Zhang Y, Long Y, 2015, V-BUDEM: a vector-based Beijing urban development model for simulating urban growth. Beijing City Lab. Working paper #60

V-BUDEM: a vector-based Beijing urban development model for simulating urban growth

Yongping Zhang and Ying Long

Abstract BUDEM (Beijing Urban Development Model) is a raster-based Cellular Automata (CA) model for supporting city planning and policies evaluation in Beijing. In this paper, we developed a vector-based version of BUDEM (V-BUDEM). In this model, urban space consists of irregular parcels, and a parcel's neighborhood is defined as all parcels surrounding it within a certain distance. Additionally, a framework of parcel subdivision was adopted to subdivide existing parcels. After describing the conceptual model of V-BUDEM, including the parcel subdivision framework, we tested it in Beijing's Yanqing Town for simulating urban growth from 2010 to 2020. Results show the V-BUDEM can be used to predict urban growth scenario, and prove the validity of our parcel subdivision framework. The main contributions of this study are as follows: (1) the model adopts a vector-based CA method using land parcels to represent urban space, composing the landscape a user would perceive as meaningful, and can simulate urban growth in a way more close to real world situation; (2) the model integrates a process of parcel subdivision, and the proposed parcel subdivision framework comprehensively considered the impacts of existing parcel boundaries in the existing land use map and planned parcel boundaries in the urban plan, and developed a straight-forward and automatic parcel subdivision tool, included in the framework as the fourth step, to partition existing

---

Yongping Zhang

Nijmegen School of Management, Radboud University Nijmegen, Nijmegen 6500 HC, the Netherlands

School of Planning and Geography, Cardiff University, Cardiff CF10 3WT, the United Kingdom

School of Urban Planning and Design, Peking University, Beijing 100871, China

Ying Long (*)

Beijing Institute of City Planning, Beijing 100045, China

University of Cambridge, the United Kingdom

Corresponding author e-mail: longying1980@gmail.com

large parcels; (3) compared with other urban models, V-BUDEM is developed specifically to identify policies required for implementing the planned urban form desired by planners and decision makers.

Keywords Vector-based; cellular automata; urban growth; Beijing

1 Introduction

Over last two decades, Cellular Automata (CA) has been widely applied in generating realistic urban growth scenarios for its ability to simulate the dynamic spatial process from a bottom-up perspective (Clark and Gaydos 1998; Landis 1995; Ward and Murray 1999; Guan et al. 2005; Liu 2012; Moghadam and Helbich 2013). Traditional CA consists of five components, namely *a space* represented as a regular grid composed of a collection of homogeneous cells, a set of possible *cell states*, and *the transition rules* which determine the evolution of the state of each cell based on states of cells within its *neighborhood* and some external constraints at each *time step* (Batty et al. 1999; White and Engelen 2000). For most existing CA models, the geographic space is typically represented as a regular raster grid and the neighborhood is defined as an assembly of adjacent cells. However, recent studies have demonstrated that the simulation results of such raster-based CA models are sensitive to the cell size and the neighborhood configuration (Moreno et al. 2009). For example, Jenerette and Wu (2001) used two different cell resolutions in their CA model to study urban expansion and showed that it generated significantly different land use patterns. Chen and Mynett (2003) investigated the impact of cell size and neighborhood configuration in a prey-predator CA model and observed that they affected both the resulting spatial patterns and the system stability. Jantz and Goetz (2005) examined the results of SLEUTH model, in response to different cell sizes and indicated that the cell size at which the land use data were represented could impact the quantification of land use patterns and the ability of the model to replicate spatial patterns. On the other hand, these raster-based CA models will be challenged and have difficulties in generating reliable land-use patterns at fine spatial resolutions (e.g. 5 m resolution) (Wang and Marceau 2013). When the resolution is increased, spatial entities in real world, such as blocks, census tract boundaries, and even individual parcels, can be identified. The use of a grid of regular cells creates areas of assumed homogeneous land use that may contain variability in reality, thereby cannot precisely represent real entities, which have irregular sizes and shapes (Stevens and Dragicevic 2007).

