Environmental Modelling & Software 75 (2016) 163–175

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Incorporation of extended neighborhood mechanisms and its impact on urban land-use cellular automata simulations

Jiangfu Liao [a, b], Lina Tang [a, *], Guofan Shao [a, c], Xiaodan Su [a], Dingkai Chen [a], Tong Xu [a]

[a] Key Lab of Urban Environment and Health, Institute of Urban Environment, Chinese Academy of Sciences, Xiamen, 361021, China
[b] Computer Engineering College, Jimei University, Xiamen, 361021, China
[c] Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 5 January 2015
Received in revised form
15 October 2015
Accepted 17 October 2015
Available online 11 November 2015

Keywords:
Urban expansion
Cellular automata
Extended enrichment factors
Neighborhood effects
Neighborhood rules

ABSTRACT

Urban cellular automata (CA) models are broadly used in quantitative analyses and predictions of urban land-use dynamics. However, most urban CA developed with neighborhood rules consider only a small neighborhood scope under a specific spatial resolution. Here, we quantify neighborhood effects in a relatively large cellular space and analyze their role in the performance of an urban land use model. The extracted neighborhood rules were integrated into a commonly used logistic regression urban CA model (Logistic-CA), resulting in a large neighborhood urban land use model (Logistic-LNCA). Land-use simulations with both models were evaluated with urban expansion data in Xiamen City, China. Simulations with the Logistic-LNCA model raised the accuracies of built-up land by 3.0%–3.9% in two simulation periods compared with the Logistic-CA model with a 3 × 3 kernel. Parameter sensitivity analysis indicated that there was an optimal large window size in cellular space and a corresponding optimal parameter configuration.

1. Introduction

Land-use dynamics constitute an open and complex spatio-temporal evolution process that involves multi-element composite effects from natural, social, and economic factors (Arsanjani et al., 2013; Fuglsang et al., 2013; Hewitt et al., 2014). Environmental modeling can support scientific decision-making processes, and thus contribute to sustainable development associated with land-use changes. Spatial simulations and quantitative analyses of urban land-use dynamics are effective ways to improve the understanding of the evolution of urban landscapes. Cellular automata (CA) have drawn increasingly more attention in the field of land-use and land-cover analysis and simulation. The 'bottom–up' approach of CA fully reflects the concept that complex global patterns emerge from interactions governed by local rules. In addition, CA are ideal for simulating and predicting complex geographic phenomena (Liu et al., 2008a).

Based on the pioneering work by Tobler (1979) and Couclelis (1988), many researchers have developed urban land-use CA models over the last three decades, resulting in significant achievements (Batty and Xie, 1994; Clarke et al., 1997; Li and Yeh, 2000; Liu et al., 2007; Stevens et al., 2007; Takeyama and Couclelis, 1997; Verburg et al., 2004b; Wu, 2002). These models generally included a combination of drivers and spatiotemporal interactions among land uses in neighborhoods.

Identifying transition rules is a key issue in urban CA. Typically, a variety of biophysical and socioeconomic factors are included in transition rules as driving forces of urban development. Researchers have proposed various methods to determine the contributions of different spatial variables and to calibrate urban CA models (Al-Ahmadi et al., 2009; Dai et al., 2005; Feng and Liu, 2013; Feng et al., 2011; Kocabas and Dragicevic, 2007; Li and Yeh, 2002, 2004; Liao et al., 2014; Liu et al., 2008a; Verstegen et al., 2014; Wang et al., 2013; Wu, 2002; Wu and Webster, 1998; Yang et al., 2008). The binary logistic regression method developed by Wu (2002) has been widely used in urban land-use modeling because of its strict theoretical basis of statistical learning and empirical characteristics, and it has become a classic calibration method for urban CA (Cheng and Masser, 2003; Dendoncker et al., 2007; Hu and Lo, 2007; Verburg et al., 2004a). More recently, new socio-economic factors such as per-capita gross domestic product (GDP), land price, employment potential, and population density have been incorporated into the driving forces of urban CA models and integrated with logistic regression and Markov chain analysis to

* Corresponding author.
E-mail address: lntang@iue.ac.cn (L. Tang).

http://dx.doi.org/10.1016/j.envsoft.2015.10.014
1364-8152/

predict future scenarios of urban development (Arsanjani et al., 2013; Guan et al., 2011; Mas et al., 2014). However, these models mainly considered a 3 × 3 kernel, which is a relatively small neighborhood, though studies have noted that the logistic regression urban CA model is sensitive to scale (Pan et al., 2010).

Neighborhood interaction rules are an important subset of transition rules and play a key role in the calculation of cellular conversion probabilities. To quantify and analyze the neighborhood effects generated by surrounding cells at different distances from a central cell, Verburg et al. (2004b) defined an enrichment factor formula for measuring the over- and under-representation of specific land uses in cellular space. More recently, other studies have achieved better simulation results by applying this enrichment factor to determine the neighborhood parameters of urban cellular models or as empirical data for calibrating neighborhood interaction rules (Hansen, 2008; Van Vliet et al., 2013). However, these models generally considered a small neighborhood scope with a relatively short distance from the central cell under a specific resolution during the application. For example, Van Vliet et al. (2013) used a neighborhood radius covering 0–4 unit distances (the discrete ring of a cell with a width of 500 m) to simulate urban land-use dynamics at a country scale in Germany.

The external effects generated by concentrative and dispersive forces play an important role in urban dynamics and are seen as the organizing forces of urban patterns (Harrop, 1973; Krugman, 1999; Rodrigue, 2004). Hagoort et al. (2008) pointed out that neighborhood interaction rules specify how the combined effects of spatial externalities work over distance in cellular space. Spatial externalities are considered to represent the aggregated effects of a specific land-use type on another in the neighborhood (Hagoort et al., 2008; Hansen, 2008). Research on neighborhood effects has shown that a neighborhood scope greater than a relatively small window size (i.e., a large neighborhood window) still has a significant influence on the development of the center cell (Hagoort et al., 2008; White and Engelen, 2003). In fundamental urban CA, the decay coefficient of a small neighborhood function will eventually approach zero as the radius of the neighborhood increases (Van Vliet et al., 2013). Thus a small neighborhood function cannot effectively express the impact of spatial externalities existing in a relatively large neighborhood window on the development of the central cell.

In summary, neighborhood interactions in urban CA models have mainly been limited to a 3 × 3 kernel or relatively small moving window, partially due to the aim of simplifying the models (White and Engelen, 2000). The neighborhood rules established in this case are unsuitable for detecting complex neighborhood effects over a larger scope. This problem is not prominent when the spatial resolution of geospatial data is low. However, high-spatial resolution remote sensing data have become readily available and increasingly popular. Thus, interaction rules designed for complex neighborhood effects in urban CA models are encountering unprecedented challenges. The goal of this paper is to characterize the role of complex neighborhood effects over a relatively large scope associated with urban sprawl simulation and prediction. A modeling exercise was designed to answer the following questions: 1) do large neighborhood effects exist on urban sprawl processes? 2) if yes, how can large neighborhood rules in urban CA modeling be calibrated? and 3) what is the expected increase in locational accuracy of the urban CA when large neighborhoods are incorporated?

This study addresses extended neighborhood effects on urban dynamics by using an extended neighborhood structure that is composed with cells with various influence weights based on their distances from the central cell. We used the extended neighborhood structure and calibrated parameter values to establish a large-window neighborhood function. Based on this, we developed an extended neighborhood model of urban land-use change, Logistic-LNCA, and applied it to simulate land-use changes in Xiamen City of China from 1990 to 2000. We then validated this method by using independent data acquired between 2000 and 2010.

The methodology for this study is given in the next section, together with a concise flowchart of the Logistic-LNCA model. Simulation experiments and result evaluations are presented in section three. Results are discussed in section four, and conclusions and further research directions are provided in section five.

2. Modeling methods

2.1. Model calibration based on logistic regression

Urban models simulate urban morphology evolution under various scenarios by characterizing a series of development profiles, which include physical attributes, socioeconomic status, planning and zoning constraints, and the effects of complex neighborhood interactions. Spatiotemporal models based on CA can reveal the agglomeration effects of land use at a local scale or the level of development through the iterative calculation of local and simple rules. Thus, the two interrelated processes of urban land development—spontaneous growth and self-organized growth—can be reproduced in urban cellular lattices (Wu, 2002). However, calibrating the contributions of the various aforementioned attributes to land development is a critical step to achieve more realistic and reliable urban CA simulations. Logistic regression or the multinomial logit model can be used to estimate the relationship between urban land-use changes and corresponding locational features (Bishop, 2006; McCullagh and Nelder, 1989; McMillen, 1989). More specifically, logistic regression can be seen as a process to extract the coefficients of the empirical relationships between observed land-use changes and driving forces in the integration with urban CA simulation (Wu, 2002).

Sample size and sampling strategy are two basic issues that affect the results of logistic regressions (Hirzel and Guisan, 2002; Huang et al., 2009; Munroe et al., 2004; Xie et al., 2005). Because sample size and resultant errors have an inverse relationship, a large sample size can better represent the characteristics of the study area but requires greater computing resources.

The sampling methods used in logistic regression models generally include systematic and random sampling. Systematic sampling can reduce spatial autocorrelation but may lose detailed information on some relatively isolated cells. Random sampling may better represent the population, but cannot effectively reduce spatial autocorrelation, especially local spatial dependence (Xie et al., 2005). A reasonable sampling scheme should maintain a balance between spatial autocorrelation and effective population representation (Huang et al., 2009). Considering the multiple characteristics of urban land-use modeling, we integrated systematic and random sampling, namely, proportional random-stratified sampling and extracted adequate samples to eliminate the spatial dependence of the population (Xie et al., 2005).

