International Journal of Geographical Information Science

ISSN: 1365-8816 (Print) 1362-3087 (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Coupling fuzzy clustering and cellular automata based on local maxima of development potential to model urban emergence and expansion in economic development zones

Xun Liang, Xiaoping Liu, Guangliang Chen, Jiye Leng, Youyue Wen & Guangzhao Chen

To cite this article: Xun Liang, Xiaoping Liu, Guangliang Chen, Jiye Leng, Youyue Wen & Guangzhao Chen (2020): Coupling fuzzy clustering and cellular automata based on local maxima of development potential to model urban emergence and expansion in economic development zones, International Journal of Geographical Information Science, DOI: 10.1080/13658816.2020.1741591

To link to this article: https://doi.org/10.1080/13658816.2020.1741591

Published online: 23 Mar 2020.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
https://doi.org/10.1080/13658816.2020.1741591

RESEARCH ARTICLE

Coupling fuzzy clustering and cellular automata based on local maxima of development potential to model urban emergence and expansion in economic development zones

Xun Liang [a,b], Xiaoping Liu [a], Guangliang Chen [c], Jiye Leng [a,d], Youyue Wen [e] and Guangzhao Chen [a]

[a]Guangdong Key Laboratory for Urbanization and Geo-simulation, School of Geography and Planning,Sun Yat-sen University, Guangzhou, PR China; [b]School of Geography and Information Engineering, China University of Geosciences, Wuhan, Hubei Province, China; [c]Guangzhou Landtu Geographic Information Technology Co., Ltd., Guangzhou, PR China; [d]Department of Geography and Planning, University of Toronto, Toronto, ON, Canada; [e]South China Institute of Environmental Science. MEE, Guangzhou, China

ABSTRACT

Modeling urban growth in Economic development zones (EDZs) can help planners determine appropriate land policies for these regions. However, sometimes EDZs are established in remote areas outside of central cities that have no historical urban areas. Existing models are unable to simulate the emergence of urban areas without historical urban land in EDZs. In this study, a cellular automaton (CA) model based on fuzzy clustering is developed to address this issue. This model is implemented by coupling an unsupervised classification method and a modified CA model with an urban emergence mechanism based on local maxima. Through an analysis of the planning policies and existing infrastructure, the proposed model can detect the potential start zones and simulate the trajectory of urban growth independent of the historical urban land use. The method is validated in the urban emergence simulation of the Taiping Bay development zone in Dalian, China from 2013 to 2019. The proposed model is applied to future simulation in 2019–2030. The results demonstrate that the proposed model can be used to predict urban emergence and generate the possible future urban form, which will assist planners in determining the urban layout and controlling urban growth in EDZs.

ARTICLE HISTORY
Received 29 July 2018
Accepted 8 March 2020

KEYWORDS
Economic development zone; urban emergence simulation; fuzzy C-means algorithm; cellular automata

1. Introduction

Urban development is considered one of the most widely discussed issues in urban studies (Lagarias 2012, Li *et al*. 2019, Liao *et al*. 2017a, Liu *et al*. 2018). Along with socioeconomic development and population growth, urban sprawl has recently been very intense in many regions, resulting in an increase in infrastructure costs (Tayyebi *et al*. 2014), the loss of cropland (Amour *et al*. 2017), unnecessary land resource usage and energy consumption (Chen *et al*. 2013, Liao *et al*. 2017b), urban heat wave (Liao *et al*. 2018),and the degradation of ecosystems (Sohl *et al*. 2012). Lack of planning and land use policies is one of the reasons for the uncontrolled development of urban land use

CONTACT Xiaoping Liu liuxp3@mail.sysu.edu.cn

(Lagarias 2012) because urban expansion is not only driven by factors such as social-cultural characteristics, accessibility, and terrain but is also associated with decision-making on the allocation of new urban areas (Lagarias 2012). To maintain the sustainable development of society and human-land coordination, planners should make exact judgments on the extent of urban growth (He *et al*. 2018).

Urban expansion models are commonly adapted to support urban planning and management at different scales (Arsanjani *et al*. 2018, Guy *et al*. 1997, White and Engelen 2000, Barredo *et al*. 2003, Li and Liu 2006, Liu *et al*. 2008, Jjumba and Dragićević 2012, Long *et al*. 2012), which is essential for rapidly developing region (Xue *et al*. 2016, Inkoom *et al*. 2017, Yao *et al*. 2017). Many current studies address the issue of urban expansion using simulation modeling based on cellular automata (CA) (Chen *et al*. 2014, Kamusoko and Gamba 2015, Feng *et al*. 2016). CA models have been widely applied in projecting future urban expansion patterns because they can effectively simulate complex geographical processes through local rules (Al-Ahmadi *et al*. 2009).

Over the past two decades, various kinds of CA models have been developed for simulating the urban growth processes (Arsanjani *et al*. 2018, Bren D, Amour *et al*. 2017, Deal and Schunk 2004). Among these studies, CA models based on the analysis of historical urban forms and various driving factors can provide insights into the nature of urban land-use dynamics (Guy *et al*. 1997, Li and Yeh 2002, Kamusoko and Gamba 2015). A large part of CA studies determines the transition rules or transition potential by using intelligence algorithms. These empirical transition potential models include logistic regression (Tayyebi *et al*. 2014), random forests (Kamusoko and Gamba 2015), and artificial neural networks (Li and Yeh 2001). These kinds of CA models can help researchers obtain robust explanations of land-use patterns (Sohl *et al*. 2007) and are easy to implement (Verburg *et al*. 2004). In recent years, many researchers have focused on integrated approaches that combine top-down and bottom-up dynamics in land-use modeling (Sohl and Sayler 2008). These integrated models are constructed by coupling a CA model with various techniques, such as system dynamics (Sohl and Sayler 2008), multisection models (Li *et al*. 2017, Dong *et al*. 2018), and Markov chains (Arsanjani *et al*. 2011). The development trends of the individual land use types and the local-scale dynamics are both addressed in the integrated models (Verburg and Overmars 2009). These models can synthesize various interactions, ranging from environmental to socioeconomic factors at different spatial-temporal scales (He *et al*. 2006). Macroscale and local-scale planning policies under different scenarios can be addressed in the integrated models, and thus, the impact of policies on urban growth can be understood (Barredo *et al*. 2003, Huang *et al*. 2014). For example, a CA-based future land use simulation (FLUS) model that combines the top-down allocation of land use change to grid cells with a bottom-up determination of conversions for specific land use transitions was proposed to simulate land cover change at various scales and serve multiple purposes (Li *et al*. 2017, Liang *et al*. 2018b, Liu *et al*. 2017). Various aspects of urban growth have been explored by previous CA studies, such as land use change patterns (Chen *et al*. 2016), urban growth boundaries (Liang *et al*. 2018a), changes in animal habitats (He *et al*. 2017), landscape connectivity (Huang *et al*. 2018a, Huang *et al*. 2018b), agricultural land loss (Amour *et al*. 2017), urban renewal (Zheng *et al*. 2015) and the effects of various planning policies on urban growth (Liang *et al*. 2018b, Shu *et al*. 2017).

Although widely used in urban dynamics simulation, previous CA models can only simulate the urban change in a region that has already been under development for a certain period because future urbanization of these regions will most likely develop around already developed high-density urban land or large urban clusters (Fragkias and Seto 2009). They rely on the availability of sufficient historical land use data in the study region to obtain neighborhood effects and learn the urban development potential (Long *et al.* 2012). Therefore, these models are unable to simulate the urban development and urban form in the places that have no historical urban land. For example, the urban emergence and expansion in some economic development zones (EDZs) that far away from central cities.

Although urban development potential still can be obtained by analyzing the current driving factors with the analytic hierarchy process (AHP) (Wu 1998) without considering historical urban land. However, AHP involves human subjectivity, which necessitates the use of decision-making under uncertainty and introduces vagueness into the model process (Tesfamariam and Sadiq 2006). Another way to solve this problem is using unsupervised algorithms that only analyze the driving factors without relying on the training labels (e.g., historical urban land) to discover the development potential or transition rules of urban growth. However, there are no past studies applying unsupervised methods to modeling urban growth. Although Omrani *et al.* (2019) proposed an LTM-cluster framework that combined the LTM model and K-means cluster to simulate land use change. However, the cluster method in this study is a zonal tool, which splits the input data into several clusters for handling mass simulation data. The simulation procedure in this study is still carried out by the Artificial Neural Network-based LTM model. In addition, most of the clustering methods (e.g., K-means, ISODATA, hierarchical clustering) can only obtain the classification label of the land use types. Their clustering results lost gradient information and are unable to couple with simulation models. Therefore, these unsupervised classification methods are commonly used in land use classification studies (Li *et al.* 2018) but rare in simulation researches. To simulate urban emergence in the places with no historical urban land, CA models should be integrated with a clustering method which considers the memberships between various classifications and simultaneously assigns the memberships of different categories to each cell.

In addition, only generating urban development potential without using historical urban land is not enough to simulate the urban emergence (e.g., AHP-CA). Because traditional CA models are based on neighborhood effects, they can't simulate the urban emergence in a region that almost has no historical urban area and the neighborhood effect are zero across the whole region. Therefore, a CA model that can find the potential urban development points is also needed for simulating urban emergence. This paper proposes a clustering-based FLUS (CFLUS) model that couples an improved FLUS model with an unsupervised fuzzy clustering algorithm (FCM) to simulate urban growth in areas without historical urban land use. Through the analysis of various planning traffic networks and current driving factors (e.g., historical traffic networks) with a fuzzy C-means clustering algorithm. The improved FLUS model includes an urban emergence mechanism based on local maxima of urban development potential, which can simulate the occurrence of new urban patches. This method can identify the most likely hotspot areas in a location with no historical urban land and thus can project the emergence and development of urban land in the EDZs. This study provides a new method for assisting urban designers to determine the start zones and final urban form for the EDZs, aiming to prevent poor urban designs (Clarke 2014) in EDZs, which can't be done by previous models.

