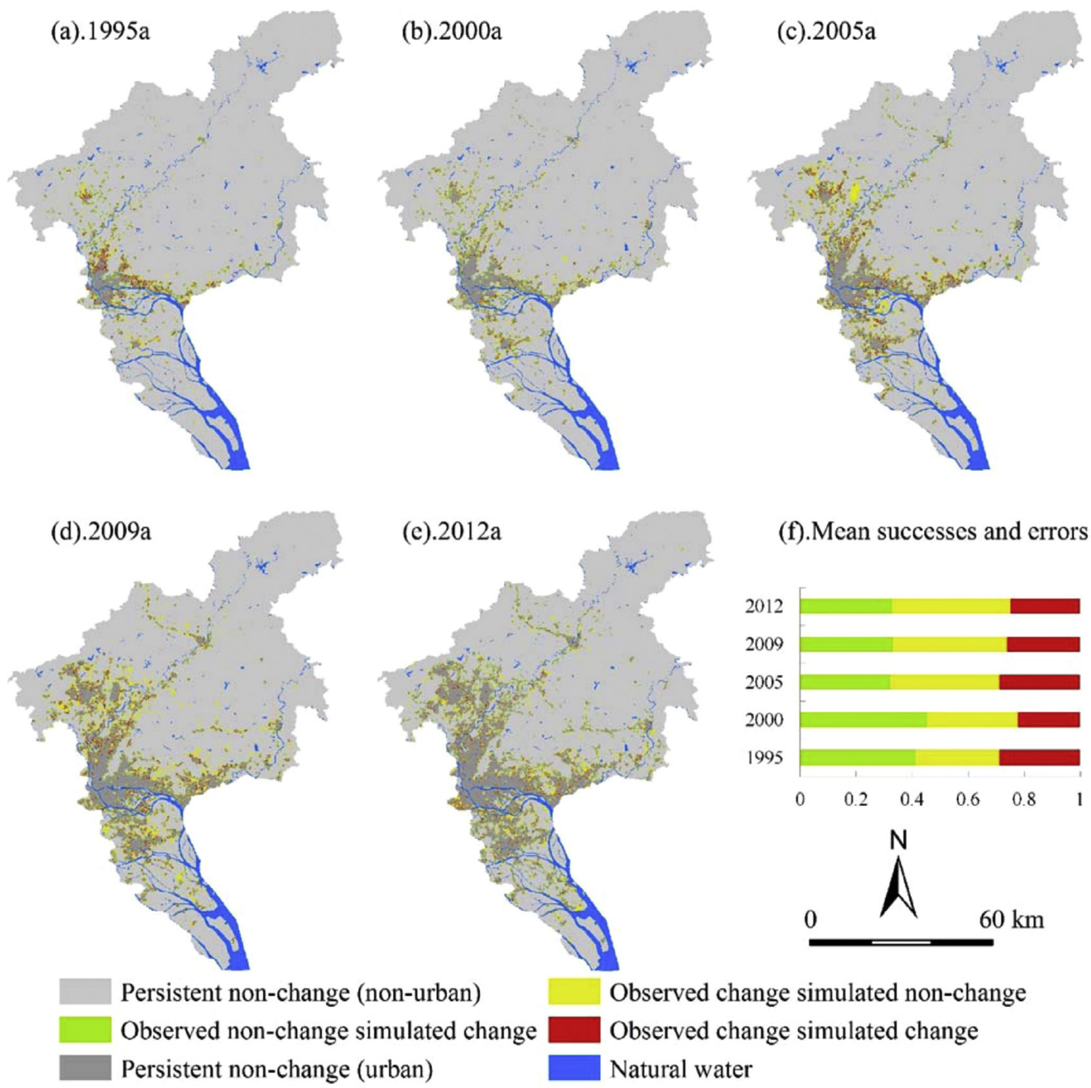

Fig. 8. Successes and errors distribution of the simulations during the period 1995–2012 produced by the Logistic CA.

cate that the AIS-based CA has better performance of replicating actual growth patterns than the Logistic CA.

Conclusions

It is not only theoretically significant but also of extensive application value to simulate complex geographical processes, which cannot be well revealed with conventional equation models. However, CA following the bottom-up approach can effectively simulate the evolution process of complex systems. The core of CA is to define proper transition rules for reflecting the urban evolution mechanism. The static transition rules discovered with traditional CA models generally pose problems in revealing the complexity, self-organization and dynamic features of geographical processes over time.