Generally, urban land-use change models quantify the local transition suitability of each cell from a set of demographic, econometric, and physical factors (Arsanjani et al., 2013; Fuglsang et al., 2013; Lauf et al., 2012). Cells with higher suitability are given higher probabilities in transition rules. In urban expansion simulations, the cellular space can be classified into two types, developed cells (built-up land) and undeveloped cells (non built-up land). The local development suitability at a location can be considered a function of various independent spatial variables, including elevation, slope, distance to the city center, distance to the town center, distance to the main road, distance to the railway, distance to the coast, and so on. Thus, because of the binary features of the dependent variable, it is suitable to evaluate the contributions of different driving forces to urban development using the binary logistic regression approach (Wu, 2002). The regression equation used to calculate local development suitability is as follows:

$$P(y=1|\mathbf{X}) = \frac{1}{1+e^{-(\alpha_0+\sum \boldsymbol{\alpha}\mathbf{X})}} \tag{1}$$

where P is the probability of the dependent variable, representing the assessed value of local development suitability; $\mathbf{X}$ is the vector of the independent variables, representing a set of development factors [spatial variables, $\mathbf{X}=(x_1,x_2,\cdots,x_k)$]; and $\boldsymbol{\alpha}$ is the vector of estimated parameters [$\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_k)$].

2.2. Fundamental logistic-based urban cellular automata model (Logistic-CA)

In general, a CA primarily consists of cells, states, neighborhoods, and transition rules. From the perspective of model runs, various transition rules and their combinations form the core engine that drives the changes in system state over time. The key of this engine is the identification of various model parameters contained in the transition rules, which define the characteristics of the process represented by the transition rules. Therefore, a general urban CA can be defined as (Verstegen et al., 2014):

$$S_{ij}^t = f\left(S_{ij}^{t-1},\ \mathbf{I}^t,\ \mathbf{A}^t\right), \quad \text{for each} \quad t=1,2,\ldots,T \tag{2}$$

where S_{ij}^t and S_{ij}^{t-1} represent the state of cell (i, j) at moment t and $t-1$, respectively, including built-up land, non built-up land, and water in this paper; the vector $\mathbf{I}^t$ represents all inputs, which usually include spatial attributes and boundary conditions; the vector $\mathbf{A}^t$ is a collection of all parameters, which constrains the contributions of various model inputs to a great extent; and f is a series of transition rules that lead to change in the system state over time.

Given the uncertainty of urban systems, it is feasible to express the conversion potential of a location with probability rules rather than deterministic rules. If model inputs are further divided into local development factors, neighborhood conditions, constraints and stochastic disturbance factors, the probability of converting the cell (i, j) from the undeveloped state to the developed state at moment t can be expressed as follows (Feng et al., 2011; White and Engelen, 1993; Wu, 2002):

$$P_{ij}^t = (P_l)_{ij}\cdot(P_\Omega)_{ij}\cdot con()\cdot P_r \tag{3}$$

where P_{ij}^t is the total development probability of cell (i,j) at moment t; $(P_l)_{ij}$ is the assessed value of local development suitability based on a variety of spatial distance variables such as the slope or distance to the city center (in a logistic-based urban CA (Logistic-CA), parameters for obtaining $(P_l)_{ij}$ are calibrated using the method mentioned in Section 2.1 and the value of $(P_l)_{ij}$ is computed with Eq. (1) (Arsanjani et al., 2013; Feng et al., 2011)); $(P_\Omega)_{ij}$ is the development density of a relative small cellular neighborhood; $con()$ is the constraint conditions of land development in this region; and P_r is the stochastic disturbance factor of urban evolution.

It must be noted that probabilistic transition rules can be expressed in forms other than Eq. (3) (Lagarias, 2012; Liu et al., 2012; Van Vliet et al., 2013). However, this research mainly refers to the logistic regression model proposed by Wu (2002), and Eq. (3) defines the Logistic-CA model in this paper.

Neighborhoods are important factors affecting the simulation of urban evolution. To generate a more compact space layout for land-use change simulations, a neighborhood window (3 × 3 Moore) is generally used to calculate interactions among different land-use types. The neighborhood function can be defined as (Liu et al., 2008b, 2014):

$$(P_\Omega)_{ij} = \frac{\sum_{3\times3} con\left(s_{ij}^t = dev\right)}{3\times3-1} \tag{4}$$

where $(P_\Omega)_{ij}$ is the development density of the neighborhood and $\sum\limits_{3\times3} con(s_{ij}^t = dev)$ is the number of developed cells in the neighborhood window.

Numerous objective development constraints such as water, woodland, and farmland protection zones and limited planning areas must be considered in the model, and the probability of converting these areas to urban land is generally low. The term $con(s_{ij}^t = \text{suitable})$ can be used to represent the development constraints of the center cell. In order to simplify the calculation, the value of $con()$ is usually 0 or 1.

Urban expansion is influenced by various forces such as natural and geographic factors, political factors, socioeconomic factors, and accidental events, and it can become even more complex due to human interventions. Therefore, to make the simulation results better represent actual development, a stochastic disturbance factor is introduced into the model (White and Engelen, 1993).

$$P_r = (1+(-\ln\gamma)^\alpha) \tag{5}$$

where $(1+(-ln\gamma)^\alpha)$ is the stochastic factor; γ is a random number in the range of [0, 1]; and α is an integer ranging from 1 to 10 that controls the effect of the stochastic factor (White and Engelen, 1993).

After calculating the development probability of the center cell according to Eq. (3), a threshold value in the range of [0, 1] is generally predefined. The model can decide whether a particular cell is converted by comparing the development probability with the threshold value in each iteration.

$$S_{ij}^t = \begin{cases} Developed, & P_{ij}^t \geq P_{\text{threshold}} \\ Undeveloped, & P_{ij}^t < P_{\text{threshold}} \end{cases} \tag{6}$$

where S_{ij}^t is the state of cell at moment t; P_{ij}^t is the value of development probability; and $P_{\text{threshold}}$ is the threshold value of cellular conversion.

2.3. Measuring large neighborhood effects and developing a new urban CA model (Logistic-LNCA)

A key task of land-use change modeling involves identifying the most important driving forces of urban expansion and determining how to express these factors in the model (Hansen, 2008). Previous research has indicated that a large neighborhood significantly influences the development of the central cell (Hagoort et al., 2008; Hansen, 2008; Van Vliet et al., 2013). Verburg et al. (2004b) defined a spatial indicator or enrichment factor to quantify and analyze neighborhood characteristics. The enrichment factor is defined as an over- or under-representation of a certain land-use type in the neighborhood of a specific location compared with the overall average level of this land use in the entire study area. It can be used to calculate the neighborhood effects of all land-use types for each neighborhood radius.

To measure the large neighborhood effect in a specific location, we followed Verburg et al. (2004b) and modified the formula for the enrichment factor. The extended formula was designed to measure the over- or under-representation of a specific land-use type in various sub-neighborhoods with different distances to the central cell in a relatively large neighborhood. The formula for this extended enrichment factor is as follows:

$$F_{i,l,\bigcup d_j} = \frac{n_{i,l,\bigcup d_j} \Big/ n_{i,\bigcup d_j}}{|N_l|/|N|}$$

$$\bigcup d_j = \left\{ \bigcup_r^{r+\Delta r} d_j \middle| r < d_j \le r + \Delta r,\ \ r \in (R_{\min}, R_{\min} + \Delta r, R_{\min} + 2\Delta r, \ldots, R_{\max}) \right\} \tag{7}$$

where $(R_{\min}, R_{\max})$ is the scope of the large neighborhood, which is divided into several sub-neighborhoods with a separation distance Δr; $\bigcup d_j$ is sub-neighborhood with a specific distance to location i and a distance range in the interval $(r, r+\Delta r)$; distance d_j is the Euclidean distance between two cells; $(0, R_{\min})$ represents a relatively small neighborhood scope around the central cell (e.g., the value of $R_{\min}$ might be 1, 4, or 7; Fig. 1 shows the shape of the large neighborhood used in this study); $F_{i,l,\bigcup d_j}$ is the enrichment of land use l on location i in the distance set $\bigcup d_j$; $n_{i,l,\bigcup d_j}$ is the number of cells of land use l located in the sub-neighborhood $\bigcup d_j$ of location i; $n_{i,\bigcup d_j}$ is the total number of cells located in the sub-neighborhood $\bigcup d_j$ of location i; and $|N_l|$ and $|N|$ are the number of cells with land use l and the total number of cells in the study area, respectively.

An enrichment factor for a certain land-use type between 0 and 1 indicates an under-representation of this land-use type in the neighborhood. For example, when the neighborhood of a particular location contains 10% agricultural land, but the average proportion of agricultural land in the study area is 40%, the neighborhood model under-represents this land-use type and the enrichment factor is 0.25, which is between 0 and 1. If the ratio of a certain land-use type in the neighborhood is equal to the average ratio of the study area, the value of the enrichment factor for this land-use type in the neighborhood is equal to 1. The over-representation of a certain land use type results in an enrichment factor greater than 1 (Verburg et al., 2004b).