2. Methods

The proposed CFLUS model is an adapted version of the FLUS model. The original FLUS model is characterized by combining artificial neural networks (ANN) with a CA model that has a self-adaptive inertia competition mechanism (Li *et al*. 2017, Liang *et al*. 2018a, Liu *et al*. 2017). To model the urban emergence in the EDZs, the CFLUS model has the following improvements. First, instead of the conventional supervised classification method (e.g., ANN), an unsupervised classification method, the fuzzy C-means algorithm (FCM), is implemented to determine the urban development potential without requiring historical urban land. The reason why we choose the FCM algorithm instead of the other commonly used clustering method (e.g., K-mean algorithm) will be elaborated below. We also designed a multistep clustering framework and a fine-tuned method for the FCM algorithm for the process of determining urban development potential; Second, we designed an urban emergence mechanism based on local maxima to select urban development points for EDZs that have no historical urban land. The general structure of the CFLUS model is illustrated in Figure 1.

2.1. Multistep fuzzy clustering framework

The fuzzy C-means (FCM) algorithm is used to cluster various driving factors of urban development, such as slope, proximity to city centers, proximity to planning traffic lines, etc., to obtain the urban development potential and non-urban development potential on each pixel. The theory of the FCM algorithm is elaborated below.

2.1.1. Fuzzy C-means algorithm

The fuzzy C-means algorithm is one of the best-known soft clustering algorithms (Rahimi *et al*. 2004). Unlike the most commonly used K-means method or ISODATA clustering, which directly assign specific categories for each sample, the fuzzy C-means algorithm allows intermediate values between various classifications while simultaneously assigning

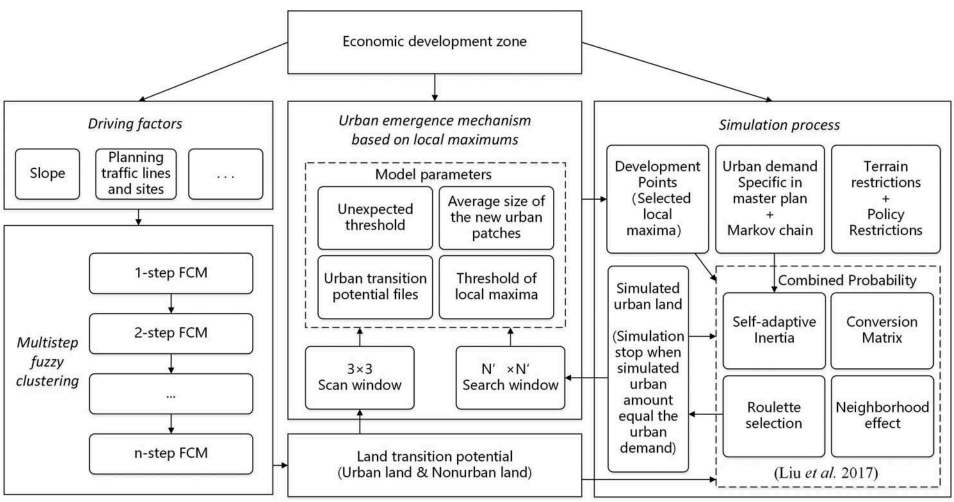

Figure 1. The framework of the proposed CFLUS model.

the memberships of different categories to the same sample, which can be regarded as the transition potential of land use types. The FCM algorithm can be expressed as the minimization of the following objective function J_m:

$$J_m = \sum_{i=1}^{N} \sum_{j=1}^{C} {\mu_{ij}}^m \left\| x_i - v_j \right\|^2 \tag{1}$$

where x_i is a vector consists of the values that are extracted from all driving factors by the same pixel i, N means the amount of the pixels in the study region; v_j denotes the clustering center of land use type j, which has the same length with x_i; C denotes the class number for clustering (Note: there are two land use types in this study: urban and non-urban respectively, thus C is equal to 2); μ_{ij} is a function that represents the degree of membership of x_i in cluster j, it means the transition potential for pixel i to change to land use type j, such that $\sum_{j=1}^{C} \mu_{ij} = 1$; and m is any real number greater than 1, which is an additional weighted exponent for the fuzzy membership. The operator $\|\cdot\|$ can be any type of inner product norm; typically, the Euclidean norm is used (Li *et al.* 2003).

The membership function μ_{ij} defines the fuzziness of an image and the information contained in the image (Selvakumar *et al.* 2012). In this study, the image is a multiband data that is composed of multiple driving factors. The fuzziness result represents the transition potential of land use type j at pixel i, which can be given by:

$$\mu_{ij} = \frac{1}{\sum_{j=1}^{C} \left(\frac{\|x_i - v_l\|}{\|x_i - v_j\|} \right)^{\frac{1}{m-1}}} \tag{2}$$

where $1 \leq l \leq$ C, $1 \leq i \leq$ N. The μ_{ij} will be iteratively updated until the clustering is finished.

The clustering centers v_l will be randomly initialized at the beginning of clustering, and also be iteratively updated during the clustering process based on the following function:

$$v_l = \frac{\sum_{i=1}^{N} {}_i x_i \cdot \mu_{il}^m}{\sum_{i=1}^{N} \mu_{il}^m} \tag{3}$$

The function J_m and Max_{ij} are iteratively minimized during the clustering process. Max_{ij} can be regarded as an indicator for determining if the iteration stops. The clustering process is finished centers until:

$$Max_{ij} = \left\{ \left| {\mu_{ij}}^{(k+1)} - {\mu_{ij}}^{(k)} \right| \right\} < \varepsilon \tag{4}$$

where ε denotes the termination value or the constant between 0 and 1, which is a parameter of FCM algorithm, and k is the number of iteration steps. The fuzzy C-means algorithm produces the membership of urban land considering various driving factors and is regarded as the transition potential of future urban areas in this study.

2.1.2. Fine-turn of the FCM algorithm

The FCM algorithm in this study only has two cluster centers because it clusters the driving factors into two groups (urban land and nonurban land). The traditional FCM initializes all the cluster centers randomly and updates cluster centers in the clustering procedures according to formula (3). However, in this study the clustering center corresponding to the urban area is set to an unchanged vector constructed by several zero values as in the following format:

$$v_0 = [0,\ 0,\ 0,\ \ldots\ ,\ 0] \tag{5}$$

The length of the vector equals the number of driving factors. The modification of the FCM algorithm corresponds to an integration of the driving factors in the simulation. Most of the driving factors we considered are proximity measurements such as the distance to the traffic lines (e.g., arterial traffic planning) and the distance to important places (e.g., downtown planning). In addition, there is an underlying assumption in most urban simulation studies that the closer the developing areas are located to such an infrastructure, the more likely that urban land is to grow (Clarke and Gaydos 1998). Other influencing factors, such as the slope, are also consistent with this assumption. Therefore, cluster centers that represent the urban areas can be defined as a vector in which all values are low, and 0 is the lowest value in this study because all the driving factors are normalized to a range of 0–1.0.

2.1.3. Multistep applications of the FCM algorithm

In the proposed method, the FCM algorithm may be implemented in several steps. In each step, we cluster the multiple driving factors into two groups (two clusters). The FCM algorithm then exports two membership images of two clusters that represent the transition potential of the future urban area and nonurban land, the FCM algorithm can ensure that the sum of these two transition potentials are 1. In the first step, we cluster in the whole study region. However, the urban development potential in the first step may not show the local spatial features and heterogeneity because the clustering region may be too large, which may not satisfy the conditions for practical use. Therefore, a binary image can be generated by comparing the two potentials on each pixel. If the urban development potential is higher than the nonurban development potential (urban development potential>0.5), the pixel of the binary image will be given a value of 1; otherwise (urban development potential≤0.5), the pixel of the binary image will be given a value of 0. An area on the binary image with a value of 1 can thus be regarded as a first-step mask region.

In the second step, FCM is applied to the data of various driving factors of urban development that is extracted by the first-step mask. Consequently, we obtain the second-step urban and nonurban development potentials and corresponding second-step mask. Then, we can apply the FCM based on the second-step mask for extracting urban and nonurban transition potentials of the next step, and apply again on the third-step mask until all the urban transition potential are less than 0.5(so that the next step of the mask will not be generated). During the simulation step, when the simulated urban area is larger than one-third of the mask area, the current urban transition potential file will be changed to the first step potential file. Figure 2 depicts a flowchart of the multistep fuzzy clustering method.