Some researchers have started to use vector-based, or irregular-based, CA models to avoid the questions mentioned above. The studies of vector-based CA are limited, but have gained much importance recent years. Shi and Pang (2000) presented a Voronoi-based CA in which the CA was extended by using the Voronoi as its spatial framework. Results showed that Voronoi-based CA could simulate local interactions among spatial objects to generate complex global patterns, and further simulate interactions among point, line and polygon objects with irregular shapes and sizes in a dynamic system. Semboloni (2000) used the Delaunay triangle network to represent irregular space. The space is divided into a number of contiguous cells that are generated by randomly distributed nodes connected by Delaunay triangulation, and the cells are represented by polygons whose vertices are located in the centroid of each triangle. O'Sullivan (2001a, 2001b) combined CA and graph theory to present a graph-CA model, where the space is represented as a planar graph composed of vertices and edges. These studies have demonstrated it's workable to develop a CA model using a vector-based representation, but the cells in these models could not represent actual spatial units in real world, composing the landscape a user would perceive as meaningful (Moreno et al. 2009).

In other researches, the irregular cell is able to represent entities in real world. Hu and Li (2004) developed an object-based CA, in which geographical entities, e.g. a land parcel, a block, a road, a school etc. are represented as points, lines, or polygons according to their real shapes and sizes. Geographic automata system (GAS) (Torrens and Benenson 2005) incorporated irregular vector objects as automata to represent spatial entities such as roads, buildings and parks. Stevens and Dragicevic (2007) developed iCity, in which an urban area is partitioned into discrete land use units based on cadastral information and represented as a collection of polygons. Hammam et al. (2007) introduced vector agents (VAs) into GAS (Torrens and Benenson 2005) to realize the change of each agent's geometry while interacting with other agents in its neighbourhood using a set of rules. Shen and Kawakami (2008) developed a geosimulation model using a vector-based CA to visualizing land use patterns in urban partitions. In the entity-based CA model presented by Moreno et al. (2008, 2009), the shape and size of each object can also change and a dynamic neighborhood was semantically implemented. Pinto and Antunes (2010) developed an irregular CA model based on census blocks to determine the land use demand by considering the evolution of population and employment densities over time. Agent iCity, developed by Jjumba and Dragicevic (2012), is an upgraded version of iCity, it can simulate land use change using irregular spatial units and capable of simulating the interactions of different stakeholders involved in the process of land use change. Wang and Marceau (2013) and Marceau et al. (2013) presented a patch-based CA model designed to simulate land use changes at a fine spatial resolution, and tested it in the eastern part of the Elbow River watershed in southern Alberta, Canada.

When simulating urban growth using a vector-based CA method, the parcels can be treated as basic spatial units. The parcel, as the basic component of built-up spaces, is the container of various urban infrastructures, such as buildings, green lands, municipal infrastructures, and public services. Since it has traffic connection with the road network, the parcel has close relationship with both of the block and road network. The change of parcel boundaries is common in the urban growth process. For example, when a parcel has a large size initially, locates in the suburban area, but would be developed in the future, it would be highly possible to be subdivided into several small parcels when the urban development happens. When simulating urban growth scenarios using the parcel as the basic spatial unit, it's unavoidable to consider the changing mechanisms or rules of parcel subdivision, which has a significant influence on simulation results.

Land subdivision is a standard practice in land surveying that aims to partition a tract of land into smaller sized lots, which means a small parcel in this chapter (Easa, 2008), or a process splitting up of a larger land tract, which means a large parcel here, into streets and smaller subspaces variously called parcels and lots (Dahal and Chow 2014). In general, land subdivision is carried out using expert knowledge and skills, rigorous field work, both spatial and non-spatial data, and through incorporating a range of zoning and development rules (Chen and Jiang, 2000; Wakchaure, 2001).