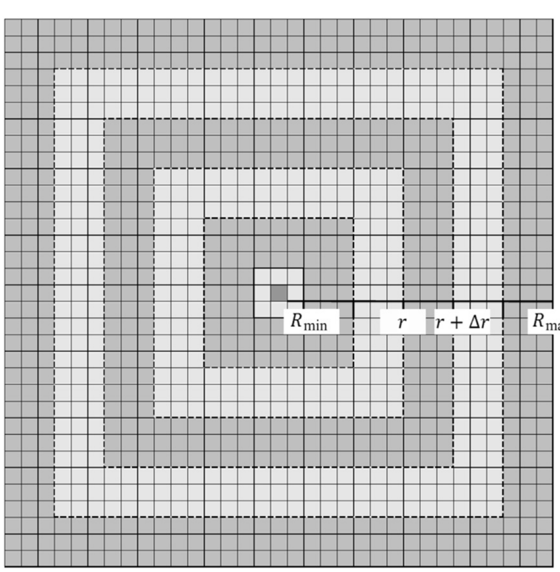

Fig. 1. Illustration of a large neighborhood configuration in the Logistic-LNCA model.

Eq. (7) has been applied in several European-based studies, which demonstrated that the location relationships of some land-use combinations have strong correlations (Van Vliet et al., 2013; Verburg et al., 2004a). Enrichment factors can be computed for different land uses and specific subsets (Hansen, 2012; Van Vliet et al., 2013). For example, Hagoort et al. (2008) calculated the average enrichment factor for a certain land-use type over an area at a particular moment T. To express the effects of various land-use types on observed land-use changes during a specific period, we assume that the set of all locations of a new land-use type k during this period is $\boldsymbol{K}$. Thus, the average extended enrichment factor for the land-use set $\boldsymbol{K}$ can be calculated as follows (Van Vliet et al., 2013; Verburg et al., 2004b):

$$\overline{F}_{k,l,\bigcup d_j} = \frac{1}{|N_k|} \sum_{i \in \boldsymbol{K}} F_{i,l,\bigcup d_j} \tag{8}$$

where $|N_k|$ is the number of cells that are changed into land use k between T_1 and T_2; $\overline{F}_{k,l,\bigcup d_j}$ is the average enrichment of land use l on all cells that are changed into land use k at a distance range $\bigcup d_j$; i is a location in land use set $\boldsymbol{K}$; and $F_{i,l,\bigcup d_j}$ is the extended enrichment factor measured by the land use map at T_1. The average extended enrichment factor can be expressed in logarithmic form and used to compare the over- or under-representations of different land-use types in different subsets of a large neighborhood.

Eq. (8) was used to test the existence of large neighborhood effects of land-use types on the central cell i. However, characterizing the extent of this influence requires further investigation. For different study areas and corresponding spatial resolutions, configurations of a large neighborhood may differ in size and shape. A large neighborhood generally extends over a relatively large spatial area, in which various attractive and repulsive effects among land-use combinations exist. The large neighborhood can be divided into multiple subsets via Eq. (7), and the neighborhood effects of land use l on location i in these subsets can then be estimated. For the purpose of investigating urban expansion dynamics in this paper, the proposed model primarily expresses the influence of neighborhood effects on central cell through neighborhood development distributions. Therefore, the extended enrichment factor defined in Eq. (7) is mainly used to measure the influences of built-up land on the central cell during the model calibration stage. Thus, land use l in Eq. (7) is limited to built-up land in Section 2.3, and the corresponding subscript in the extended enrichment factor is expressed as dev. By treating land-use changes during T_1 and T_2 as dependent variable and extended enrichment factors of different neighborhoods of the $\bigcup d_j$ distance sets as independent variables, the logistic regression method descripted in Section 2.1 can be used to calibrate the large neighborhood effects:

$$P(\mathrm{y}=1|\mathbf{F}) = \frac{1}{1 + e^{-\left(\beta_0 + \beta_1 F_{i,\mathrm{dev},\bigcup d_1} + \beta_2 F_{i,\mathrm{dev},\bigcup d_2} + \ldots + \beta_n F_{i,\mathrm{dev},\bigcup d_n}\right)}} \tag{9}$$

where P is development probability of cell i considering only large neighborhood effects; $\mathbf{F}$ is the vector of extended enrichment factors [spatial attributes, $\mathbf{F} = (F_{i,\mathrm{dev},\bigcup d_1}, F_{i,\mathrm{dev},\bigcup d_2}, \cdots, F_{i,\mathrm{dev},\bigcup d_n})$]; β_n is the coefficient of a different subset of the large neighborhood; $F_{i,\mathrm{dev},\bigcup d_n}$ is the enrichment factor calculated with Eq. (7), especially for measuring the over- or under-representation of built-up land in the extended neighborhood scope; and $\bigcup d_n$ is a subset of the extended neighborhood.

Eq. (9) can be extended to include the effects of different land-use types. With the increase in the spatial extent of the moving window of the large neighborhood, there will be many independent variables in the formula, which can be adjusted by setting the neighborhood subset interval Δr. Thus, the large neighborhood function can be constructed and the development probability affected by the large neighborhood can be calculated as follows:

$$P_{ln} = \left(\frac{1}{1 + e^{-\left(\beta_0 + \beta_1 F_{i,\mathrm{dev},\bigcup d_1} + \beta_2 F_{i,\mathrm{dev},\bigcup d_2} + \ldots + \beta_n F_{i,\mathrm{dev},\bigcup d_n}\right)}} \right)^{\delta} \tag{10}$$

where P_{ln} is the value of the large neighborhood function; $F_{i,\mathrm{dev},\bigcup d_n}$ is the same extended enrichment factor input used in Eq. (9); β_n is the parameter value in the large neighborhood module estimated by Eq. (9); and δ is a control factor used to adjust for the effect of the large neighborhood in the urban model.

Eq. (3) is a fundamental form of constrained CA, a commonly used model in the field of urban simulation. Conceptually, this type of model includes the components of development suitability, neighborhood effects, constraints, and stochastic perturbations (White et al., 1997). Adding a large neighborhood as a moving window to an urban CA model allows Eq. (3) to be adjusted as follows (García et al., 2012; Li et al., 2014):

$$P_{ij}^{t} = (P_l)_{ij} \cdot (P_\Omega)_{ij} \cdot (P_{ln})_{ij} \cdot con() \cdot P_r \tag{11}$$

where P_{ij}^{t} is the total development probability of cell (i, j) at moment t in the new urban CA model; $(P_l)_{ij}$ is the local development suitability of cell (i, j), which is calculated with the method proposed in Section 2.1 and has the same inputs and parameters as the Logistic-CA model proposed in Section 2.2; $(P_\Omega)_{ij}$ is the impact of the 3 × 3 kernel neighborhood at moment $t-1$; $con()$ is the constraints on urban growth at moment $t-1$; $(P_{ln})_{ij}$ represents large neighborhood effects at moment $t-1$; and P_r is the stochastic perturbation term. The model presented here is called Logistic-LNCA, which considers extended neighborhood effects. The Logistic-CA model proposed in Section 2.2 does not have such considerations.

The components/procedures of the urban CA considering large neighborhood effects are illustrated in Fig. 2.

3. Simulation experiments and result evaluations

3.1. Experimental design

The simulation experiment was designed to include a model calibration in 1990–2000 (Sections 3.4, 3.5, and 3.6) and an independent validation in 2000–2010 (Section 3.7). Referring to previous studies on urban CA modeling and testing (Feng and Liu, 2013; Feng et al., 2011; Li et al., 2014; Van Vliet et al., 2013), the generated transitional rules were applied to urban expansion simulations during the same two periods. During the validation stage, observed land-use changes were treated as an independent data set and the transition rules calibrated during the first period were used to predict urban expansion dynamics at the end of second period (Van Vliet et al., 2013).

According to Eqs. (2) and (3), urban CA models are mainly characterized by model inputs and parameter settings. From this point of view, the Logistic-CA model is defined by a series of spatial variables and corresponding parameter values used for determining local development suitability. In contrast, the Logistic-LNCA model also includes inputs and parameters for the extended neighborhood rule, except those for the local development suitability component. During the independent validation period (2000–2010), the parameter configurations related to the aforementioned transition rules in the two urban CA models were all obtained during the calibration period (1990–2000), and the model inputs remained unchanged.

3.2. Study area and experimental data

Xiamen City in southeast China, with an area of 1574 km^2, was selected as the study area for this research. In 2010, Xiamen City had a total population of 1,802,060 (Tang et al., 2013; XCSB, 2011).

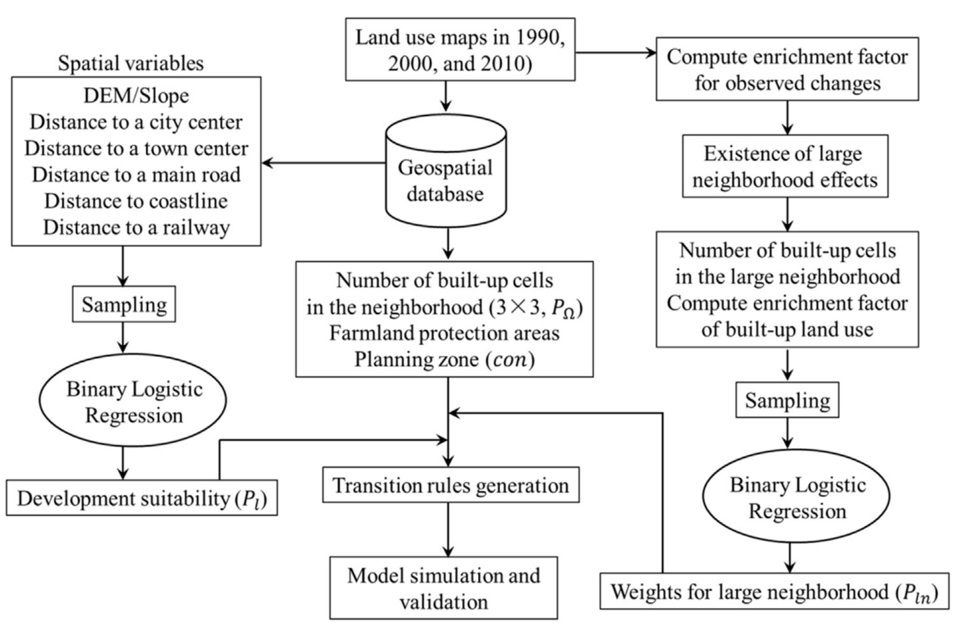

Fig. 2. A flowchart of the urban CA model integrating the calibration of large neighborhood rules.