2.2. Urban emergence mechanism based on local maxima

Based on the transition potential map exported by the fine-turn FCM algorithm, an urban emergence mechanism based on local maxima of urban development potential is proposed to identify the potential development points in local areas (3 × 3 neighborhood) of the study area. In a local region (3 × 3 windows) of grid space, when the value of the core

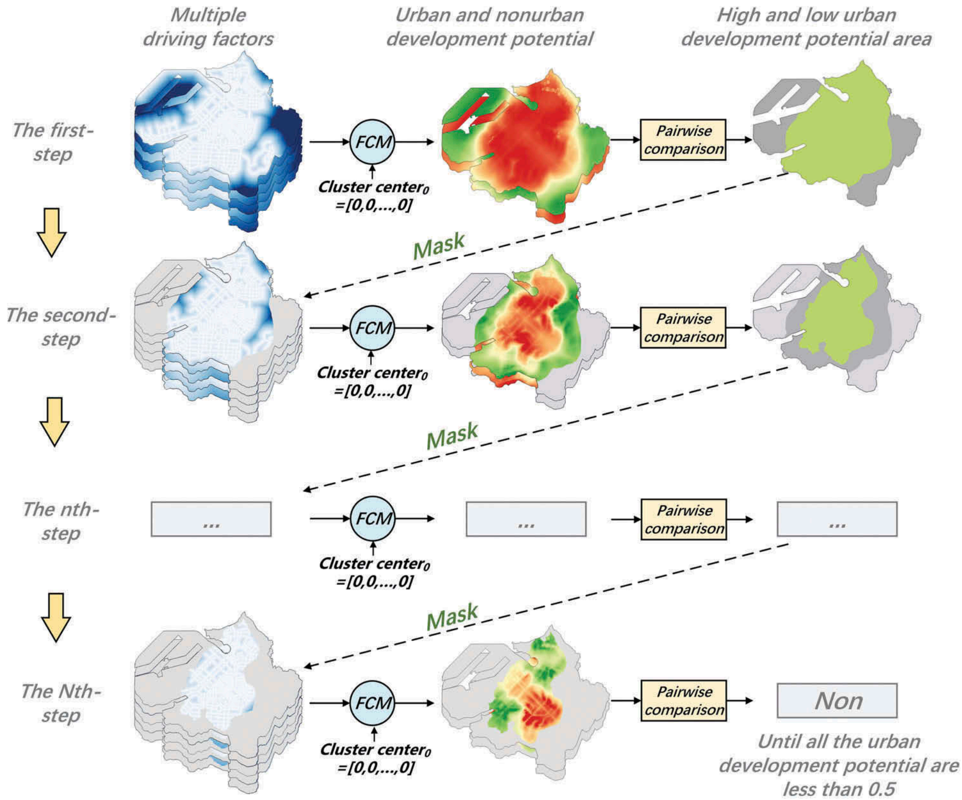

Figure 2. Schematic framework of the multistep applications of the FCM algorithm.

pixel is larger than surrounding eight other pixels, the position of the core pixel is defined as local maxima. The local maxima on the probability surface of urban development are the most likely development points at the local scale. In this mechanism, a 3×3 slide window is employed to select some local maxima of the urban development potential surfaces before simulation. The reason we used a 3×3 slide window is that the 3×3 slide window is the smallest window to find local maxima. Only by using a 3×3 slide window can we find all the local maxima on the urban development potential surface. The local maxima were selected using the following rules:

$$\text{Localmax} = \begin{cases} \text{Pswin}^{n=3}_{(i=0,j=0)} > \max\left(\text{Pswin}^{n=3}_{(i\neq0,j\neq0)}\right) & -1 \leq i,j \leq 1 \\ \text{Pswin}^{n=3}_{(i=0,j=0)} > \epsilon \text{ or } randn \times \left(1 - Rswin^{N'}_{urban}\right) < \sigma \end{cases} \tag{6}$$

where Localmax denotes the local maxima of a sliding window. If the central pixel of the slide window ($\text{Pswin}^{n=3}_{(i=0,j=0)}$) has a higher transition potential than other pixels ($\text{Pswin}^{n=3}_{(i\neq0,j\neq0)}$) in the 3×3 slide window, the central pixel is one of the local maxima of the transition potential surface. ϵ denotes a threshold within [0,1] defined by the modeler to select the local maxima with the highest urban development potential. *randn*is a random number between 0 and 1, and $Rswin^{N'}_{urban}$ is the ratio of urban cells in an N′×N′ search window, where N′ is an odd number. The search window is different from the previous 3×3 scan window in two aspects:

first, the 3 × 3 scan window searches on the urban development potential surface while the N′×N′ search window scans on the intermediate result of simulated urban land. Second, the 3 × 3 scan window is employed to select some local maxima with relative high urban development potential, and he N′×N′ search window is used to give priority to the local maxima that closer to new urban land and allows the local maxima with low urban development potential can be chosen as urban development points. If $randn \times \left(1 - Rswin^{N'}_{urban}\right)$ is less than a small unexpected threshold σ, the local maxima can also be placed in the geographic space. This condition allows the local maxima close to the original urban land to develop into urban land, though the urban development potential is low (lower than ϵ). Because the growing cities still need the support of the original urban area.

Then, the random development points will be selected from these local maxima if the following conditions are fulfilled:

$$\text{DPoint} = \begin{cases} Amount_{dpoint} < \frac{Demand_{urban} - Amound_{urban}}{APsize_{urban}} \\ Type_{Localmax} \neq urban\ land \\ P_{Localmax} > \text{RandVal} \qquad 0 \leq \text{RandVal} \leq 1 \end{cases} \tag{7}$$

where $Amount_{DPoint}$ depicts the max number of development points, which are estimated before simulation. $Demand_{urban}$ is the future urban amount specified by policymakers or predicted by 'top-down' models. $Amound_{urban}$ denotes the current urban amount. $APsize_{urban}$ represents the average size of the new urban patches, which is an input parameter of this model. The number of new development points will increase until it reaches the max amount of development points. During this process, if the position of a Localmax still not develop to urban land, the third condition will be implemented – when the transition potential of local maxima $P_{Localmax}$ is greater than a random value RandVal within (0, 1), a development point is placed in the cell. This rule gives priority to the local maxima with high urban development potential but it also allows the local maxima with low urban development potential can be chosen as development points. As a result, the total probability of the CFLUS model can be expressed by the following form:

$$TP^t_{i,k} = \begin{cases} P_{i,k} \times randn \times \text{Inertia}^t_k \times con_{c \to k} & if\ a\ DPoint\ is\ placed \\ P_{i,k} \times \Omega^t_{i,k,N} \times \text{Inertia}^t_k \times con_{c \to k} & others \end{cases} \tag{8}$$

where $TP^t_{i,k}$ is the total probability that grid cell i will be converted from the original land use into the target land use k at iteration time t (only nonurban land can be converted to urban land in this study); $P_{i,k}$ denotes the transition potential of land use type k in grid cell i, which was generated by the FCM algorithm; and $con_{c \to k}$ is a transition matrix that defines the conversion possibility from the original land use type c to the target land use k (1 denotes possible conversion, and 0 denotes impossible conversion). *randn* is a random value between [0,1], which allows new urban land to emerge in places where the neighborhood effect is zero. $\Omega^t_{i,k,N}$ denotes the neighborhood effect of land use type k in grid cell i at time t, N is the neighborhood size. $Inertia^t_k$ denotes the inertia coefficient of land-use type k at iteration time t, which is determined by $Demand_{urban}$and $Amound_{urban}$ (Liu *et al*. 2017).

Finally, we have both urban and nonurban total probabilities. These are adjusted in proportion in the following formula to ensure that the sum is 1:

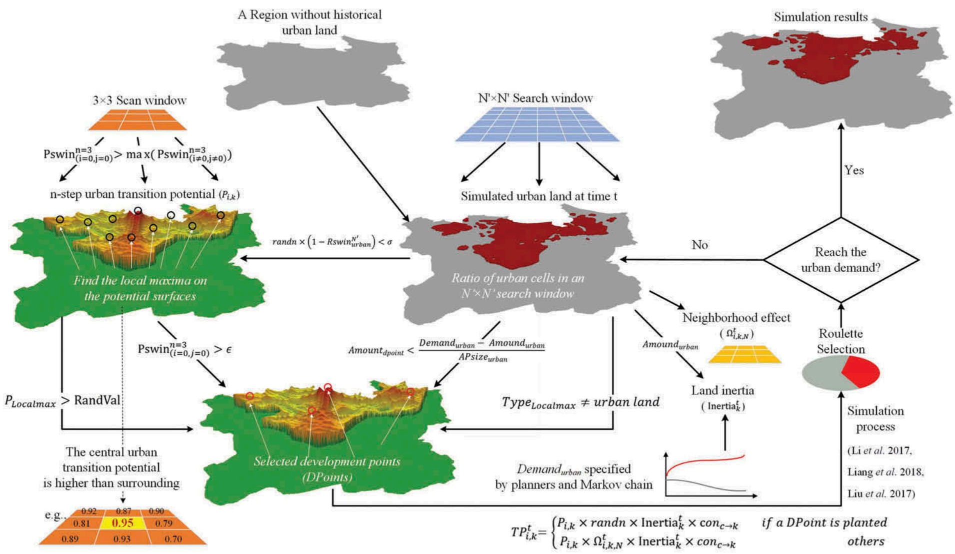

Figure 3. Framework of the urban emergence mechanism based on local maxima.

$$TP_{i,k}^{t} = \frac{TP_{i,k}^{t}}{\sum_{k=1}^{2} TP_{i,k}^{t}} \tag{9}$$

Then, the total probabilities of two land use types will construct a roulette selection mechanism, in which the two land use types will compete in each pixel. If the urban land use type wins on a nonurban pixel, the nonurban pixel will change to an urban pixel. The roulette selection mechanism enables the CFLUS model to better simulate the uncertainties and randomness in the urban growth process (Chen *et al*. 2013). Figure 3 depicts a flowchart of the urban emergence mechanism based on local maxima.