In this study, a self-adaptive geographical CA based on AIS was proposed. In the model, the antigen library is dynamically updated with the consideration of additional temporal data as new antigens, and the self-adaptive and self-learning AIS system is capable of generating corresponding antibodies to recognize new antigens. The new antibodies will be also provided for the antibody library to help calibrate model parameters. In this way, self-adaptive transition rules and dynamic urban evolution patterns can be derived in a time sequence. Furthermore, with the self-adaptive, self-learning and memorizing capabilities, the AIS-based CA model can accumulate sufficient experience with the variation of the external environment, and the transition rules can be correspondingly adjusted. It can be found that the model will be gradually calibrated with a series of spatial data, and dynamic transition rules can be automatically extracted to reveal the self-adaptive features of complex systems. Therefore, it is apparently advantageous in simulating complex geographical scenarios.

The self-adaptive AIS-based CA model was further applied to the urban growth simulation of Guangzhou city. With the TM satellite images collected in different years as the main observation data, sample points were randomly selected to constitute the initial antigen library. Then dynamic transition rules were automatically discovered with the self-adaptive AIS-based CA model. This was correspondingly used to simulate the urban evolution in the years 1995, 2000, 2005, 2009 and 2012 for the study area. The model was further validated with the established quantitative index 'FoM' and the pattern-level similarity. Compared with the conventional Logistic CA model, it shows that the cell level precision and pattern level similarity for the AIS-based CA model are both obviously higher, and the spatial pattern derived from the model is closer to actual urban growth situations. The AIS-based CA model is apparently suitable for simulating similar complex geographical phenomena such as hydrological process or landslide process due to the self-adaptive feature.

Acknowledgments

The authors would like to thank the editor, associate editor and anonymous reviewers for their helpful comments and suggestions. This work was supported by the National Natural Science Foundation of China under grant 41301418, and the China 973 Fundamental Research and Development Project under grant 2011CB707103.

References

Batty, M., 1993. Using Geographical Information Systems in Urban Planning and Policy Making. Geographical Information Systems: Spatial Modeling and Policy Evaluation. Springer-Verlag, Berlin, pp. 51–69.

Carter, J.H., 2000. The immune system as a model for pattern recognition and classification. J. Am. Med. Inform. Assoc. 7 (1), 28–41.

Chen, Y.M., Li, X., Liu, X.P., et al., 2014. Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy. Int. J. Geogr. Inform. Sci. 28 (2), 234–255.

Chun, J.S., Kim, M.K., Jung, H.K., et al., 1997. Shape optimization of electromagnetic devices using immune algorithm. IEEE Trans. Magn. 33 (2), 1876–1879.

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. Inform. Sci. 12 (7), 699–714.

Clarke, K.C., Hoppen, S., Gaydos, L., 1997. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environ. Plan. B: Plan. Des. 24, 247–261.

Couclelis, H., 1985. Cellular worlds: a framework for modeling micro–macro dynamics. Environ. Plan. A 17, 585–596.

Couclelis, H., 1989. Macrostructure and microbehavior in a metropolitan area. Environ. Plan. B 16, 141–154.

Couclelis, H., 1988. Of mice and men: what rodent populations can teach us about complex spatial dynamics. Environ. Plan. A 20, 99–109.

Dasgupta, D., Forrest, S., 1995. Tool breakage detection in milling operations using a negative-selection algorithm. In: Technical Report CS95-5. Department of Computer Science, University of New Mexico.

Dietzel, C., Oguz, H., Hemphil, J.J., et al., 2005. Diffusion and coalescence of the Houston metropolitan area: evidence supporting a new urban theory. Environ. Plan. B: Plan Des. 32 (2), 231–246.

De Casro, L.N., Von Zuben, F.J., 2000. Clonal selection algorithm with engineering applications. In: Proceedings: GECCO'00, LasVegas, Nevada, USA, pp. 36–37.

De Castro, L.N., Timmis, J., 2002. Artificial Immune Systems: A New Computational Intelligence Approach. Springer, London.

Feng, Y., Liu, Y., Tong, X., et al., 2011. Modeling dynamic urban growth using cellular automata and particle swarm optimization rules. Landsc. Urban Plan. 102 (3), 188–196.

Goodchild, M.F., 1992. Integrating GIS and spatial data analysis: problems and possibilities. Int. J. Geogr. Inform. Syst. 6 (5), 327–334.

Guan, D.J., Li, H.F., Inohae, T., et al., 2011. Modeling urban land use change by the integration of cellular automaton and Markov model. Ecol. Model. 222 (20), 3761–3772.

He, C., Zhao, Y., Tian, J., et al., 2013. Modeling the urban landscape dynamics in a megalopolitan cluster area by incorporating a gravitational field model with cellular automata. Landsc. Urban Plan. 113, 78–89.