Its urban landscape consisted of two urban centers (Siming District and Huli District on Xiamen Island) and four sub-centers outside the island (Jimei, Haicang, Tongan, and Xiangan). Recently, Haicang, Jimei, and Xiangan have gradually become the new industrial centers.

Xiamen City is one of the first four special economic zones in China. Its urbanization has accelerated over the past three decades, and its built-up area increased from 34.01 km^2 in 1989 to 197 km^2 in 2008 (Chen and Xu, 2005; XCSB, 2008). Such rapid urban expansion poses serious challenges for the region in terms of sustainable development, environmental loads, and ecosystem services (Shaker, 2015).

Land-use maps of the study area in 1990, 2000, and 2010 were obtained from the Institute of Remote Sensing and Digital Earth of the Chinese Academy of Sciences. Each land-use map has a 30 $\times$ 30 m spatial resolution and includes farmland, woodland, grassland, urban residential land, industrial land, water, and rural areas (Fig. 3).

First, the land-use datasets from 1990 to 2000 were used to compute the extended enrichment factors of different land-use types in the neighborhood of residential land during the periods of 1990–2000 and 2000–2010 and to verify the existence of neighborhood effects. Subsequently, the land-use maps from 1990, 2000, and 2010 were regrouped into three categories (built-up land, non built-up land, and water bodies). Industrial land and residential land were merged into built-up land, and the spatial resolution remained unchanged during the aggregation process. The reclassified data for 1990 and 2000 were used to calculate the extended enrichment factor and configure the initial states of the Logistic-LNCA model during the simulation periods.

3.3. Measuring large neighborhood effects and calibrating the urban CA model

When measuring neighborhood effects with the enrichment factor formula, the enrichment factors of all land-use types can be computed for all locations of a specific type or only for locations that have changed their land-use states/types (Van Vliet et al., 2013). We used the second approach to calculate the enrichment factors for each land-use type for locations that changed to industrial or urban residential land during the simulation period. We also determined the average enrichment factor through logistic transformation for comparison and curve drawing in the specific neighborhood scope.

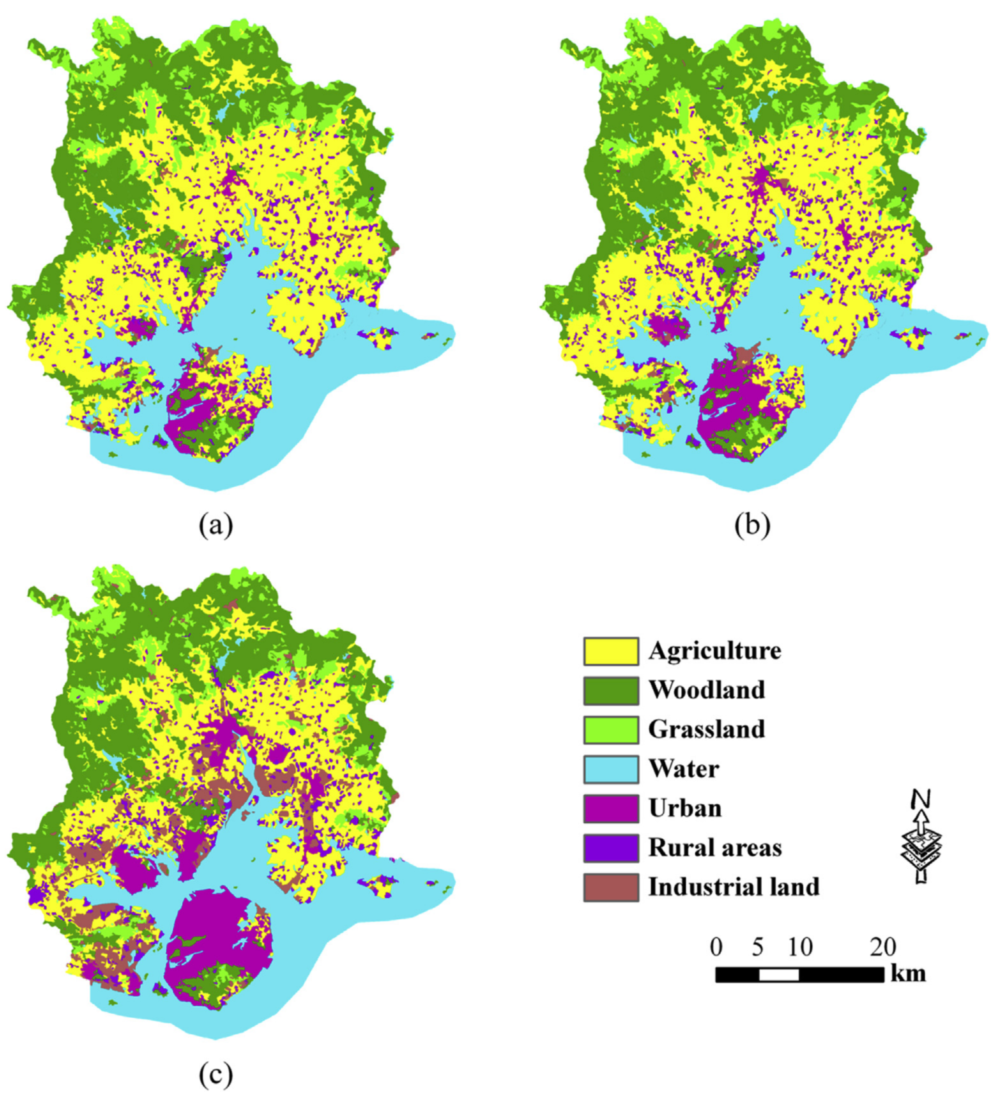

Fig. 3. Land-use patterns in Xiamen City in 1990 (a), 2000 (b), and 2010 (c). These maps were used to compute enrichment factors and verify large neighborhood effects.

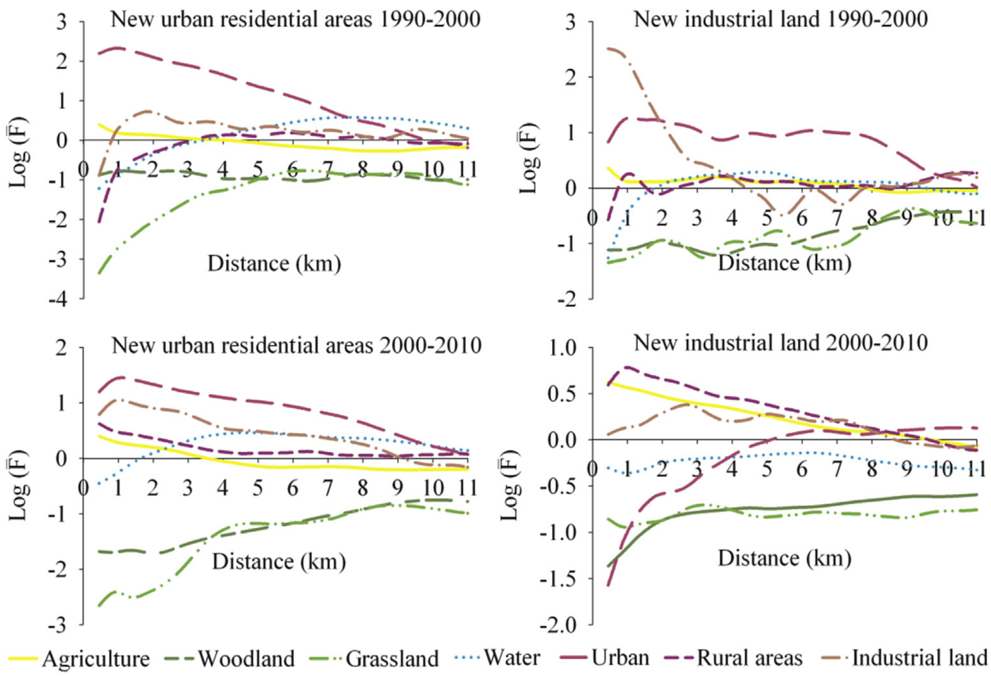

Fig. 4. Neighborhood characteristics (logarithm of the average extended enrichment factor, Log($\overline{F}_{k,l,\bigcup d_j}$)) as a function of distance for observed urban and industrial land-use changes during 1990–2000 and 2000–2010.