3. Study area and datasets

3.1. Economic development zones (EDZs)

EDZs, specified and managed by the urban designers, are important parts of urban growth in many fast-developing regions (Wu and Webster 1998, Yang and Wang 2008, Al-Ahmadi *et al*. 2009, Ong 2014), like China (He *et al*. 2016, Liang *et al*. 2018a, Liu *et al*. 2017, Yeh and Li 1998). The number of EDZs in China reached 4120 in 1996. By 2003, it had rocketed to 6866 and still increase with years (Huang *et al*. 2017). Zeng (2010) has pointed out that EDZs are often managed in a top-down way which created a peculiar landscape and played a critical role in China's processes of urbanization and industrialization (Huang *et al*. 2017). Therefore, it is very important for urban designers to guide the reasonable layout of the urban shapes in the EDZs.

EDZs used to be established in remote areas outside of central cities, while in recent years, some of them have integrated into the urban space as a whole, serving as new subcenters of the main urban area (Wuttke 2011). Urban designers often make plans for transportation networks, economic hubs, or other infrastructures for such regions. Considering relevant urban development policies, previous CA models can project future urban form in these regions by analyzing the relationship between historical urban distribution, and historical and planned urban infrastructure (Liang *et al.* 2018b). However, some EDZs locate in places with almost no developed infrastructure and urban areas. Thus, existing simulation models, which obtain transition rules by mining the relationship between the historical urban pattern and various driving factors with a supervised classification method (e.g., ANN and RF), are not available for these regions. This study aims to use a fuzzy unsupervised classification method to mine the urban development potential through analyzing construction conditions (e.g., terrain), existing traffic networks and planning infrastructure in the Taiping Bay development zone, a port EDZ far away from central cities with almost no original urban land.

3.2. Study area

The Taiping Bay development zone is located in northwest Dalian, the most developed city in Liaoning province and Northeast China (Figure 4). The Taiping Bay development zone is 130 km away from the downtown area of Dalian located approximately 240 km to the south of Shenyang, the provincial capital city of Liaoning province. Taiping Bay is an important port and logistics hub, which plays an important role in promoting the regional development of the whole Liaodong Peninsula.

The study area considered in this research has a total planning area of 340 km^2, including a land area of 240 km^2 and a planning reclamation area of 100 km^2. The terrain of Taiping Bay is flat. The harbor of Taiping Bay is richly endowed by nature with a broad estuary, deep water, slow-flow seawater, and a soft seabed. The traffic lines are very convenient in the Taiping Bay development zone. There are many important traffic systems and traffic sites around the study region, including the Shenyang–Haikou Expressway, the Harbin–Dalian high-speed railway and its corresponding high-speed railway stations, and high speedway gates. In summary, the Taiping Bay development zone has many location-related advantages for

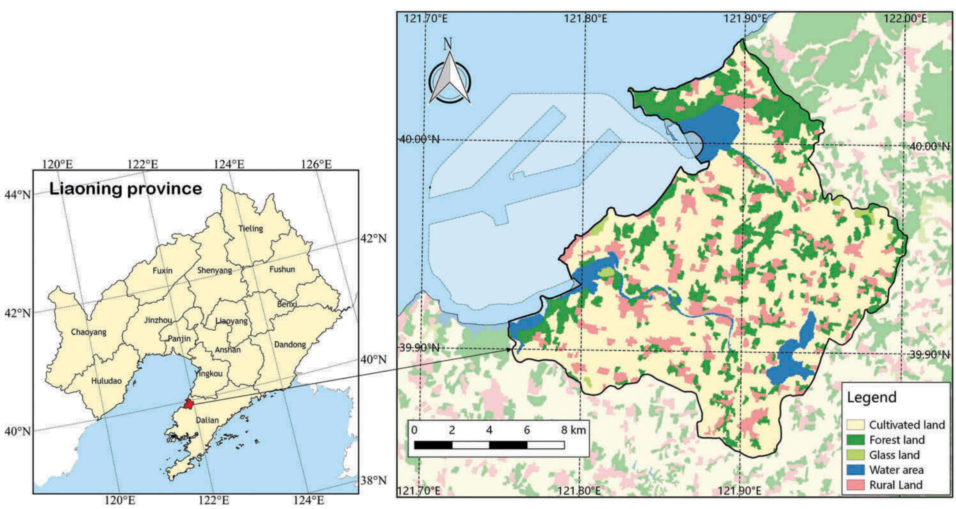

Figure 4. Spatial location of the Taiping Bay development zone.

developing as a port city, including favorable traffic infrastructure, excellent natural conditions and ample space for development.

The development of the Taiping Bay development zone has led to 29.35 billion dollars in investments. The decision-makers aim to build a modern port city in which humanity and nature coexist harmoniously. Therefore, accurately identifying the most suitable area for developing new cities, properly designing the urban form, and optimizing the layout of urban space and land-use structure is of crucial importance for the Taiping Bay development zone.

3.3. Data processing

Most of the driving factors are provided by the Master plan (2013–2020) in the Taiping Bay development zone. However, the construction of the Taiping Bay EDZ has been delayed by attracting investment and environmental protection issues in 2013 and 2017. The construction of Taiping Bay EDZ is imperative because it has finished its attraction of investment in 2017. With this in the background, the master plan from 2013 to 2020 has been extended to 2030. (https://baijiahao.baidu.com/s?id=1616369792827214639&wfr=spider&for=pc). As a result, only a few new urban patches are built during this period. The new urban patches are extracted from the Google high-resolution imagery in 2013 and 2019 using visual interpretation. The true urban patches (Figure 8) are too few to reflect the distribution pattern of urban emergence in Taiping Bay EDZ, nor can they be regarded as the training samples of most of the supervised classification methods. Therefore, we used these urban patches to validate the simulation results of the CFLUS model.

Two categories of driving factors are considered in the clustering process: 1) planning factors and 2) existing factors. Planning factors include various kinds of planning roads, city and business centers; existing factors contains all level of existing roads, terrain and the proximity to rural residential areas. A total of 18 driving factors are collected in this study (Figure 5). However, considering that the multicollinearity between variables may bring negative effects on the clustering results (Dormann *et al*. 2013), we conduct a co-linearity test with the Variance Inflation Factor (VIF)(Figure 6). The result shows that the VIF values of 8 driving factors are larger than 10, which indicates that these factors have high intercorrelations or inter-association with other independent variables. These driving factors are excluded in the clustering process.

Finally, there are 10 factors pass the VIF test, include 6 planning factors and 4 existing factors. These data used in this study are listed in Table 1. After coordinate correction and rasterization, all of the spatial datasets were calculated or resampled to the same resolution of 10×10 m. We only considered two types of land use (urban land and nonurban land) in the simulation. In addition, the water area, natural hilly region, and areas with relatively high slope (>5°) are regarded as restricted areas that are not allowed to develop into urban land.

4. Model implementation and results

The applicability of the proposed CFLUS model is first tested by discussing the urban development potential determined by the multistep fuzzy clustering. Then, the spatial simulation process of the CFLUS model is implemented to generate the simulation results

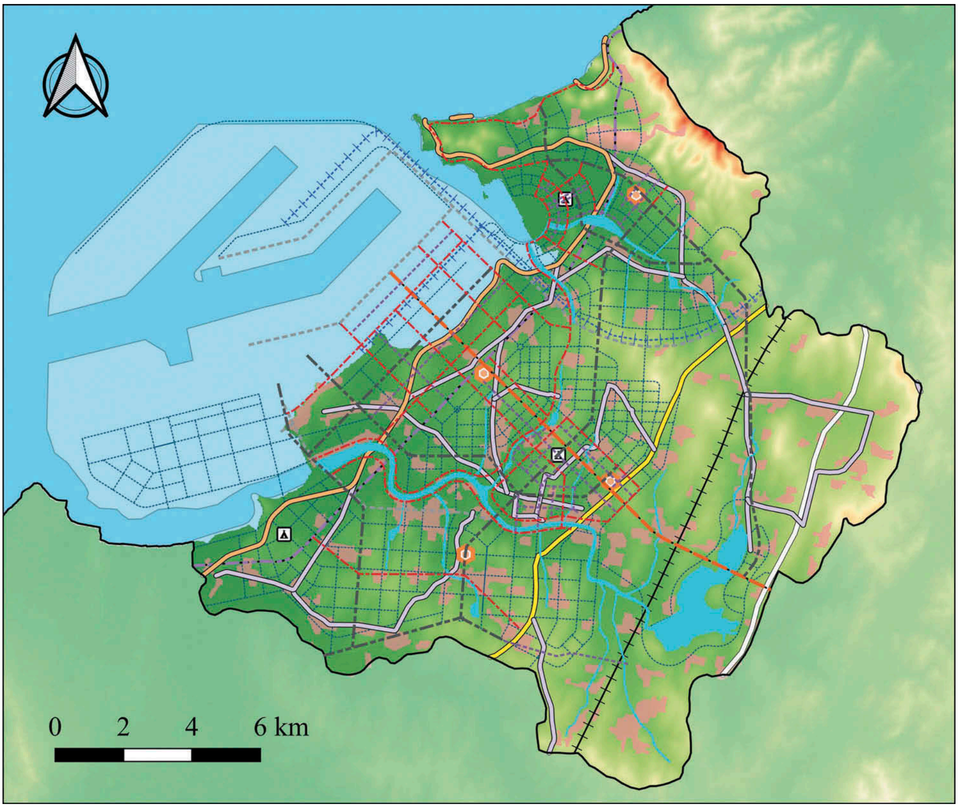

Figure 5. Driving factors collected by this study.

regarding the urban amount decided by the master plan. In this study, the FCM algorithm is applied twice to generate the two-step urban development potential, which divides the driving factors listed in Table 1 into two categories: potential urban land and nonurban land. The additional fuzzy exponent (m) is set to 2.0 according to the study proposed by Rahimi *et al*. (2004), for better clustering results can be obtained based on the value of 2.0. The termination threshold (ε) of the FCM algorithm is set at 5×10^{-10}. The termination threshold is encouraged to be set to lower values. The CFLUS will stop and output the clustering result when the change of all the values of the fuzziness result (development potential) is less than this threshold. The distribution pattern of fuzziness results will not change a lot with the variation of the termination threshold. The Euclidean distance is used to measure the distance between data points and cluster centers.