The simulation of the land subdivision is very useful (e.g. speeding up the process and saving the cost of field work) in many applied and research areas, including for urban growth prediction. Ko et al. (2006) developed the Fragmented Land Ownership Spatial Simulator (FLOSS) to generate ownership patterns. Alexandridis and Pijanowski (2007) used a parcelization algorithm to simulate how a seller agent subdivides its land for sale in an agent-based landscape model. However, these two studies are developed for raster-based landscape models, and are performed at relatively coarse resolutions (Wickramasuriya et al. 2011). Wakchaure (2001) developed an ArcView tool for supporting land subdivision at a single parcel level, but the tool is semi-automated. Vanegas et al. (2009a, 2009b) developed an automated urban layout generation module using the algorithm of recursive binary division, and their solution can create streets, lots arrangements. Wickramasuriya et al. (2011) presented a fully-automated land subdivision tool that uses vector data and is capable of generating layouts with both lot and street arrangements for land parcels of any shape. Vanegas et al. (2012) presented a method for interactive procedural generation of parcels within the urban modeling pipeline using two algorithms. Agent iCity (Jjumba and Dragicevic 2012) has a land subdivision module, which can subdivide land into city blocks, then small cadastral parcels depending on the anticipated growth within the city as determined by the planning agent, but the authors failed to report further details about the corresponding algorithm. Wickramasuriya et al. (2013) improved a method to dynamically subdivide parcels in land use change models. Demetriou et al. (2013) and Demetriou (2013) developed a module called Land Parcelling System (LandParcelS) to automate the land partitioning process by designing and optimizing land parcels in terms of their shape, size and value using a spatial genetic algorithm. Dahal and Chow (2014) developed a GIS toolset, called Parcel-divider, for automated subdivision of land parcels. The highlight of this toolset is that it can provide various subdivision styles based on different geometric attributes of parcels, and can be used either as a module to be integrated into an urban growth model or as a stand-alone application. Except the reports in published literature, CityEngine (ESRI, 2013) already provided the parameters for block subdivision to create lots, using recursive OBB (oriented bounding box) algorithm, offset algorithm, and skeleton subdivision algorithm. Although the studies of land subdivision simulation are comparatively mature, most of them have not been integrated into urban growth models. In addition, none of them considers the impacts of urban planning on the land subdivision when urban growth happens.

In the background of the prosperity of macro-economy, Beijing, the capital of China, has achieved unprecedented development of urban construction, especially after adopting the Reform and Open Policy in 1978. To cope with the challenge caused by rapid urban growth, the Beijing Urban Development Model (BUDEM) was developed (Long et al. 2009) to support urban planning and the evaluation of corresponding policies in Beijing Institute of City Planning (BICP), which is a leading and professional institution responsible for establishing various urban plans in Beijing and providing planning-relevant services for Beijing's government. BUDEM is a spatio-temporal dynamic urban model using a combination of the raster-based CA and Agent-based Modelling (ABM) approaches. In the first phase, the model was used to identify urban growth mechanisms in various historical phases since 1986, to retrieve urban policies needed to implement the desired (planned) urban form in 2020, and to predict the urban growth scenario until 2049. The model has been proved to be capable of analyzing historical urban growth mechanisms and predicting future urban growth for metropolitan areas in China, and its vision is to be a comprehensive urban model. Besides the urban growth simulation, some other studies have been conducted, for example, retrieving spatial policy parameters from an alternative plan (Long et al. 2012), calculating transportation energy consumption and environment impact (Long et al. 2013a), and comparing simulation and artwork of urban growth boundaries (Long et al. 2013b).

In this chapter, we aim to improve initial the raster-based BUDEM model into a vector-based one (called V-BUDEM here), using urban parcels as the basic spatial units. At this stage, we are focused on the part of urban growth simulation using this model. The V-BUDEM model was described in detail in Section 2. Then the model was applied to simulate urban growth scenario of Yanqing Town, a small town in Beijing, rather than testing it in the whole

Beijing metropolitan area in Section 3. Finally, conclusions and discussion are presented in Section 4.