The extended enrichment factor of urban or industrial land near new residential land was greater than that of agriculture, woodland, grassland, water, or rural land, but its value decreased with distance (Fig. 4), suggesting that new urban and industrial land are more likely to emerge in the neighborhood of urban land use than other land-use types. The depicted enrichment factor curves indicated that the influences of neighboring land-use types decrease with increasing neighborhood radius. However, the measured extended enrichment factors for urban or industrial land in certain large neighborhood scopes (1–3 km, equal to 30–100 cells in the neighborhood radius) of new urban land still maintained relatively high values. The extended enrichment factor curves presented an undulating pattern. These neighborhood characteristics demonstrated that there was a considerably large neighborhood effect in

Table 1
Weights for various sub-neighborhoods at different distances in the large neighborhood model.

Simulation periods	Weights for sub-neighborhoods ($\bigcup d_j$)							
	β_0	β_1 [12,21]	β_2 [22,31]	β_3 [32,41]	β_4 [42,51]	β_5 [52,61]	β_6 [62,71]	β_7 [72,81]
1990–2000	−2.273	0.348	0.180	0.137	0.156	0.208	0.059	0.332
2000–2010	−0.720	0.337	−0.044	0.211	−0.056	0.105	0.051	0.465

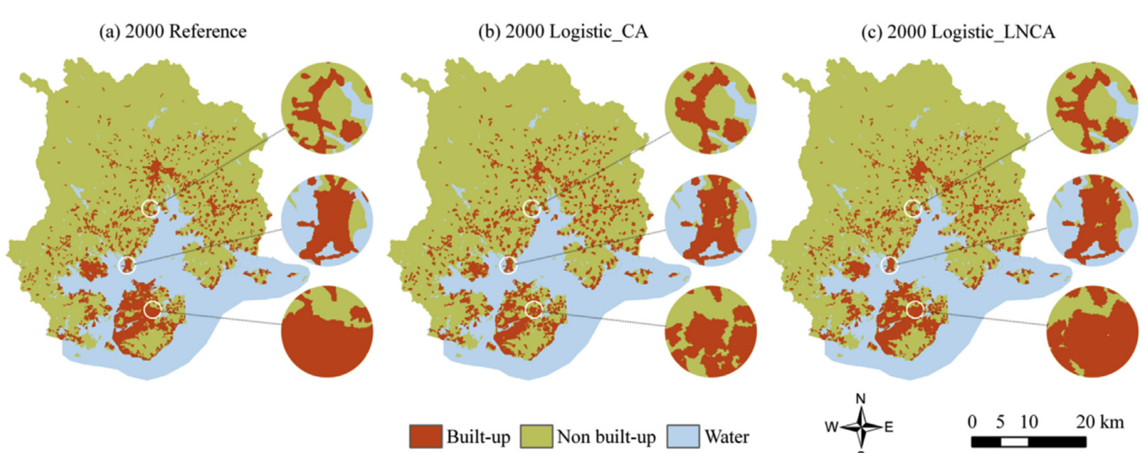

Fig. 5. Comparison of remotely sensed (a) and simulated (b/c) land-use datasets in 2000. Logistic-CA and Logistic-LNCA models were used for the land-use change simulations.

Table 2
Accuracy assessment of Logistic-LNCA and Logistic-CA simulations with confusion matrices generated with spatial overlays between remotely sensed reference maps and maps simulated with Logistic-LNCA or Logistic-CA.

2000		Reference		
		Built-up cells	Non built-up cells	User's accuracy (%)
Simulation with logistic-LNCA	Built-up cells	202,694	25,667	88.8
	Non built-up cells	25,126	1,358,395	98.2
	Producer's accuracy (%)	89.0	98.1	
Overall accuracy = 96.8%; Kappa coefficient = 0.87				
2000		Reference		
		Built-up cells	Non built-up cells	User's accuracy (%)
Simulation with logistic-CA	Built-up cells	195,838	32,523	85.8
	Non built-up cells	31,982	1,351,539	97.7
	Producer's accuracy (%)	86.0	97.7	
Overall accuracy = 96.0%; Kappa coefficient = 0.84				

cellular space. Compared with traditional small neighborhood windows that generally cover several cells (for instance, 1–4 cells, up to a 9 × 9 Moore neighborhood) in the neighborhood radius, it is more appropriate to construct the transition rules of the urban CA model using a large neighborhood.

To further estimate the influence of a particular large neighborhood, the values of Δr, $R_{\min}$, and $R_{\max}$ in Eq. (7) were set to 10, 11, and 81, respectively. Thus, the large neighborhood of the central cell was divided into seven annular bands, and the radius of each annular band was 10 cells. It is worth noting that the window size used here was visually obtained from Fig. 4, and its influence on the model results is still unknown. The large neighborhood module dynamically calculated the extended enrichment factors for the central cell at different annular bands. Similarly, whether the cell was to become built-up land was treated as the dependent variable, and 20% of the sample points were extracted through proportional random-stratified sampling. The obtained sample data were used to perform a binary logistic regression in SPSS. Eq. (9) was used to calibrate the model, and parameter values of the large neighborhood module were obtained.

Before executing the logistic regression calibration, the enrichment factor for the large neighborhood was normalized and the standardization was realized as:

$$F_{\text{norm}} = \frac{F_{\text{orig}} - F_{\min}}{F_{\max} - F_{\min}} \quad (12)$$

where F_{norm} is the normalized enrichment factor, ranging from 0 to 1. The obtained parameter value generally estimates the relationship between observed land-use changes and measured extended enrichment factors in a sub-neighborhood. Thus, specific contribution weights were given to developed cells falling in corresponding neighborhood scopes in the urban CA model (Table 1). The regions close to the central cell in the large neighborhood had the greatest impact on its development. However, the seventh ring scope still had a notable level of contribution.

3.4. Simulation results

The simulation of urban land-use changes was carried out with the Logistic-LNCA model by incorporating the probability map obtained by the large neighborhood module into a classic Logistic-CA model. The iteration number, conversion demand, and observation time during the simulation period are important considerations in urban CA simulations. An adequate iteration can help reveal the spatial details of local interactions and generate more accurate simulation results (Li and Yeh, 2004; Liu et al., 2008a). Experiments revealed that 400 iterations are sufficient to allow the Logistic-LNCA model to fully conduct the neighborhood calculation. The Logistic-CA model used the same iteration number when obtaining comparable simulated maps. As to the calculation of transition functions (Eq. (3) for Logistic-CA and Eq. (11) for Logistic-LNCA), the local development suitability (P_l) was updated only once at the start of the iterations; the 3 × 3 kernel neighborhood rule (P_Ω) and extended neighborhood rule (P_{ln}) were updated during each iteration; constraint factors ($con()$) remained unchanged

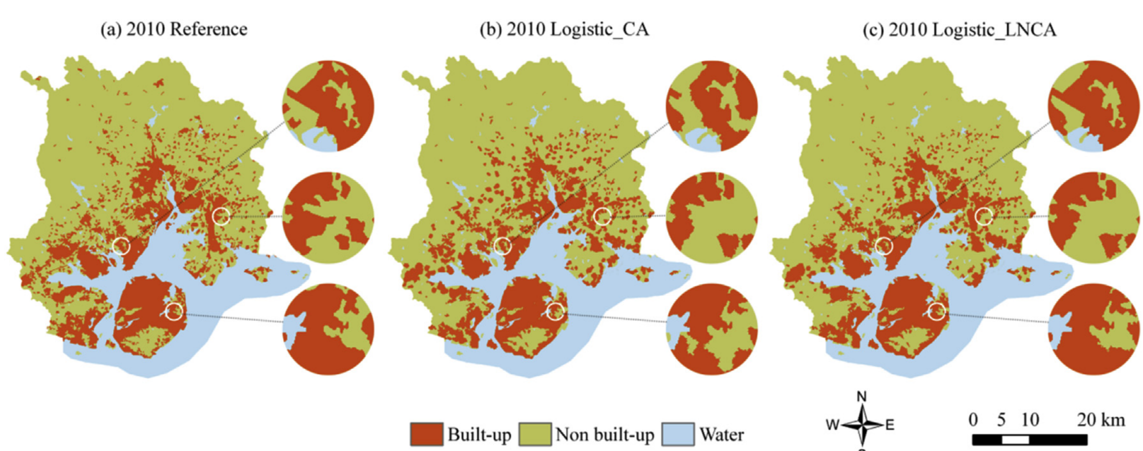

Fig. 6. Comparison of simulated urban growth patterns in Xiamen from 2000 to 2010 using the Logistic-CA and Logistic-LNCA models.

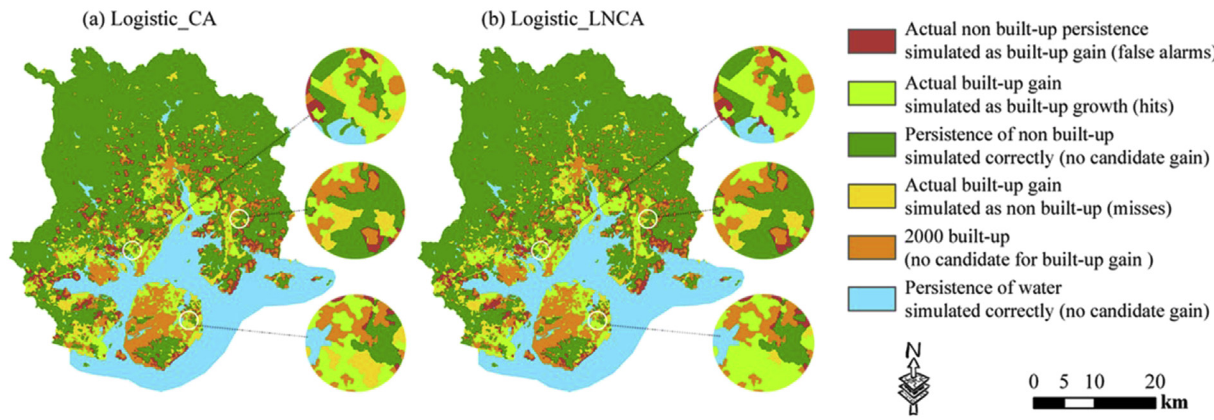

Fig. 7. Overlaid maps of observed and simulated built-up land expansions between 2000 and 2010 using the Logistic-CA and Logistic-LNCA models.

during model runs; and stochastic perturbations term (P_r) was also updated during each iteration.