4.1. Urban development potential analysis

Figure 7(a,b) depict the output urban development potentials of the first-step fuzzy clustering and second-step fuzzy clustering, respectively. The first-step clustering roughly identified the high development potential area (the mask region) for the second-step clustering. The second-step urban development potential on each cell is less than 0.5, so we used the second-step clustering result to run the simulation. Based on the mask region, the second-step

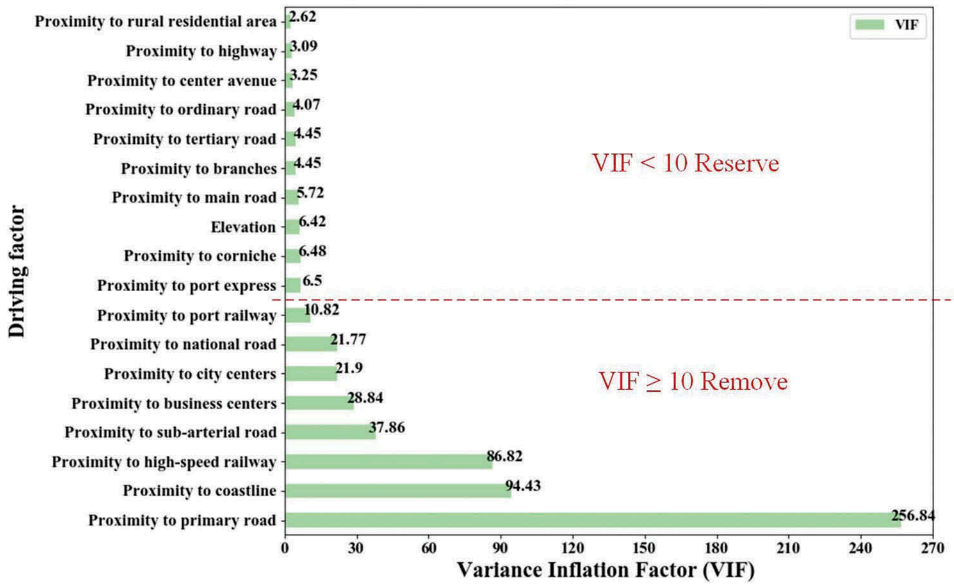

Figure 6. Variance Inflation Factor (VIF) of the driving factors. A VIF value larger than 10 indicates that the corresponding driving factors have high intercorrelations with other variables.

Table 1. Independent variables that pass the VIF test in this study.

Category	Data	Data resource
Planning factors	Proximity to port express	Master Plan
	Proximity to corniche	
	Proximity to main road	
	Proximity to branches	
	Proximity to ordinary road	
	Proximity to center avenue	
Existing factors	Proximity to rural residential area	OpenStreetMap
	Proximity to tertiary road	
	Proximity to highway	
	Elevation	http://www.gscloud.cn/
Restricted data	Water and terrain restrictions	Master plan & Slope>5°

clustering output a more precise urban development potential with more spatial features of road networks and spatial heterogeneity. The higher urban development potential region in the mask can also be recognized in the second-step clustering.

4.2. Urban growth simulation

4.2.1. Model validation and sensitive analysis

We first validated the CFLUS model with real data (new urban patches from 2013 to 2019). Based on the second-step clustering result, we simulated the emergence of new urban patches from 2013 to 2019. We have also developed a sensitivity analysis for testing the model sensitivity to the average size of the new urban patches (ASNUP), the most important parameter for determining the pattern of urban emergence. Other parameters are listed in Table 2. All the parameters are determined by the user's experience and

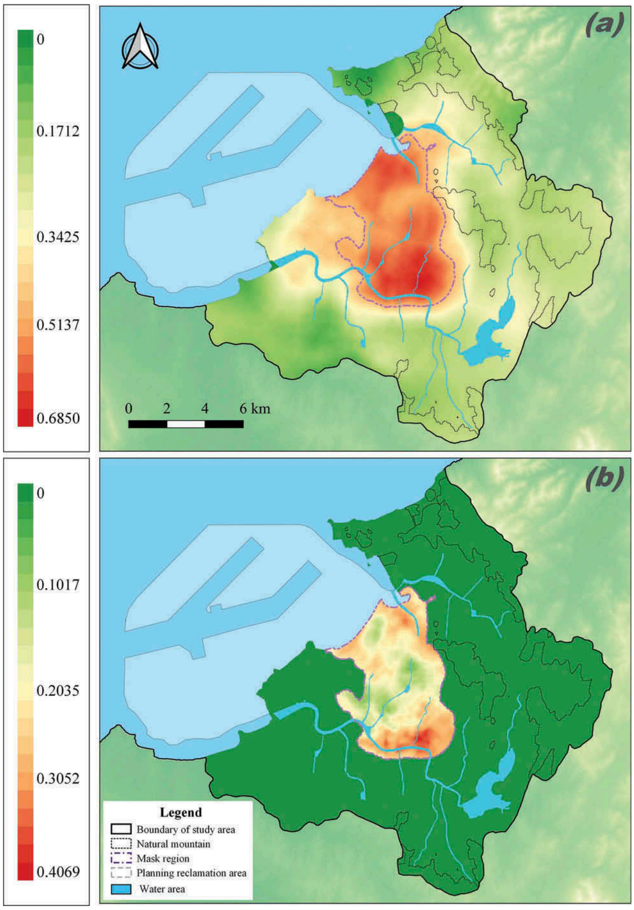

Figure 7. Urban development potential in the Taiping Bay development zone: (a) the first-step urban development potential; (b) the second-step urban development potential.

calibrated using the trial-and-error approach (Feng and Tong 2020). Thus the values of these parameters are determined by the simulation data, they may change if the CFLUS is applied to other regions.

We overlaid the simulated urban emergence with the actual new urban patches and find that most of the simulated urban patches appear nearby the actual urban patches. However, considering that the urban emergence from 2013 to 2019 are very few, it is very difficult for simulation models to hit the actual urban emergence at cell scale. Thus we aggregate the actual and simulated urban emergence (10 m resolution) to the hexagon grid with the side length of 2000 m and compare the spatial patterns of actual and simulated results with the correlation coefficient (R) and the ratio of the number of correct hexagons that contain both actual and simulated new urban patches (C).

Figure 8 shows the R and C values of the simulation results under different ASNUP parameters. The correlation coefficient between actual and simulated patterns are still relatively low (ranging from 0.1792 to 0.2581), the CFLUS model reached the highest R-value when the ASNUP is 150. However, the CFLUS model performed much better at

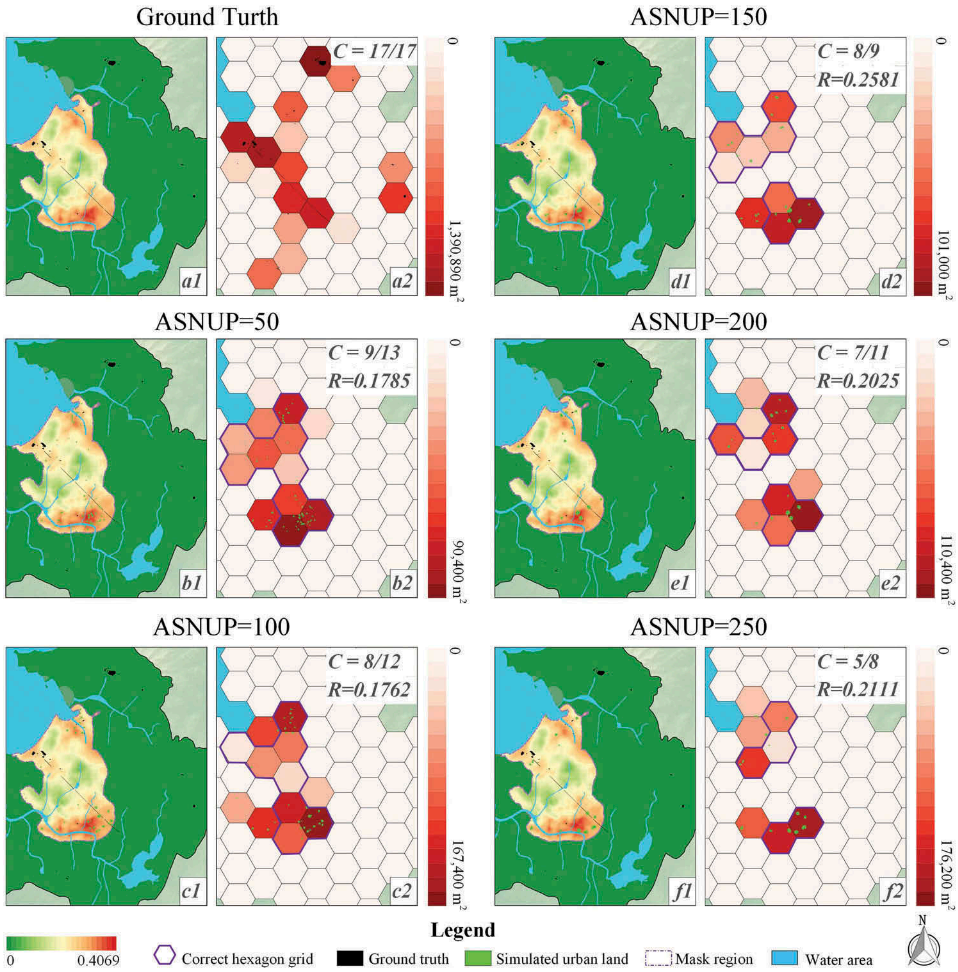

Figure 8. Validation and sensitivity analysis of the CFLUS model. Panel a1, a2 are the ground truth, Panel d1-f1 show the second-step development potential map, and the comparison maps between actual and simulated urban emergence under different ASNUP values. Panel b2-f2 show the simulated urban emergence and their aggregate hexagons. R is the correlation coefficient between ground truth and simulation results at the hexagon grid scale. C shows the ratio between the number of correct hexagons and the simulated hexagons.

predicting the potential areas of urban emergence. For example, when the ASNUP value is 50, the CFLUS model successfully identified 9 correct hexagons (17 correct hexagons in total) in which new urban patches emerged from 2013 to 2019. When the ASNUP is 150, the correct number of hexagons has dropped to 8 but only one incorrect hexagon appeared in the simulation. When the ASNUP is set to 250, the CFLUS model predicted the least amount of hexagon areas (8) and the correct hexagon (5).