2 The V-BUDEM model

2.1 Constrained cellular automata

Considering the complexity of urban growth, many researches have introduced constraints into the CA model, namely constrained CA, thereby rendering the simulation of urban growth closer to real world outcomes. The V-BUDEM is also a constrained CA model, of which the conceptual model is shown as equation 1.

$$V_i^{t+1} = f\left\{V_i^t, A_{status}, A_{loc}, A_{gov}, A_{nei}\right\} \quad (1)$$

Where V_i^t is the status at cell i of iteration t; f is the transition rules of the constrained CA. In the V-BUDEM model, the parcel, with various shape and size, is treated as the cell, and the cell status represents 0 for undeveloped or 1 for developed from non-urban built-up parcel to urban built-up parcel. A parcel's neighborhood is defined as all parcels surrounding it within a certain distance. Constrained conditions in the urban growth process consist of four aspects, including self-status constraints A_{status}, location constraints A_{loc}, government (or institutional) constraints A_{gov}, and neighborhood constraint A_{nei}. Locational, and institutional constraints are assumed to remain static during the future urban growth process, and they do not change across simulation iterations. Self-status (e.g. whether the parcel is agricultural land) constraints can also be treated as keeping static, because the parcel will be excluded in the simulation process if its self-status has changed in previous iteration. The neighborhood effect, however, continue to change with simulation iterations of the constrained CA.

The transition rules in constrained CA are illustrated in equation 2.

$$1.\ LandDemand = \sum_{t} stepArea^{t}$$

$$In \text{ iteration } t+1$$

$$2.\ S^{t} = x_{0} + \sum_{k=1}^{n-1} x_{k} * \text{a}_{k} + x_{n} * a_{n}^{t}$$

$$3.\ p_{g}^{t} = \frac{1}{1+e^{-s^{t}}}$$

$$4.\ p^{t} = \exp[\alpha(\frac{p_{g}^{t}}{p_{g,\max}^{t}} - 1) * RI^{t}] \tag{2}$$

$$5.\ RI^{t} = 1 + (\gamma^{t} - 0.5)/k$$

$$If\ \ p_{i}^{t} = p_{\max}^{t}\ \ and\ \ Area_{i} \le (stepArea^{t} - deveArea^{t})$$

$$then\ V_{i}^{t+1} = 1$$

$$p_{i}^{t} = p_{i}^{t} - p_{\max}^{t}$$

$$p_{\max}^{t} update$$

Where $LandDemand$ is the total area, which will be developed during the whole simulation process; $stepArea^{t}$ is the total area, which will be developed during the iteration t; s^{t} is the development suitability; p_{g}^{t} is the initial transition probability; $p_{g,\max}^{t}$ is the maximum value of p_{g}^{t} in iteration t; p^{t} is the final transition probability; n is the total number of considered constraint variables; x_{0} is the constant item; a_{k} is the static constraint variable (excluding the neighborhood constraint), x_{k} is the weight of a_{k}; a_{n}^{t} is the dynamic constraint variable (namely neighborhood constraint); x_{n} the weight of a_{n}^{t}; RI^{t} is the random item in iteration t; γ^{t} is the random value varying from 0 to 1; k is the random index used to regulate RI^{t}; α is the dispersion parameter ranging from 1 to 10, indicating the rigid level of development. The larger α indicates stricter development control and lower development probability with the same suitability. p_{i}^{t} is the final transition probability of cell i; $p_{\max}^{t}$ is the maximum value of p_{g}^{t} in iteration t, and originally equals $p_{g,\max}^{t}$; $Area_{i}$ is the area of cell i; $deveArea^{t}$ is the developed area in iteration t. In each iteration, the beginning value of $deveArea$ is 0, and the final value is no more than

$stepArea^t$.