Both the Logistic-CA and Logistic-LNCA models determined whether a location changed to built-up land based on the threshold proposed in Eq. (6). The determination of the threshold of conversion probability affects the conversion demand at each iteration and the linear relationship between the iteration number and observation time (Feng and Liu, 2013). We divided the total amount of land conversion during the simulation period by the iteration number to get the land demand converted at each iteration. Thus, each iteration produced a specific dynamic threshold of conversion probability. To reflect the uncertainty in urban expansion, the parameter α, which controls the size of random variable P_r in the random disturbance term, was set to 2 (White and Engelen, 1993). To simplify the calculation in Eq. (11), the large neighborhood scope ($R_{\min}$, $R_{\max}$), neighborhood subset interval Δr, and control factor δ remained unchanged during the simulation periods. In this study, radius 1 of 3 $\times$ 3 kernel neighborhood was set for $R_{\min}$ for model comparison; a neighborhood radius of 46 cells was set for the large neighborhood size $R_{\max}$, and a Euclidean distance of five cells was set as the neighborhood subset interval based on a trial and error process; and the control factor delta was set to 1 to reflect the baseline influences of the calibrated extended neighborhood rules. The large neighborhood size and interval here were also the optimal settings in Table 5.

Two urban CA models, Logistic-LNCA and Logistic-CA, were employed to simulate urban dynamics in Xiamen City, China. Using the land-use dataset in 1990 as the initial state, we simulated land-use change dynamics from 1990 to 2000 (Fig. 5). Based on a visual assessment, the three maps of Xiamen City in 2000 have many similarities, but visual analysis has obvious limitations because it is affected by the display scales and the analyst's experience.

3.5. Confusion matrix analysis

We overlaid the simulated results and reference land-use maps obtained from remote sensing images in 2000 and conducted a cell-by-cell comparison. We calculated the overall accuracies and Kappa coefficients of the two land-use models with the confusion matrix method (Liu et al., 2008a; Pontius et al., 2004) (Table 2).

The Logistic-LNCA model simulations resulted in land-use maps with an overall accuracy of 96.8% and a Kappa coefficient of 0.87 in 2000. The overall accuracy and Kappa coefficient values were both greater than those obtained with the Logistic-CA model. So were the accuracies for individual land-use types, either simulation- or reference-based (Table 2).

3.6. Independent validation

The validation of an environmental model requires observation/reference data that have not been used for the model's development. We configured the neighborhood rules at independent validation stages with parameter settings obtained from the model calibration during 1990–2000. We then used both the Logistic-CA and Logistic-LNCA models to simulate urban expansion from 2000 to 2010 under the same transition rules as in the model calibration period. The predicted urban dynamics in 2010 are shown in Fig. 6. Visual comparison indicated that the simulated pattern resulting from the Logistic-LNCA model is more similar to the reference map in 2010 than that from the Logistic-CA model.

Table 3
The cell-by-cell comparison accuracies of the Logistic-CA and Logistic-LNCA models during the validation period from 2000 to 2010.

2010		Reference		
		Built-up cells	Non built-up cells	User's accuracy (%)
Simulation with logistic-LNCA	Built-up cells	371,016	80,531	82.2
	Non built-up cells	83,825	1,118,044	93.0
	Producer's accuracy (%)	81.6	93.3	
Overall accuracy = 90.1%; Kappa coefficient = 0.75				
2010		Reference		
		Built-up cells	Non built-up cells	User's accuracy (%)
Simulation with logistic-CA	Built-up cells	353,521	98,026	78.3
	Non built-up cells	101,320	1,100,549	91.6
	Producer's accuracy (%)	77.7	91.8	
Overall accuracy = 87.9%; Kappa coefficient = 0.70				

Table 4
Spatial metrics of compared observed and predicted changes during the validation period from 2000 to 2010 (Bennett et al., 2013).

Name	Logistic-CA	Logistic-LNCA	Description	Equation
Accuracy (fraction correct)	0.906	0.923	It is heavily influenced by the most common category, usually "no event".	$\frac{hits+correct\ negatives}{total}$
Bias score (frequency bias)	0.985	0.985	Measures the ratio of the frequency of modeled events to that of observed events. Indicates whether the model has a tendency to underestimate (BIAS < 1) or overestimate (BIAS > 1).	$\frac{hits+false\ alarms}{hits+misses}$
Probability of detection (hit rate)	0.554	0.631	Sensitive to hits, but ignores false alarms. Good for rare events.	$\frac{hits}{hits+misses}$
False alarm ratio	0.438	0.360	Sensitive to false alarms, but ignores misses.	$\frac{false\ alarms}{hits+false\ alarms}$
Probability of false detection (false alarm rate)	0.069	0.057	Sensitive to false alarms, but ignores misses.	$\frac{false\ alarms}{correct\ negatives+false\ arlams}$
Threat score (critical success index, CSI)	0.387	0.466	Measures the fraction of observed cases that were correctly modeled. Penalizes both misses and false alarms.	$\frac{hits}{hits+misses+false\ alarms}$
Success index	0.627	0.670	Equally weights the ability of the model to correctly detect occurrences and non-occurrences of events.	$\frac{1}{2}\left(\frac{hits}{hits+misses}+\frac{correct\ negatives}{observed\ no}\right)$

Each of the two simulated maps was overlapped with the reference map in 2010 (Fig. 7), and the confusion matrices were employed to calculate simulation accuracies and Kappa coefficients (Table 3). Producer's and user's accuracies of built-up land each increased by 3.9% between the predictions with the Logistic-CA and Logistic-LNCA models. The Logistic-LNCA model improved the overall accuracy by 2.2% and the Kappa coefficient by 0.05 compared with the Logistic-CA model. This indicates that the consideration of large neighborhood effects improved the simulation capability of the urban model.

We also used the figure of merit method to measure the hit and error distribution of the simulated maps (Chen and Pontius, 2010). This method limits the statistical data mainly to those cells in the simulated and observed maps that change during the simulation periods (García et al., 2012; Santé et al., 2010).

The predicted correct and non-correct cells during the simulated period from 2000 to 2010 were classified into six categories (Fig. 7): no candidate gain for built-up land, non built-up land, water, hits, misses and false alarms. Based on this classification, we executed quantitative evaluations with seven spatial metrics (Bennett et al., 2013) (Table 4). A comparison of these indicators demonstrated that the simulated map generated by the Logistic-LNCA model obtained a higher probability of detection (hit rate) and a lower false alarm ratio and probability of false detection (false alarm rate) than the Logistic-CA model. In addition, the Logistic-LNCA model achieved greater threat score (critical success index or CSI) and success index values.

The simulation iterations were set from 10 to 600 and both models ran repeatedly with all other conditions remaining unchanged. The simulation accuracies of the two models showed substantial improvement with increases in iterations from 10 to 200 and remained relatively unchanged beyond 200 iterations (Fig. 8). Moreover, the differences in the simulation accuracies between the two models grew with increases in iterations. This suggests that increases in iterations can help the large neighborhood module fully simulate the impact of neighborhood interactions on urban development and result in more reliable spatial patterns of urban expansion.

3.7. Sensitivity and performance analysis

Changes in neighborhood window size can have a significant impact on the performance of an environmental model (Tang et al., 2012). In order to compare with the 3 × 3 kernel neighborhood model (Logistic-CA), R_{min} in Eq. (7) was set to 1 during parameter sensitivity analysis. Thus the extended neighborhood module has two adjustable parameters, window size and radius interval, which refer to R_{max} and Δr in Eq. (7), respectively. A trial and error method was used to determine the values of these two parameters during the parameter sensitivity analysis. The extended neighborhood rule required recalibrated according to the methods proposed in Sections 2.1 and 2.3 after the values of the two parameters changed. Table 5 lists the effects of several different window sizes and radius intervals on the simulation accuracies and Kappa coefficients, and simulation results of the Logistic-CA model. When the window size was equal to 46 cells and the radius interval is equal to five cells, the Logistic-LNCA model achieved the highest accuracy based on the tested parameters. The simulation results were quite different when these two parameters were set to other values, but they were generally more accurate than the simulations generated by the 3 × 3 kernel neighborhood model. The parameter sensitivity analysis indicated that the parameter settings of a larger neighborhood have a greater impact on simulation results. Therefore, there must be an optimal window size and corresponding parameter settings in cellular space.