The sensitivity analysis shows that the aggregate pattern of the simulation results is sensitive to the change of the ASNUP. A low value of ASNUP results in a more dispersed distribution pattern of new urban patches. Thus, the CFLUS model tends to predict more hexagon areas and the correlation coefficients at the hexagon scale

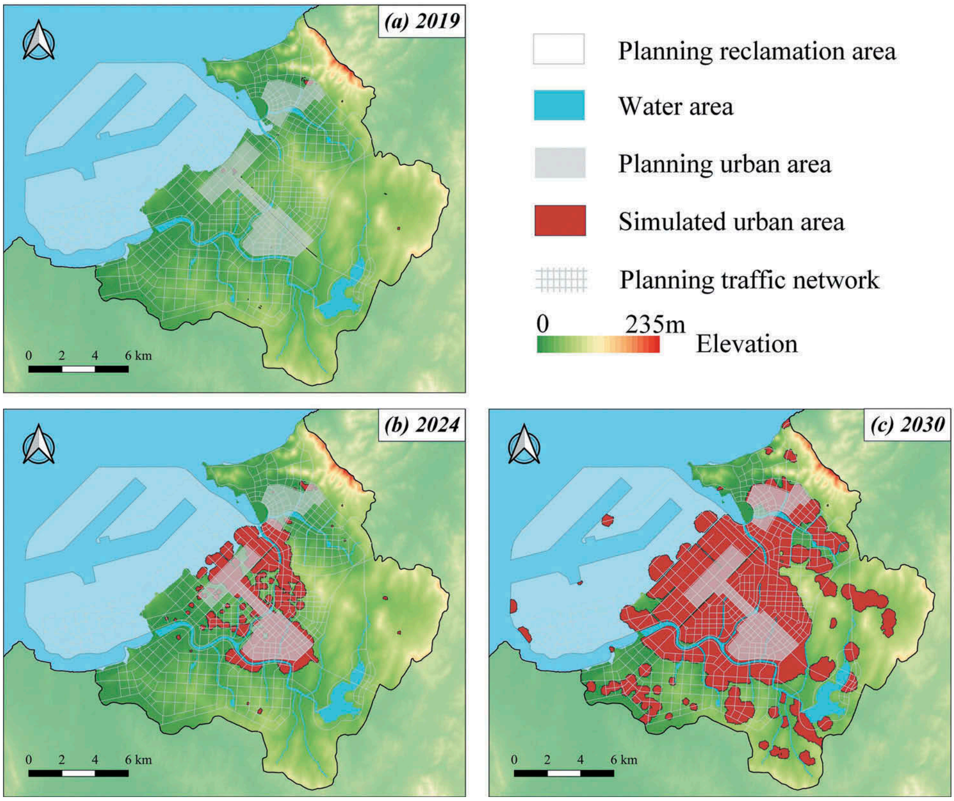

Figure 9. Simulated urban pattern in the Taiping Bay development zone in 2024 and 2030.

Table 2. Simulating parameters of the CFLUS model.

Simulation periods	2013-2019	2019-2024 (early stage)	2024-2030 (long term)	Symbols
Urban amount (square km)	0.3805	33.40	111.30	$Demand_{urban}$
Slide window (pixel)	3 (fixed)	3 (fixed)	3 (fixed)	-
Neighborhood (pixel)	3	3	3	N
Size of search window (pixel)	10	20/10	10	N'
Average size of the new urban patches (pixel)	50/100/150/200/250	800/2000	2000	$APSize_{urban}$ (ASNUP)
Threshold of local maxima	0.01	0.10	0.10	ϵ
Unexpected threshold	0.01	0.1/0.05	0.05	σ
Urban development potential file	2nd step	2nd/1st step	1st step	-

change accordingly. But because of the randomness in the simulation mechanism of the CFLUS model, we didn't find any obvious change law between R and the growth of ASNUP. This validation process also indicates that the CFLUS model can effectively detect the potential regions for new urban patches. Although the simulation accuracies are not high in the validation process, there is still room to improve the performance of the CFLUS model if the future urban demand becomes larger. Therefore, we continued to simulate the urban emergence and expansion from 2019 to 2030 with the CFLUS model.

4.2.2. Scenario development

Based on the transition potential given by the second-step fuzzy C-means algorithm, the CFLUS model is employed to simulate future urban development in the Taiping Bay development zone. We determine the land demand in 2019–2030 based on the expected population growth. According to the master plan in Taiping Bay EDZ, the population capacity of the early stage is 300 thousand and the construction area is 33.4 square kilometers. The population capacity in the long term is 1 million, and according to the ratio between construction area and population in the early stage, the construction area in the long term is about 111.3 square kilometers. We defined the early stage as 2019–2024 because the 'five-year plan' is commonly used in the infrastructure development of a region (Planning Commission 2008), thus the long-term stage can be defined as 2024–2030. Water areas are regarded as conversion constraints for protecting the future environment, and converting these into an urban area is prohibited. The simulation parameters and urban demands in corresponding years are listed in Table 2.

We used the second-step urban development potential (Figure 7(b)) during the early stage periods (2019–2024) to simulate the emergence and development of the start zones. When the simulated urban area is larger than one-third of the mask area, the current urban development potential file is changed to the first step potential file. Note that in the period 2019–2024, when the simulated urban area is larger than one-third of the first step mask area, the urban development potential file is changed to the 1st step potential file, and other parameters are set to the same as the parameters in the simulation period in 2024–2030 (Table 2).

4.2.3. Urban growth simulation and comparison with master plan

Under the driven of future urban demand, we used a land-use map in 2019 as the start map for the simulation, together with the simulation parameters provided in Table 1, the future urban development in the Taiping Bay development zone was projected using the proposed CFLUS model, and compared the simulation results with planned urban land provided by the master plan in Taiping Bay EDZ. Note that despite our simulation is based on both planning traffic networks and current driving factors, the planned urban form provided in the master plan is not involved in the modeling process of the CFLUS model. Figure 9(b) shows the urban planning area delineated in the master planning and the simulation result in the early stage (2019–2024). The new urban areas primarily appear at the center of the Taiping Bay development zone, emerge along and nearby the central avenue, which has a high consistency to the master planning. The northeast side of the central avenue develops more urban areas than the urban area on the south side, this trend also shows the same tendency as the master planning. However, the northern urban block specified by the master planning is farther away from the central avenue than the simulated urban patches at this stage, which indicates that the currently existing and planning infrastructure still can't fully support the urban pattern specified by the master planning at this stage. The simulation result in 2030 shows that the urban area will cover the central part of the study region, and a series of urban patches will appear around the central urban block Figure 9(c). The simulated urban shape in 2030 has displayed the possible future urban layout in the Taiping Bay development zone.

5. Discussion

In this study, a CA-based model with an urban emergence mechanism and the FCM algorithm were incorporated as a CFLUS model to enhance the simulation of the urban emergence and expansion in EDZs. The combination of the CA model and the unsupervised fuzzy clustering method enables the CFLUS model to identify development hotspots without analyzing the relationship between historical urban land use and various driving factors. Although because the validation data is too few, the simulation accuracies of the CFLUS model are not high in the validation process. But the CFLUS model still showed its ability to efficiently detect the potential regions for new urban patches. So there is still plenty of room to improve the performance of the CFLUS model if the future urban demand becomes larger.

Due to the ability of the CFLUS model to discover the potential development areas and predict probable urban patterns for regions that do not have any original urban areas. It can help urban designers decide where to build the start zones for the EDZs by providing specific locations. Furthermore, the rudiments of the future urban shape can be generated in the simulation process, which can provide key information for planners on designing the future urban form in EDZs.

Although economic development zones can be regarded as planned urban seeds in the coming years, and this study is to evaluate the urban development potential and spatially allocate the potential urban development with a seed-based CA model. But it doesn't mean that there is no need to simulate urban growth in the places that can be regarded as planned urban seeds in the coming years. Because even though we know these regions will have planned urban seeds in the future, but we still don't know where to put these planned seeds. Although EDZs are usually small scale area and planned to be built, questions such as 'where to put the future urban seeds within EDZs?'; 'how to manage the layout of urban shape in EDZs?'; 'which parts of the EDZs are most likely or best suited to build new urban areas?' still exist. The CFLUS model can be used to answer these questions.