The parameters that need to be calibrated in the constrained CA model include $LandDemand$, $stepArea^t$, x_k , and x_n . $LandDemand$ is determined by the difference between the total area of existing urban built-up parcels and that of planned urban built-up parcels. $stepArea^t$ is assumed to be constant through the entire simulation period and can be calculated simply as follow:

$$stepArea^t = LandDemand / t \quad (3)$$

To calibrate x_k and x_n to reflect the weights of constraint policies, the logistic regression and heuristic approaches can both be applied. The weight of x_k for static constraint variables can be retrieved by the logistic regression in which whether a parcel is developed from non-urban to urban is the dependent variable (0 or 1) and the static constraint variables are the independent variables. Holding x_k static, x_n can be calibrated using the MonoLoop method (for details see Long et al. 2009), where x_n is continually sampled from 0 to $x_{n,\max}$ with an interval of $x_{n,\max} / M$. $x_{n,\max}$ and M can be set based on the user's experience. The sampled x_n and the calibrated x_k are used as the input variables for the constrained CA model. The simulated urban form is compared with the observed urban form (namely the planned urban form) to obtain the Kappa index, which is significantly correlated with overall accuracy for comparing two maps. In this paper, the Kappa index is calculated to analyze the degree of similarity (goodness-of-fit) between the simulated and the planned urban forms parcel by parcel. The x_n with the maximum Kappa index is specified as the final weight for a_n . Finally, x_k and x_n as the calibrated weights are used in the constrained CA model to simulate future urban growth. Kappa index was initially introduced by Cohen (1960) and adapted for the accuracy assessment in remote sensing applications by Congalton and Mead (1983). The adopted method for calculating the vector-based Kappa is different with the kappa index used in raster-based CA models, and is shown as equation 4.

$$K = \frac{Area_t * (Area_{dd} + Area_{uu}) - Area_{d\_} * Area_{\_d} - Area_{u\_} * Area_{\_u}}{Area_t^2 - Area_{d\_} * Area_{\_d} - Area_{u\_} * Area_{\_u}} * 100 \quad (4)$$

Where K is the Kappa index, $Area_t$ is the total area of the study region; $Area_{dd}$ is the total area of the parcels, which are developed in both simulated and planned urban forms; $Area_{uu}$ is the total area of the parcels, which are undeveloped in both simulated and planned urban forms; $Area_{d\_}$ is the total area of the parcels, which are developed in the simulated urban form; $Area_{\_d}$ is the total area of the parcels, which are developed in the planned urban form; $Area_{u\_}$ is the total area of the parcels, which are undeveloped in the simulated urban form; $Area_{\_u}$ is the total area of the parcels, which are undeveloped in the planned urban form.

2.2 Constraint variables

Fourteen constraint variables (means $n = 14$), the same with those in the BUDEM model, are shown in Table 1. Some constraint variables may be different with those in other urban growth models, but they are suitable in China's or Beijing's situation. For example, Variables *isagri* and *isrural* stand for the transition from agricultural land and rural built-up land into urban built-up land, which is a big concern for the Beijing Municipal Government. The detailed explanations for the selection of constraint variables can be found in Long et al. (2009).

Table 1. Constraint variables in V-BUDEM

Type	Name	Value	Description
A_{status}	*isagri*	0, 1	Whether the cell is agricultural land
	isrural	0, 1	Whether the cell is rural built-up land
A_{gov}	*con_f*	0, 1	Whether in the forbidden development zone
	landresource	1-8	Land suitability classified for agriculture
	planning	0, 1	Whether planned as urban built-up land
A_{loc}	*d_bdtown*	≥0	Minimum distance to town boundaries
	d_city	≥0	Minimum distance to new cities
	d_river	≥0	Minimum distance to rivers
	d_road	≥0	Minimum distance to roads
	d_tam	≥0	Minimum distance to Tian'anmen Square

	d_town	≥0	Minimum distance to towns
	d_vcity	≥0	Minimum distance to important new cities
	d_vtown	≥0	Minimum distance to important towns
A_{nei}	*neighbour*	0-1.0	Neighborhood development intensity