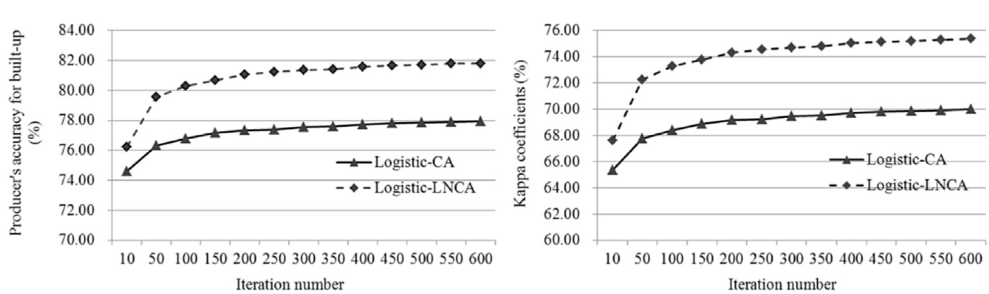

Fig. 8. Comparison of the changes in simulation accuracies and Kappa coefficients with the increase in iterations during the validation period from 2000 to 2010 between the Logistic-CA and Logistic-LNCA models.

Table 5
Impacts of window size (WS) and radius interval (RI) on simulation results of the Logistic-LNCA model and simulation results of the 3 × 3 kernel neighborhood model (Logistic-CA).

	1990–2000			2000–2010		
	Built-up (%)	Overall (%)	Kappa	Built-up (%)	Overall (%)	Kappa
WS = 5, RI = 1	86.7	96.2	0.84	78.7	88.5	0.71
WS = 9, RI = 2	87.3	96.4	0.85	79.2	88.8	0.72
WS = 21, RI = 4	88.3	96.7	0.86	80.1	89.3	0.73
WS = 46, RI = 5	89.0	96.8	0.87	81.6	90.1	0.75
WS = 91, RI = 10	88.6	96.8	0.87	81.0	89.8	0.74
WS = 151, RI = 15	88.5	96.7	0.86	80.3	89.4	0.73
Logistic-CA (3 × 3)	86.0	96.0	0.84	77.7	87.9	0.70

4. Discussion

Various evaluations indicated that the incorporation of a large neighborhood was an effective approach to improve the urban CA model's simulation performance. To some extent, the extended enrichment factor proposed in this paper can be considered as a generalization of the enrichment factor proposed by Verburg et al. (2004b), and it is suitable to detect larger neighborhood scopes at a specific resolution. Thus, this research may have potential contributions to improve other cellular-based urban models. However, the extended enrichment factor was mainly used to calibrate an urban expansion model during specific simulation periods, which only considered over- and under-representations of built-up land in the neighborhood. A meaningful simulation practice would be to incorporate the extended enrichment factor into a land-use urban CA model and examine the corresponding neighborhood effects.

Simulation accuracy is an important indicator to evaluate the performance of urban CA models but its values obtained from the results of different models, study areas, and simulation periods can differ significantly (Liu et al., 2008a, 2012; Santé et al., 2010). In this study, the Logistic-LNCA model resulted in a 3.0% increase in the mean producer accuracy and user accuracy of built-up land during the calibration periods (1990–2000) and a 3.9% increase in the validation period (2000–2010) compared with the Logistic-CA model. These differences indicated that the urban CA model with a large neighborhood module could better simulate urban expansion. Although the overall accuracies of the two simulation periods were relatively high (96.8% and 90.1%, respectively), their improvement was slight (0.8% and 2.2%, respectively). This minor improvement might indicate the limited effectiveness of large-neighborhood considerations or could be due to the fact that built-up land was not the dominant land-use type in the study area. In addition to the assessments above, we also performed error budget assessment (Pontius et al., 2004) and enrichment factor evaluation (Van Vliet et al., 2013). These two assessments also indicated that the Logistic-LNCA model was superior to the Logistic-CA model. Detailed information about these assessments is available from the authors upon request.

It is essential to ensure adequate local interactions to achieve accurate urban expansion simulations. Some studies have run CA models at yearly time intervals, for which the land conversion quantity in one year was constrained to the land demand at each iteration (Wu, 2002). Thus, if the time lag between two land-use coverage values is 5, 10, or 20 years, then the CA simulation will proceed on a time step equal to corresponding yearly intervals. As depicted by the curves in Fig. 8, urban CA models cannot obtain optimum simulation results. However, some studies have suggested that the discrete simulation time used in urban CA models is different from continuous real time (Yeh and Li, 2006). Urban CA models are a 'bottom–up' approach based on a complex system created by the interactions of simple subsystems (Liu et al., 2007). Local interactions are important for generating realistic urban morphology and adequate time steps are necessary to resolve spatial details during simulations (Yeh and Li, 2006). Therefore, we used 400 iterations for our urban-expansion simulations.

5. Conclusions

We developed a logistic regression urban CA model by considering the effects of a relatively large surrounding cellular space or neighborhood (Logistic-LNCA). We extended the definition of the enrichment factor to characterize the neighborhood effects of various sub-neighborhoods at different distances within a large neighborhood. The curves based on the extended enrichment factors and land-use changes during the simulation period demonstrated a strong neighborhood effect at relatively long distances (1–3 km, equal to 30–100 cells in the neighborhood radius) from the center. Both the Logistic-LNCA model and the traditional urban CA model (Logistic-CA) were used to simulate the spatio-temporal processes of urban expansion in Xiamen City, China for two periods of time: 1990–2000 and 2000–2010. The data from the first period were used for model calibration whereas the data from the second period were used for model validation. The simulation results showed that the Logistic-LNCA model could achieve higher accuracy values and Kappa coefficients than the Logistic-CA model during both the calibration and validation periods. A source budget analysis of the error budget method also supported the fact that the large neighborhood model could reconstruct historical spatial patterns with higher agreement with the reference land-use map.

This simulation exercise indicated that the window size and radius interval values in the large neighborhood module have important effects on the performance of the Logistic-LNCA model. The values of these two parameters were determined through a trial-and-error method after setting parameter boundaries. The simulation results varied with window sizes and radius intervals, and their accuracies were generally better than the accuracy of the 3 × 3 kernel neighborhood model according to parameter sensitivity analysis. Over- and under-representation of built-up land in the neighborhood of new built-up land for observed and simulated land-use changes during the calibration and validation periods were further investigated. Results showed that the enrichment factors measured from simulated maps of the Logistic-LNCA model were more agreeable to the enrichment factors of observed land-use changes. The execution of the large neighborhood model required up to as twice computation power as that of the Logistic-CA model.

As the large neighborhood method is suitable for logistic-based CA modeling, it is reasonable to believe that such an approach may be also be useful for other types of urban CA models such as the CLUE-S model (Verburg et al., 2002), LUCIA model (Hansen, 2007), or SLEUTH model (Clarke and Gaydos, 1998). It would be interesting to investigate the large neighborhood effects associated with various urban CA models based on artificial intelligence and evolutionary computation. Because the study area in this research was a typical coastal city, the large neighborhood model needs to be tested in other geographical areas to further investigate the performance and application of the Logistic-LNCA model. It would also be useful to incorporate the large neighborhood method into urban simulation models to examine multiple land-use changes and predict their future states.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (41101143).