As an adapted version of a CA-based FLUS model, the CFLUS model is developed to simulate the urban emergence and expansion in EDZs that usually have a smaller area and scale than the cities they belong to. It seems that the CFLUS model's function and scope of application are for a specific use and not as widely used as its predecessor FLUS model which has been used in several simulations range from the global scale to city scale (Li *et al*. 2017). However, it still has a broad application because China has been experiencing a 'development zone fever' since 1990 (Yang and Wang 2008); thus, many development zones or industrial parks will be established with the fast development of China. What's more, this method may contribute to the urban design and urban growth simulation of the Xiong'an New Area (https://en.wikipedia.org/wiki/Xiong%27an), a new area of national significance, following the Shenzhen Special Economic Zone and Shanghai Pudong New Area in China.

6. Conclusion

EDZs are often used by planners to guide the development of new urban areas in fast developing countries. Simulating urban development under the influence of different planning policies in these regions can help planners examine the consequences and outcomes of varying policies. However, when there is no historical or original urban land in the EDZs, simulation models are not available. To solve this problem, we developed a CFLUS model to simulate urban growth in a place almost without urban or developed areas.

The proposed CFLUS model was validated in the simulation period from 2013 to 2019. The results showed that it can identify the potential regions for urban emergence. The CFLUS model was applied to simulate the urban growth in the Taiping Bay development zone during 2019–2030. The results show that the CFLUS model can predict urban emergence and generate the layout of the future urban shape. Urban designers can also obtain the probable long-term urban development trajectory in the Taiping Bay development zone through the simulation of the CFLUS model. Understandably, the proposed model can prevent the creation and adoption of poor urban designs during the planning of development zones and can provide an analytical basis and scientific support for exploring future development alternatives.

Acknowledgments

We sincerely thank the editors and the three anonymous reviewers for their useful comments and suggestions that significantly strengthened this manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was funded by the National Key R & D Program of China (2017YFA0604404); the National Natural Science Foundation of China (Grant No. 41671398 and Grant No. 41901332); and the China Postdoctoral Science Foundation funded project (Grant No. 2019M652729).

Notes on contributors

Xun Liang is currently a post-doctor researcher in the School of Geography and Information Engineering, China University of Geosciences. His research focuses on land use simulation and Spatio-temporal modeling.

Xiaoping Liu is a Professor in the School of Geography and Planning, Sun Yat-sen University. His research interests include land-use planning, urban expansion, big data, and land-use simulation.

Guangliang Chen is currently an engineer at Guangzhou lantu Geographic Information Technology Co., Ltd. His research focuses on land use simulation and urban computing.

Jiye Leng is a Master student in the Department of Geography and Planning, University of Toronto. His research interests include remote sensing in global carbon cycle modeling.

Youyue Wen is currently an engineer in the South China Institute of Environmental Sciences. MEE. His research focuses on terrestrial ecosystem modeling, environmental remote sensing application, environmental pollution improvement.

Guangzhao Chen is a PhD Candidate in the School of Geography and Planning, Sun Yat-sen University. His research interests include land-use simulation, urban expansion, and environmental analysis.

ORCID

Xun Liang http://orcid.org/0000-0001-9401-7353
Xiaoping Liu http://orcid.org/0000-0003-4242-5392
Jiye Leng http://orcid.org/0000-0002-2059-1341
Youyue Wen http://orcid.org/0000-0001-7506-3549
Guangzhao Chen http://orcid.org/0000-0001-7537-2288

Data and codes availability statement

The data and codes that support the findings of this study are available at figshare.com with the identifier [DOI: 10.6084/m9.figshare.11676222].

References

Al-Ahmadi, K., *et al*. 2009. A Fuzzy Cellular Automata Urban Growth Model (FCAUGM) for the City of Riyadh, Saudi Arabia. Part 2: scenario testing. *Applied Spatial Analysis and Policy*, 2 (2), 85–105. doi:10.1007/s12061-008-9019-z.

Amour, C.B.D., *et al*. 2017. Future urban land expansion and implications for global croplands. *Proceedings of the National Academy of Sciences of the United States of America*, 114 (34), 8939–8944. doi:10.1073/pnas.1606036114.

Arsanjani, J.J., Fibæk, C.S., and Vaz, E., 2018. Development of a cellular automata model using open source technologies for monitoring urbanisation in the global south: the case of Maputo, Mozambique. *Habitat International*, 71, 38–48. doi:10.1016/j.habitatint.2017.11.003

Arsanjani, J.J., Kainz, W., and Mousivand, A.J., 2011. Tracking dynamic land-use change using spatially explicit Markov Chain based on cellular automata: the case of Tehran. *International Journal of Image & Data Fusion*, 2 (4), 329–345. doi:10.1080/19479832.2011.605397.

Barredo, J.I., *et al*. 2003. Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Landscape and Urban Planning*, 64 (3), 145–160. doi:10.1016/S0169-2046(02)00218-9.

Chen, Y., *et al*. 2013. Simulating urban form and energy consumption in the Pearl River delta under different development strategies. *Annals of the American Association of Geographers*, 103 (6), 1567–1585. doi:10.1080/00045608.2012.740360.

Chen, Y., *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. *International Journal of Geographical Information Science*, 28 (2), 234–255. doi:10.1080/13658816.2013.831868.

Chen, Y., *et al*., 2016. Capturing the varying effects of driving forces over time for the simulation of urban growth by using survival analysis and cellular automata. *Landscape & Urban Planning*, 152, 59–71. doi:10.1016/j.landurbplan.2016.03.011.

Clarke, K.C., 2014. Why simulate cities? *GeoJournal*, 79 (2), 129–136. doi:10.1007/s10708-013-9499-5.

Clarke, K.C. and Gaydos, L.J., 1998. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographical Information Science*, 12 (7), 699–714. doi:10.1080/136588198241617.

Deal, B. and Schunk, D., 2004. Spatial dynamic modeling and urban land use transformation: a simulation approach to assessing the costs of urban sprawl. *Ecological Economics*, 51 (1–2), 79–95. doi:10.1016/j.ecolecon.2004.04.008.

Dong, N., *et al*., 2018. Land use projections in China under global socioeconomic and emission scenarios: utilizing a scenario-based land-use change assessment framework. *Global Environmental Change*, 50, 164–177. doi:10.1016/j.gloenvcha.2018.04.001.

Dormann, C.F., *et al*. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *ECOGRAPHY*, 36 (1), 27–46. doi:10.1111/j.1600-0587.2012.07348.x.

Feng, Y., *et al*., 2016. Simulation of dynamic urban growth with partial least squares regression-based cellular automata in a GIS environment. *International Journal of Geo-Information*, 5 (12), 243.

Feng, Y. and Tong, X., 2020. A new cellular automata framework of urban growth modeling by incorporating statistical and heuristic methods. *International Journal of Geographical Information Science*, 34 (1), 74–97. doi:10.1080/13658816.2019.1648813.

Fragkias, M. and Seto, K.C., 2009. Evolving rank-size distributions of intra-metropolitan urban clusters in South China. *Computers, Environment and Urban Systems*, 33 (3), 189–199. doi:10.1016/j.compenvurbsys.2008.08.005.

Guy, E., Roger, W., and Inge, U., 1997. *Integrating constrained cellular automata models, GIS and decision support tools for urban planning and policy making*. *In:* Timmermans H.P.J, ed. *28 Decision support systems in urban planning*. London: E&FN Spon, 125–155.

He, C., *et al*. 2006. Modeling urban expansion scenarios by coupling cellular automata model and system dynamic model in Beijing, China. *Applied Geography*, 26 (3–4), 323–345. doi:10.1016/j.apgeog.2006.09.006.

He, J., *et al*., 2018. Mining transition rules of cellular automata for simulating urban expansion by using the deep learning techniques. *International Journal of Geographical Information Science*, 32 (10), 2076–2097.

He, J., Guan, X., and Yu, Y., 2016. A modeling approach for farmland protection zoning considering spatial heterogeneity: a case study of E-Zhou City, China. *Sustainability*, 8 (10), 1052. doi:10.3390/su8101052.

He, J., Huang, J., and Li, C., 2017. The evaluation for the impact of land use change on habitat quality: a joint contribution of cellular automata scenario simulation and habitat quality assessment model. *Ecological Modelling*, 366, 58–67. doi:10.1016/j.ecolmodel.2017.10.001

Huang, J., *et al*., 2018a. An ex-post evaluation approach to assess the impacts of accomplished urban structure shift on landscape connectivity. *Science of the Total Environment*, 622-623, 1143–1152. doi:10.1016/j.scitotenv.2017.12.094.

Huang, Q., *et al*. 2014. Modeling the impacts of drying trend scenarios on land systems in northern China using an integrated SD and CA model. *Science China Earth Sciences*, 57 (4), 839–854. doi:10.1007/s11430-013-4799-7.

Huang, Y., *et al*., 2018b. Simulating urban expansion and its impact on functional connectivity in the Three Gorges Reservoir Area. *Science of the Total Environment*, 643, 1553–1561. doi:10.1016/j.scitotenv.2018.06.332.

Huang, Z., He, C., and Zhu, S., 2017. Do China's economic development zones improve land use efficiency? The effects of selection, factor accumulation and agglomeration. *Landscape & Urban Planning*, 162, 145–156. doi:10.1016/j.landurbplan.2017.02.008

Inkoom, J.N., Nyarko, B.K., and Antwi, K.B., 2017. Explicit modeling of spatial growth patterns in Shama, Ghana: an agent-based approach. *Journal of Geovisualization and Spatial Analysis*, 1 (1–2), 1–2. doi:10.1007/s41651-017-0006-2.

