Computers, Environment and Urban Systems 65 (2017) 140–149

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

A segment derived patch-based logistic cellular automata for urban growth modeling with heuristic rules

Xuecao Li [a,b], Peng Gong [a,c,*], Le Yu [a,c], Tengyun Hu [d]

[a] Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China
[b] Department of Geological & Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA
[c] Joint Center for Global Change Studies, Beijing 100875, China
[d] Beijing Municipal Institute of City Planning and Design, Beijing 100045, China

ARTICLE INFO

Article history:
Received 12 July 2016
Received in revised form 8 May 2017
Accepted 5 June 2017
Available online xxxx

Keywords:
Cellular automata
Patch
Segmentation
Landscape
Urban growth types

ABSTRACT

Cellular automata (CA) models are extensively applied in urban growth modeling in different forms (i.e., pixel or patch). Studies have reported that the patch-based approach can achieve a more realistic urban landscape. However, they are subjected to uncertainties due to a variety of stochastic processes involved, which weakens their effectiveness on urban planning or decision making. Here, we propose a new patch-based urban growth model with heuristic rules that employed logistic CA model with a watershed segmentation algorithm (Segmentation-Patch-CA). The segment objects derived from features of urban CA model were regarded as potential patches for conversion, through defining a utility function that considered both the suitability and heterogeneity of pixels within the patch. Thereafter, two different urban growth types, i.e., organic growth and spontaneous growth, were identified and simulated separately by introducing a landscape expansion index (LEI) that built on neighborhood density analysis. The proposed Segmentation-Patch-CA was applied to Guangzhou City, China. Our results revealed that the proposed model produced a more realistic urban landscape (96.00% and 97.38%) than pixel-based (45.14% and 74.82%) for two modeling periods 2003–2008 and 2008–2012, respectively, when referring to an assembled indicator that closely related to urban patterns (e.g., shape, size, or distribution). Meanwhile, it also achieved a good performance when comparing to other patch-based urban CA models but with less uncertainty. Our model provided a very flexible framework to incorporate patches using segments or self-growth based on pixels, which is very helpful to future urban planning practices.

1. Introduction

Urban development and its influence on the environment, public health, land resources, and energy consumption have attracted an increasing amount of research attention (Chen, Li, Liu, & Ai, 2014; Gong et al., 2012; Li et al., 2016; Wang et al., 2012; Zhang, Cai, & Hu, 2013). As more and more people migrate to cities, large tracts of farmlands are disappearing with expanded built-up areas, particularly in rapidly developing regions (Liu, Li, Tan, & Chen, 2011; United Nations, 2015). Thus, urban growth modeling is a promising tool to better understand historical urban development and design future scenarios to support decision making and risk management (Li & Gong, 2016; Solecki, Seto, & Marcotullio, 2013).

Many urban growth models have been developed, including land use/transportation (LUT) models, agent-based models (ABM), and cellular automata (CA) models (Li & Gong, 2016). Among them, the CA-based urban growth model has been widely used in urban simulation because of its simplicity, transparency, and flexibility (Santé, García, Miranda, & Crecente, 2010). It works on a discrete grid system in which the evolution of the central cell is mainly driven by its neighbors (Batty & Xie, 1994). The well-known urban CA models include SLEUTH (Slope, Land use, Exclusion, Urban, Transportation, and Hillshade) (Clarke, Hoppon, & Gaydos, 1997), constrained CA (Li & Yeh, 2000; White & Engelen, 1993), Logistic-CA (Wu, 2002), Fuzzy CA (Liu & Phinn, 2003), intelligence CA (Li & Yeh, 2002; Liu, Li, Shi, Wu, & Liu, 2008; Liu et al., 2014), Markov CA (Muller & Middleton, 1994; Shafizadeh Moghadam & Helbich, 2013). However, most urban CA models are pixel- or cell-based (Li & Gong, 2016; Santé et al., 2010), which ignored the spatial homogeneity of urban land use evolution at local scales (Chen et al., 2014; Liu & Phinn, 2003; Meentemeyer et al., 2012; Wang & Marceau, 2013). For instance, cells with similar development probabilities may end up with different simulation results (e.g., converting or not) when adopted different thresholds (Li, Liu, & Yu, 2014b; Li & Yeh, 2000). Also, from the perspective of urban planning, urban land expansion is always based on parcels (or patches) (Hu,

* Corresponding author at: Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China.
E-mail address: penggong@tsinghua.edu.cn (P. Gong).

http://dx.doi.org/10.1016/j.compenvurbsys.2017.06.001
0198-9715/

Yang, Li, & Gong, 2016), which requires a patch-based approach to better support decision making or urban planning.

Currently, several attempts have been made that employ patch-based urban CA models using Voronoi diagrams (e.g., iCity model) (Moreno, Wang, & Marceau, 2009; Shi & Pang, 2000; Stevens, Dragicevic, & Rothley, 2007). However, patches in these models are always static during the modeling. An alternative is to generate patches through seed planting and self-growing based on pixel for achieving a reliable urban landscape (Chen, Li, Liu, Ai, & Li, 2016; Chen et al., 2014; Meentemeyer et al., 2012; Sohl, Sayler, Drummond, & Loveland, 2007). For instance, Sohl et al. (2007) proposed a self-growth method for patch generation based on planting seeds on the suitability surface. Later, this approach was replaced by a pre-collected patch library (i.e., patches with different sizes or shapes obtained from historical datasets) to improve efficiency (Sohl & Sayler, 2008). Similarly, Wang and Marceau (2013) applied the patch-based simulation at a very fine spatial resolution (5 m) in land use/cover change modeling. For particular urban growth, Meentemeyer et al. (2012) employed a patch-growth algorithm (PGA) to represent the fragmented landscape in rural–urban fringe areas. Chen et al. (2014) also proposed a Patch-Logistic-CA model that considered two different urban growth types, i.e., organic growth (expanding from initial developments) and spontaneous growth (growing apart from initial developments) (see Fig. 4 below), using a moving-window approach.