References

Al-Ahmadi, K., See, L., Heppenstall, A., Hogg, J., 2009. Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia. Ecol. Complex. 6 (2), 80–101.
Arsanjani, J.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.
Batty, M., Xie, Y., 1994. From cells to cities. Environ. Plan. B Plan. Des. 21, 531–548.
Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40, 1–20.
Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, New York.
Chen, B., Xu, H., 2005. Urban expansion and its driving force analysis using remote sensed data a case of Xiamen city. Econ. Geogr. 25 (1), 79–83.
Chen, H., Pontius, R.G., 2010. Sensitivity of a land change model to pixel resolution and precision of the independent variable. Environ. Model. Assess. 16 (1), 37–52.
Cheng, J., Masser, I., 2003. Urban growth pattern modeling: a case study of Wuhan city, PR China. Landsc. Urban Plan. 62 (4), 199–217.
Clarke, K., Hoppen, S., Gaydos, L., 1997. A self-modifying cellular automaton model of historical. Environ. Plan. B 24, 247–261.
Clarke, K.C., Gaydos, L.J., 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. Int. J. Geogr. Inf. Sci. 12 (7), 699–714.
Couclelis, H., 1988. Of mice and men: what rodent populations can teach us about complex spatial dynamics. Environ. Plan. A 20 (1), 99–109.
Dai, E., Wu, S., Shi, W., Cheung, C.-k, Shaker, A., 2005. Modeling change-pattern-value dynamics on land use: an integrated GIS and artificial neural networks approach. Environ. Manag. 36 (4), 576–591.
Dendoncker, N., Rounsevell, M., Bogaert, P., 2007. Spatial analysis and modelling of land use distributions in Belgium. Comput. Environ. Urban Syst. 31 (2), 188–205.
Feng, Y., Liu, Y., 2013. A heuristic cellular automata approach for modelling urban land-use change based on simulated annealing. Int. J. Geogr. Inf. Sci. 27 (3), 449–466.
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.
Fuglsang, M., Münier, B., Hansen, H.S., 2013. Modelling land-use effects of future urbanization using cellular automata: an eastern Danish case. Environ. Model. Softw. 50, 1–11.
García, A.M., Santé, I., Boullón, M., Crecente, R., 2012. A comparative analysis of cellular automata models for simulation of small urban areas in Galicia, NW Spain. Comput. Environ. Urban Syst. 36 (4), 291–301.
Guan, D., Li, H., Inohae, T., Su, W., Nagaie, T., Hokao, K., 2011. Modeling urban land use change by the integration of cellular automaton and Markov model. Ecol. Model. 222 (20), 3761–3772.
Hagoort, M., Geertman, S., Ottens, H., 2008. Spatial externalities, neighbourhood rules and CA land-use modelling. Ann. Reg. Sci. 42 (1), 39–56.
Hansen, H.S., 2007. An adaptive land-use simulation model for integrated coastal zone planning. In: The European Information Society. Springer, pp. 35–53.
Hansen, H.S., 2008. Quantifying and analysing neighbourhood characteristics supporting urban land-use modelling. In: The European Information Society. Springer, pp. 283–299.
Hansen, H.S., 2012. Empirically derived neighbourhood rules for urban land-use modelling. Environ. Planning-Part B 39 (2), 213.
Harrop, K.J., 1973. Nuisances and Their Externality Fields. Department of Geography, University of Newcastle Upon Tyne.
Hewitt, R., Van Delden, H., Escobar, F., 2014. Participatory land use modelling, pathways to an integrated approach. Environ. Model. Softw. 52, 149–165.
Hirzel, A., Guisan, A., 2002. Which is the optimal sampling strategy for habitat suitability modelling. Ecol. Model. 157 (2), 331–341.
Hu, Z., Lo, C., 2007. Modeling urban growth in Atlanta using logistic regression. Comput. Environ. Urban Syst. 31 (6), 667–688.
Huang, B., Zhang, L., Wu, B., 2009. Spatiotemporal analysis of rural–urban land conversion. Int. J. Geogr. Inf. Sci. 23 (3), 379–398.
Kocabas, V., Dragicevic, S., 2007. Enhancing a GIS cellular automata model of land use change: Bayesian networks, influence diagrams and causality. Trans. GIS 11 (5), 681–702.
Krugman, P., 1999. The role of geography in development. Int. Reg. Sci. Rev. 22 (2), 142–161.
Lagarias, A., 2012. Urban sprawl simulation linking macro-scale processes to micro-dynamics through cellular automata, an application in Thessaloniki, Greece. Appl. Geogr. 34, 146–160.
Lauf, S., Haase, D., Hostert, P., Lakes, T., Kleinschmit, B., 2012. Uncovering land-use dynamics driven by human decision-making–a combined model approach using cellular automata and system dynamics. Environ. Model. Softw. 27, 71–82.
Li, X., Yeh, A.G.-O., 2000. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. Int. J. Geogr. Inf. Sci. 14 (2), 131–152.
Li, X., Yeh, A.G.-O., 2002. Integration of principal components analysis and cellular automata for spatial decisionmaking and urban simulation. Sci. China Ser. D Earth Sci. 45 (6), 521–529.
Li, X., Yeh, A.G.-O., 2004. Data mining of cellular automata's transition rules. Int. J. Geogr. Inf. Sci. 18 (8), 723–744.
Li, X., Liu, X., Yu, L., 2014. A systematic sensitivity analysis of constrained cellular automata model for urban growth simulation based on different transition rules. Int. J. Geogr. Inf. Sci. 28 (7), 1317–1335.
Liao, J., Tang, L., Shao, G., Qiu, Q., Wang, C., Zheng, S., Su, X., 2014. A neighbor decay cellular automata approach for simulating urban expansion based on particle swarm intelligence. Int. J. Geogr. Inf. Sci. 28 (4), 720–738.
Liu, X., Li, X., Yeh, A.G.-O., He, J., Tao, J., 2007. Discovery of transition rules for geographical cellular automata by using ant colony optimization. Sci. China Ser. D Earth Sci. 50 (10), 1578–1588.
Liu, X., Li, X., Liu, L., He, J., Ai, B., 2008a. A bottom-up approach to discover transition rules of cellular automata using ant intelligence. Int. J. Geogr. Inf. Sci. 22 (11–12), 1247–1269.
Liu, X., Li, X., Shi, X., Wu, S., Liu, T., 2008b. Simulating complex urban development using kernel-based non-linear cellular automata. Ecol. Model. 211 (1–2), 169–181.
Liu, X., Ma, L., Li, X., Ai, B., Li, S., He, Z., 2014. Simulating urban growth by integrating landscape expansion index (LEI) and cellular automata. Int. J. Geogr. Inf. Sci. 28 (1), 148–163.
Liu, Y., Yang, S., Chen, J., 2012. Modeling environmental impacts of urban expansion: a systematic method for dealing with uncertainties. Environ. Sci. Technol. 46 (15), 8236–8243.
Mas, J.-F., Kolb, M., Paegelow, M., Camacho Olmedo, M.T., Houet, T., 2014. Inductive pattern-based land use/cover change models: a comparison of four software packages. Environ. Model. Softw. 51, 94–111.
McCullagh, P., Nelder, J., 1989. Generalized Linear Models. CRC Press, Boca Raton.
McMillen, D.P., 1989. An empirical model of urban fringe land use. Land Econ. 138–145.
Munroe, D.K., Southworth, J., Tucker, C.M., 2004. Modeling spatially and temporally complex land-cover change: the case of western Honduras*. Prof. Geogr. 56 (4), 544–559.
Pan, Y., Roth, A., Yu, Z., Doluschitz, R., 2010. The impact of variation in scale on the behavior of a cellular automata used for land use change modeling. Computers. Environ. Urban Syst. 34 (5), 400–408.
Pontius, J., Robert, G., Huffaker, D., Denman, K., 2004. Useful techniques of validation for spatially explicit land-change models. Ecol. Model. 179 (4), 445–461.
Rodrigue, J.P., 2004. Transport Geography on TheWeb. Department of Economics and Geography, Hofstra University, New York.
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.
Shaker, R.R., 2015. The well-being of nations: an empirical assessment of sustainable urbanization for Europe. Int. J. Sustain. Dev. World Ecol. 22 (5), 375–387.
Stevens, D., Dragicevic, S., Rothley, K., 2007. *iCity*: a GIS–CA modelling tool for urban planning and decision making. Environ. Model. Softw. 22 (6), 761–773.
Takeyama, M., Couclelis, H., 1997. Map dynamics: integrating cellular automata and GIS through geo-algebra. Int. J. Geogr. Inf. Sci. 11 (1), 73–91.
Tang, L., Zhao, Y., Yin, K., Zhao, J., 2013. In: City Profile: Xiamen. Cities, 31, pp. 615–624.
Tang, L., Su, X., Shao, G., Zhang, H., Zhao, J., 2012. A Clustering-Assisted Regression (CAR) approach for developing spatial climate data sets in China. Environ. Model. Softw. 38, 122–128.
Tobler, W., 1979. Cellular geography. In: Philosophy in Geography. Springer, pp. 379–386.
Van Vliet, J., Naus, N., Van Lammeren, R.J., Bregt, A.K., Hurkens, J., Van Delden, H., 2013. Measuring the neighbourhood effect to calibrate land use models. Comput. Environ. Urban Syst. 41, 55–64.
Verburg, P.H., Van Eck, J.R., de Nijs, T.C., Dijst, M.J., Schot, P., 2004a. Determinants of land-use change patterns in the Netherlands. Environ. Plan. B 31 (1), 125–150.
Verburg, P.H., de Nijs, T., Ritsema van Eck, J., Visser, H., de Jong, K., 2004b. A method to analyse neighbourhood characteristics of land use patterns. Comput. Environ. urban Syst. 28 (6), 667–690.
Verburg, P.H., Soepboer, W., Veldkamp, A., Limpiada, R., Espaldon, V., Mastura, S.S.,

2002. Modeling the spatial dynamics of regional land use: the CLUE-S model. Environ. Manag. 30 (3), 391–405.
Verstegen, J.A., Karssenberg, D., Van Der Hilst, F., Faaij, A.P., 2014. Identifying a land use change cellular automaton by Bayesian data assimilation. Environ. Model. Softw. 53, 121–136.
Wang, H., He, S., Liu, X., Dai, L., Pan, P., Hong, S., Zhang, W., 2013. Simulating urban expansion using a cloud-based cellular automata model: a case study of Jiangxia, Wuhan, China. Landsc. Urban Plan. 110, 99–112.
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 (8), 1175–1199.
White, R., Engelen, G., 2000. High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Comput. Environ. urban Syst. 24 (5), 383–400.
White, R., Engelen, G., 2003. A calibration procedure for constrained large neighbourhood cellular automata based land use models. In: 13th European Colloquium on Theoretical and Quantitative Geography, Lucca, Italy.
White, R., Engelen, G., Uljee, I., 1997. The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. Environ. Plan. B 24, 323–344.
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. Plan. B 25, 103–126.
XCSB, 2008. Xiamen City Statistics Bureau and National Bureau of Statistics of China (XCSB). In: Yearbook of Xiamen Special Economic Zone in 2008. China Statistics Press, Beijing. http://www.stats-xm.gov.cn/2008/.
XCSB, 2011. Xiamen City Statistics Bureau and National Bureau of Statistics of China (XCSB). In: Yearbook of Xiamen Special Economic Zone in 2011. China Statistics Press, Beijing. http://www.stats-xm.gov.cn/2011/.
Xie, C., Huang, B., Claramunt, C., Chandramouli, C., 2005. Spatial logistic regression and gis to model rural-urban land conversion. In: Proceedings of PROCESSUS Second International Colloquium on the Behavioural Foundations of Integrated Land-use and Transportation Models: Frameworks, Models and Applications, pp. 12–15.
Yang, Q., Li, X., Shi, X., 2008. Cellular automata for simulating land use changes based on support vector machines. Comput. Geosci. 34 (6), 592–602.
Yeh, A.G.-O., Li, X., 2006. Errors and uncertainties in urban cellular automata. Comput. Environ. Urban Syst. 30 (1), 10–28.