Jjumba, A. and Dragićević, S., 2012. High resolution urban land-use change modeling: agent iCity approach. *Applied Spatial Analysis and Policy*, 5 (4), 291–315. doi:10.1007/s12061-011-9071-y.

Kamusoko, C. and Gamba, J., 2015. Simulating urban growth using a Random Forest-Cellular Automata (RF-CA) model. *ISPRS International Journal of Geo-Information*, 4 (2), 447–470. doi:10.3390/ijgi4020447.

Lagarias, A., 2012. Urban sprawl simulation linking macro-scale processes to micro-dynamics through cellular automata, an application in Thessaloniki, Greece. *Applied Geography*, 34, 146–160. doi:10.1016/j.apgeog.2011.10.018

Li, D., *et al*. 2019. Urban heat island: aerodynamics or imperviousness? *Science Advances*, 5 (4), u4299. doi:10.1126/sciadv.aau4299.

Li, X., *et al*., 2003. Inhomogeneity correction for magnetic resonance images with fuzzy C-mean algorithm. *Proceedings of SPIE*, *5032*, 995–1005.

Li, X., *et al*. 2017. A new global land-use and land-cover change product at a 1-km resolution for 2010 to 2100 based on human–environment interactions. *Annals of the American Association of Geographers*, 107 (5), 1040–1059. doi:10.1080/24694452.2017.1303357.

Li, X. and Liu, X., 2006. An extended cellular automaton using case-based reasoning for simulating urban development in a large complex region. *International Journal of Geographical Information Science*, 20 (10), 1109–1136. doi:10.1080/13658810600816870.

Li, X. and Yeh, A.G., 2002. Neural-network-based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 16 (4), 323–343. doi:10.1080/13658810210137004.

Li, X. and Yeh, G.O., 2001. Calibration of cellular automata by using neural networks for the simulation of complex urban systems. *Environment & Planning A*, 33 (8), 1445–1462. doi:10.1068/a33210.

Li, Z., Bagan, H., and Yamagata, Y., 2018. Analysis of spatiotemporal land cover changes in Inner Mongolia using self-organizing map neural network and grid cells method. *Science of the Total Environment*, 636, 1180–1191. doi:10.1016/j.scitotenv.2018.04.361

Liang, X., *et al*., 2018a. Delineating multi-scenario urban growth boundaries with a CA-based FLUS model and morphological method. *Landscape and Urban Planning*, 177, 47–63. doi:10.1016/j.landurbplan.2018.04.016.

Liang, X., *et al*. 2018b. Urban growth simulation by incorporating planning policies into a CA-based future land-use simulation model. *International Journal of Geographical Information Science*, 32 (11), 2294–2316. doi:10.1080/13658816.2018.1502441.

Liao, W., *et al*. 2017a. Estimated influence of urbanization on surface warming in Eastern China using time-varying land use data. *International Journal of Climatology*, 37 (7), 3197–3208. doi:10.1002/joc.2017.37.issue-7.

Liao, W., *et al*. 2017b. the impact of energy consumption on the Surface Urban Heat Island in China's 32 major cities. *Remote Sensing*, 9 (3), 250. doi:10.3390/rs9030250.

Liao, W., *et al*. 2018. Stronger contributions of urbanization to heat wave trends in wet climates. *Geophysical Research Letters*, 45 (20), 11, 310–311, 317. doi:10.1029/2018GL079679.

Liu, X., *et al*., 2008. A bottom-up approach to discover transition rules of cellular automata using ant intelligence. *International Journal of Geographical Information Science*, 22 (11–12), 1247–1269.

Liu, X., *et al*., 2017. A future land use simulation model (FLUS) for simulating multiple land use scenarios by coupling human and natural effects. *Landscape and Urban Planning*, 168, 94–116. doi:10.1016/j.landurbplan.2017.09.019.

Liu, X., *et al*., 2018. High-resolution multi-temporal mapping of global urban land using landsat images based on the Google Earth Engine Platform. *Remote Sensing of Environment*, 209, 227–239. doi:10.1016/j.rse.2018.02.055.

Long, Y., Gu, Y., and Han, H., 2012. Spatiotemporal heterogeneity of urban planning implementation effectiveness: evidence from five urban master plans of Beijing. *Landscape and Urban Planning*, 108 (2–4), 103–111. doi:10.1016/j.landurbplan.2012.08.005.

Omrani, H., *et al*., 2019. The land transformation model-cluster framework: applying k-means and the spark computing environment for large scale land change analytics. *Environmental Modelling & Software*, 111, 182–191. doi:10.1016/j.envsoft.2018.10.004.

Ong, L.H., 2014. State-led urbanization in China: Skyscrapers, land revenue and "Concentrated Villages". *The China Quarterly*, 217, 162–179. doi:10.1017/S0305741014000010

Planning Commission, G.O.I, 2008. *Eleventh five year plan (2007-2012) volume III agriculture, rural development, industry,services, and physical infrastructure*. Oxford: University Press.

Rahimi, S., *et al*., 2004. A parallel Fuzzy C-mean algorithm for image segmentation. *Processing NAFIPS 04, IEEE Annual meeting*, 1, 234–237.

Selvakumar, J., Lakshmi, A., and Arivoli, T., 2012. Brain tumor segmentation and its area calculation in brain MR images using K-mean clustering and Fuzzy C-mean algorithm. *In: IEEE-International conference on advances In engineering, science and management (ICAESM-2012)*, 186–190.

Shu, B., *et al*. 2017. Modeling urban expansion by using variable weights logistic cellular automata: a case study of Nanjing, China. *International Journal of Geographical Information Science*, 31 (7), 1314–1333. doi:10.1080/13658816.2017.1283505.

Sohl, T., *et al*. 2007. The FORE-SCE model: a practical approach for projecting land cover change using scenario-based modeling. *Journal of Land Use Science*, 2 (2), 103–126. doi:10.1080/17474230701218202.

Sohl, T. and Sayler, K., 2008. Using the FORE-SCE model to project land-cover change in the southeastern United States. *Ecological Modelling*, 219 (s1–2), 49–65. doi:10.1016/j.ecolmodel.2008.08.003.

Sohl, T.L., *et al*., 2012. Spatially explicit land-use and land-cover scenarios for the great plains of the United States. *Agriculture, Ecosystems & Environment*, 153, 1–15. doi:10.1016/j.agee.2012.02.019.

Tayyebi, A., Perry, P.C., and Tayyebi, A.H., 2014. Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster–vector routines with remote sensing and GIS. *International Journal of Geographical Information Science*, 28 (4), 639–659. doi:10.1080/13658816.2013.845892.

Tesfamariam, S. and Sadiq, R., 2006. Risk-based environmental decision-making using fuzzy analytic hierarchy process (F-AHP). *Stochastic Environmental Research and Risk Assessment*, 21 (1), 35–50. doi:10.1007/s00477-006-0042-9.

Verburg, P.H., *et al*. 2004. Land use change modelling: current practice and research priorities. *GeoJournal*, 61 (4), 309–324. doi:10.1007/s10708-004-4946-y.

Verburg, P.H. and Overmars, K.P., 2009. Combining top-down and bottom-up dynamics in land use modeling: exploring the future of abandoned farmlands in Europe with the Dyna-CLUE model. *Landscape Ecology*, 24 (9), 1167–1181. doi:10.1007/s10980-009-9355-7.

White, R. and Engelen, G., 2000. High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems*, 24 (5), 383–400. doi:10.1016/S0198-9715(00)00012-0.

Wu, F., 1998. SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *International Journal of Geographical Information Science*, 12 (1), 63–82. doi:10.1080/136588198242012.

Wu, F. and Webster, C.J., 1998. Simulation of land development through the integration of cellular automata and multicriteria evaluation. *Environment and Planning B: Planning and Design*, 25 (1), 103–126. doi:10.1068/b250103.

Wuttke, C., 2011. from industrial enclaves to prototypes of the Modern Chinese City: development zones in Guangdong. *Local Economy*, 26 (26), 363–372. doi:10.1177/0269094211409137.

Xue, B., *et al*. 2016. Extended land-use coding system and its application in urban brownfield redevelopment: case study of Tiexi District in Shenyang, China. *Journal of Urban Planning and Development*, 142 (3), 5015014. doi:10.1061/(ASCE)UP.1943-5444.0000309.

Yang, D.Y. and Wang, H., 2008. Dilemmas of local governance under the development zone fever in China: a case study of the Suzhou Region. *Urban Studies*, 45 (5–6), 1037–1054. doi:10.1177/0042098008089852.

Yao, Y., *et al*. 2017. Simulating urban land-use changes at a large scale by integrating dynamic land parcel subdivision and vector-based cellular automata. *International Journal of Geographical Information Science*, 31 (12), 2452–2479. doi:10.1080/13658816.2017.1360494.

Yeh, A.G. and Li, X., 1998. Sustainable land development model for rapid growth areas using GIS. *International Journal of Geographical Information Science*, 12 (2), 169–189.

Zeng, D.Z., 2010. Building engines for growth and competitiveness in China: experience with special economic zones and industrial clusters. *World Bank Publications*, 45 (9), 1292–1293.

Zheng, H.W., *et al*., 2015. Simulating land use change in urban renewal areas: a case study in Hong Kong. *Habitat International*, 46, 23–34. doi:10.1016/j.habitatint.2014.10.008.