Although these patch-based CA models can produce more realistic urban landscape, they contain a variety of uncertainty during the key procedures such as patch generation, growth type identification and self-growth (Chen et al., 2014; Meentemeyer et al., 2012; Sohl et al., 2007). First, the initial seed for patch growth was determined using a Monto Carlo approach based on suitability surface. That is, patch seeds were picked out if generated random values were greater than suitability values underneath, which will result in different seeds and finally form patches during the modeling. Thus, a repeated running in many times is required to achieve a robust result. Second, Chen et al. (2014) introduced two different urban growth types (i.e., organic growth and spontaneous growth) in their Patch-Logistic-CA model. Nevertheless, the growth types were determined in a stochastic manner, through comparing a random value to a predefined (or calibrated) threshold of development probability, whereas there lacks a clear mechanism behind. Although these patch-based urban CA models reported better performance in modeled urban landscape, the uncertainties caused by stochastic processes are difficult to interpret.

To address these issues, we propose a new patch based urban CA model with heuristic rules. Those objects that derived from watershed segmentation were used as basic modeling unit. Both the patch selection and their growth types were quantitatively assessed using defined heuristic rules. Finally, the proposed model was applied in Guangzhou (China) city to validate the model performance.

2. Study area and data

The proposed model was tested in Guangzhou City, a rapidly developing region in China (Fig. 1). This city is one of the megacities located in the Pearl River Delta and is attracting thousands of people for residency every year (Li et al., 2014b). As a result, a rapid urbanization process occurred over the past decades. Massive croplands around the central city have been encroached by emerging urban development driven by stimulating policies (Liu et al., 2011). Hence, we selected Guangzhou as our study area for model development and test.

The datasets used here were mainly derived from Landsat Thematic Mapper images of Guangzhou dated in 2003, 2008 and 2012. After a series of image preprocessing (i.e., geometric, topographical and atmospheric corrections), the mean classification accuracies of these two images were estimated to be 83% to 85% according to field checking, including three categories (i.e., urban, non-urban, and water) (Chen, Li, Zheng, Guan, & Liu, 2011). A total of eight proximity variables were employed for transition rule extraction. These variables are distance to road, distance to expressway, distance to railway, distance to subway, distance to main centers, distance to district centers, distance to large town centers, and distance to small town centers. This dataset has been successfully tested in other studies (Li, Lin, Chen, Liu, & Ai, 2013).

3. Methodology

Our model was built on a prototype of Logistic-CA, but the unit is replaced by patches derived from segmentation (named as Segmentation-Patch-CA) (Fig. 2). The whole framework aimed at defining a clear process of patch evolution using heuristic rules. First, segments were generated using a watershed algorithm based on CA components, including suitability surface, neighborhood, stochastic term and development surface. These are potential candidates to patch growth and further aggregate into super patches. Then, they were assessed quantitatively with a utility function to determine their conversion, which differs from stochastic alternatives on seed planting. Thereafter, we introduced a neighborhood-density modified index of landscape expansion index (LEI_{nei}) to determine two different urban growth types (i.e., organic and spontaneous growth) (Liu et al., 2010a). This provides a theoretical base and reduced the uncertainty caused by stochastic process. Finally, based on the identified growth types, we implemented different growth strategies for a given patch (i.e., based on pixel or segment). Details of each component are given in the following sections.

3.1. Logistic-CA

Logistic-CA model has been widely adopted in urban growth simulation because of its clear explanation of each spatial features and ease of implementation (Hu & Lo, 2007; Li, Liu, & Gong, 2015b; Wu, 2002). The key is to use logistic regression model for eliciting transition rules. Assuming there are n drivers $[x_1,x_2,\ldots,x_n]$, the logistic regression model can be expressed as Eqs. (1) and (2).

$$z = b_0 + b_1x_1 + \ldots + b_nx_n \tag{1}$$

$$p_{suit} = \frac{exp(z)}{1 + exp(z)} \tag{2}$$

where p_{suit} is the obtained suitability value (or transition rules in Santé et al. (2010)) and b_i,x_i represent the ith coefficient and variable, respectively (i.e., b_0 is the coefficient of the constant term).

The neighborhood is a crucial component in urban CA models because it is a basic driver for dynamic simulation. The neighborhood impacts on urban expansion are closely related to its size, shape, and ambient land use (Kocabas & Dragicevic, 2006). Moore neighborhood is a commonly used configuration (Li & Gong, 2016) and its intensity can be quantified as Eq. (3).

$$\Omega = \frac{\sum_{m\times m} Con(S_{ij} = urban)}{m \times m - 1} \tag{3}$$

where Ω denotes the influence of the neighborhood and m is the window size, S_{ij} is the land use/cover status of cell (i,j), and $Con()$ is a conditional function and returns 1 when S_{ij} is urban.

Stochastic perturbation represents unconsidered factors (e.g., policies) in the Logistic-CA model. White and Engelen (1993) provided a quantitative expression of stochastic perturbation [Eq. (4)], where RA denotes stochastic perturbation, λ is a random value [0, 1], and α is a parameter that controls the degree of perturbation.

$$RA = 1 + (-\ln\lambda)^{\alpha} \tag{4}$$

Given that only urban and non-urban were considered (excluded water), here we defined the land use/cover constraint as $Land_{ij}$. Finally,