2.3 The parcel subdivision framework

Some traditional parcel subdivision tools can automatically subdivide parcels using various algorithms. They highly improve the speed of parcel subdivision process and the corresponding results can provide a good reference to urban planners when establishing urban plans. Unlike these tools, our parcel subdivision framework is used to simulate real parcel subdivision process, which actually would be greatly affected by established urban plans, when urban growth happens. At this stage, we assume the parcel subdivision would be finished before starting the urban growth simulation in our V-BUDEM model, namely not consider the dynamic changing of parcel boundaries. Therefore, the framework of our parcel subdivision methods is used to determine parcel boundaries in the base year, based on which we simulate urban growth.

Specifically, there are four steps for the achievement of our parcel subdivision framework.

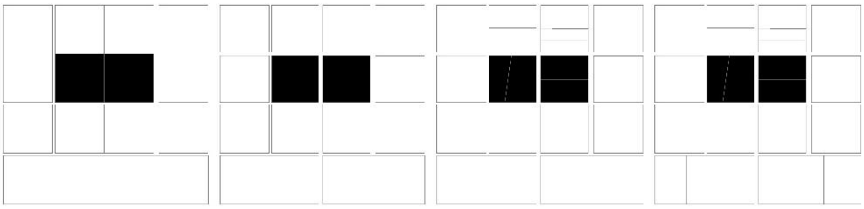
(1) Urban plan is one of key instruments for Chinese government to regulate urban development. A typical master urban plan in China can be regarded as a planned urban form, mainly consisting of three parts: road network, urban built-up parcels, and non-urban built-up parcels. Transport infrastructure is crucial to support various urban development, thereby has high priority to be constructed. Compared with other two parts of a planned urban form, road network is more likely to be achieved. In addition, road network is the basic element of an urban form, and the boundaries of parcels (including blocks) are usually determined by it. Therefore, in our parcel subdivision framework, it’s reasonable to assume that the existing parcels in the base year are firstly subdivided according to the planned road network (Figure 1A). For example, parcel A would be subdivided into two small parcels if a new road is planned to pass through it.

(2) Main social and economic activities of cities happen in urban built-up parcels. To achieve social and economic goals in a planning duration, it’s important to realize the part of a planned urban form consisting of urban built-up parcels. Therefore, the existing parcels then are assumed to be

subdivided according to the boundaries of the planned urban built-up parcels (Figure 1B).

(3) When the existing parcels are subdivided according to the planned road network, the effect of the distribution of non-urban built-up parcels has already been considered to some extent, because parcel boundaries are usually determined by road network. On the other hand, the corresponding existing parcels, which covered by the planned non-urban built-up parcels, usually are non-urban built-up ones. Although they are planned to keep undeveloped, when it comes to real situation of China, which is facing a rapid urbanization, some of them may be developed illegally. For example, people may build their own houses on right-owned parcels, as a result, the non-urban built-up parcels may be changed into urban residential parcels, and the corresponding boundaries would be changed. Because we have already considered the influence of non-urban built-up parcels to some extent, and to reflect some issues (e.g. illegal development) which may happen in real world, the boundaries of the existing parcels, which belong to the planned non-urban built-up parcels, are adopted (Figure 1C).

(4) Some existing large non-urban built-up parcels will be subdivided into some small ones, for they may be developed in a smaller size (Figure 1D). It should be mentioned that, after the urban growth simulation, if a non-urban built-up parcel is not developed, the original parcel boundary would be adopted again.



A B C D

Fig. 1. The parcel subdivision framework

Generally speaking, our approach to subdivide an existing large parcel is to subdivide the Minimum Bounding Rectangle (MBR) of the parent parcel using an algorithm, to clip the subdivided MBR using the parent parcel as a mask, and finally to eliminate the unacceptable generated small parcels. The details are as follows.