Huang, K., Liu, X., Li, X., et al., 2013. An improved artificial immune system for seeking the Pareto front of land-use allocation problem in large areas. Int. J. Geogr. Inform. Sci. 27 (5), 922–946.

Jerne, N.K., 1973. The immune system. Sci. Am. 229 (1), 52–60.

Kim, J., Bentley, P., 1999. The artificial immune model for network intrusion detection. In: Proceedings of the 7th European Conference on Intelligent Techniques and Soft Computing, Aachen, Germany.

Kumark, K., Neidhoefer, J., 1997. Immunized neurocontrol. Exp. Syst. Appl. 13 (3), 201–214.

Li, X., Yeh, A.G.O., 2000. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. Int. J. Geogr. Inform. Sci. 14 (2), 131–152.

Li, X., Yeh, A.G.O., 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. Int. J. Geogr. Inform. Sci. 16, 323–343.

Li, X., Yang, Q.S., Liu, X.P., 2008. Genetic algorithms for determining the parameters of cellular automata in urban simulation. Science in China (Series D) 50 (12), 1857–1866.

Lipowski, A., Lipowska, D., 2012. Roulette-wheel selection via stochastic acceptance. Phys. A: Stat. Mech. Appl. 391 (6), 2193–2196.

Liu, X.P., et al., 2007. Simulating complex urban development using kernel-based non-linear cellular automata. Ecol. Model. 211, 169–181.

Liu, X.P., Li, X., Liu, L., et al., 2008. A bottom-up approach to discover transition rules of cellular automata using ant intelligence. Int. J. Geogr. Inform. Sci. 22 (11–12), 1247–1269.

Liu, X.P., Li, X., Shi, X., et al., 2010. Simulating land use dynamics under planning policies by integrating artificial immune systems with cellular automata. Int. J. Geogr. Inform. Sci. 24 (5), 783–802.

Liu, X., Ma, L., Li, X., et al., 2014. Simulating urban growth by integrating landscape expansion index (LEI) and cellular automata. Int. J. Geogr. Inform. Sci. 28 (1), 148–163.

Liu, X., Li, X., Tan, Z., et al., 2011. Zoning farmland protection under spatial constraints by integrating remote sensing, GIS and artificial immune systems. Int. J. Geogr. Inform. Sci. 25 (11), 1829–1848.

McGarigal, K., et al. (2012). FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. http://www.umass.edu/landeco/research/fragstats/fragstats.html. (accessed 02.02.13.).

Parker, D., Meretsky, V., 2004. Measuring pattern outcomes in an agent-based model of edge effect externalities using spatial metrics. Agric. Ecosyst. Environ. 101 (2–3), 233–250.

Pontius, R., et al., 2007. Accuracy assessment for a simulation model of Amazonian deforestation. Ann. Assoc. Am. Geogr. 97 (4), 677–695.

Pontius, R., et al., 2008. Comparing the input, output, and validation maps for several models of land change. Ann. Region. Sci. 42 (1), 11–37.

Seto, K., Fragkias, M., 2005. Quantifying spatiotemporal patterns of urban land-use change in four cities of China with time series landscape metrics. Landsc. Ecol. 20 (7), 871–888.

Sui, D., Zeng, H., 2001. Modeling the dynamics of landscape structure in Asia's emerging desakota regions: a case study in Shenzhen. Landsc. Urban Plan. 53 (1–4), 37–52.

Tarakanov, A., Dasgupta, D., 2000. A formal model of an artificial immune system. BioSystems 55 (1), 151–158.

Timmis, J., 2000. On Parameter Adjustment of Immune Inspired Machine Learning Algorithm AINE. University Kent, Canterbury.

Timmis, J., Neal, M.A., 2001. Resource limited artificial immune system for data analysis. Knowl. Based Syst. 14 (3–4), 121–130.

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

Yang, X., Zheng, X.Q., Lv, L.N., 2012. A spatiotemporal model of land use change based on ant colony optimization: Markov chain and cellular automata. Ecol. Model. 233, 11–19.

Zhong, Y., Zhang, L., Huang, B., et al., 2006. An unsupervised artificial immune classifier for multi-/hyperspectral remote sensing imagery. IEEE Trans. Geosci. Rem. Sens. 44 (2), 420–431.

Zhong, Y., Zhang, L., 2012. An adaptive artificial immune network for supervised classification of multi-/hyperspectral remote sensing imagery. IEEE Trans. Geosci. Rem. Sens. 50 (3), 894–909.