(1) Evaluating whether a non-urban built-up parcel is large enough and should be subdivided. The criteria adopted here is simple, shown as equation 5.

$$\mathrm{Area}_i \geq Area_{average} * n \quad (5)$$

Where Area_i is the area of parcel i; $Area_{average}$ is the average area of all existing urban built-up parcels; n is the integrate number, and can be set according to user's experience. If the area of parcel i is no less than $Area_{average} * n$, meaning parcel i is a large parcel, and should be subdivided.

(2) Determining the area of the subdivided parcel $Area_{sub}$. In our tool, we assume the areas of all subdivided parcels are the same, and equal $Area_{average}$. Existing researches usually generate the road first, then subdivide the parent parcel according to the value of $Area_{sub}$. Different from these, we don't consider the road generation, but assume the created subdivided parcel has contain the road area, which can be used to support the connectivity of subdivided parcels. Therefore, the step of road generation is emitted in our method, but the area of subdivided parcel should contain the road area. Showing as equation 6.

$$Area_{sub}^{'} = Area_{sub} + Area_{road} = Area_{sub} * (1 + \partial) \quad (6)$$

Where $Area_{sub}^{'}$ is the area containing the road area, $Area_{road}$ is the corresponding road area, and ∂ is the ratio of parcel and road areas. The method has an assumption that all subdivided parcels need to be adjacent to the road. This adoption can speed up the parcel subdivision process, considering the creation of road is comparatively complicated. The simulation result can still represent the urban growth trend based on this method.

(3) Creating the MBR of the parent parcel (Figure 2B); the parcel's MBR can reflect its direction.

(4) Subdividing the MBR according to $Area_{sub}^{'}$ (Figure 2C). The row number (Row_{num}) and column number ($Column_{num}$) of the MBR are determined by equation 7

$$\begin{aligned} & t = \sqrt{Area_{MBR} / Area_{sub}^{'}} \\ & if\ t < \text{int}(t) + 0.5: \\ & \quad Row_{num} = Column_{num} = \text{int}(t) \qquad (7) \\ & else: \\ & \quad Row_{num} = Column_{num} = \text{int}(t) + 1 \end{aligned}$$

Where, $Area_{MBR}$ is MBR's area.

(5) Clipping the subdivided MBR using the parent parcel as a mask (Figure 2C).

(6) Eliminating unacceptable small parcels, resultant of clipping, by merging them into adjacent parcels (Figure 2D). The criteria to determine whether a small parcel should be eliminated is by equation 8.

$$Area_{created} \le Area_{sub}^{'} / m \quad (8)$$

Where, $Area_{created}$ is the area of a created small parcel; m is the integrate number, and can be set according to user's experience. The criteria to determine which adjacent parcel would be selected to merge with the small created parcel is that the parcel which has the longest shared border.

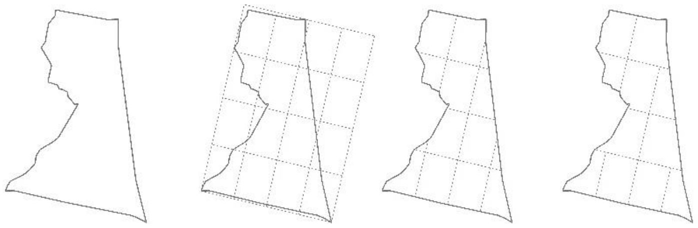

A B C D

Fig. 2. The method to subdivide a large parcel

2.4 The simulation procedure

The V-BUDEM's simulation procedure is shown in Figure 3. $Parcel_{i,\max}$ is the parcel i, which has the maximum probability to be developed among all remaining non-urban built-up parcels in iteration t, and $Area_{i,\max}$ is its area. In each iteration, the final value of $deveArea^{t}$ is always less than that of $stepArea^{t}$ (although very close). As a result, if $t = n$, the value of $stepArea^{t}$ is determined by $LandDemand - \sum_{t} deveArea^{t}$, in which $\sum_{t} deveArea^{t}$ is the total developed area during the finished $n-1$ iterations. This consideration can promise the final developed area is close to $LandDemand$.

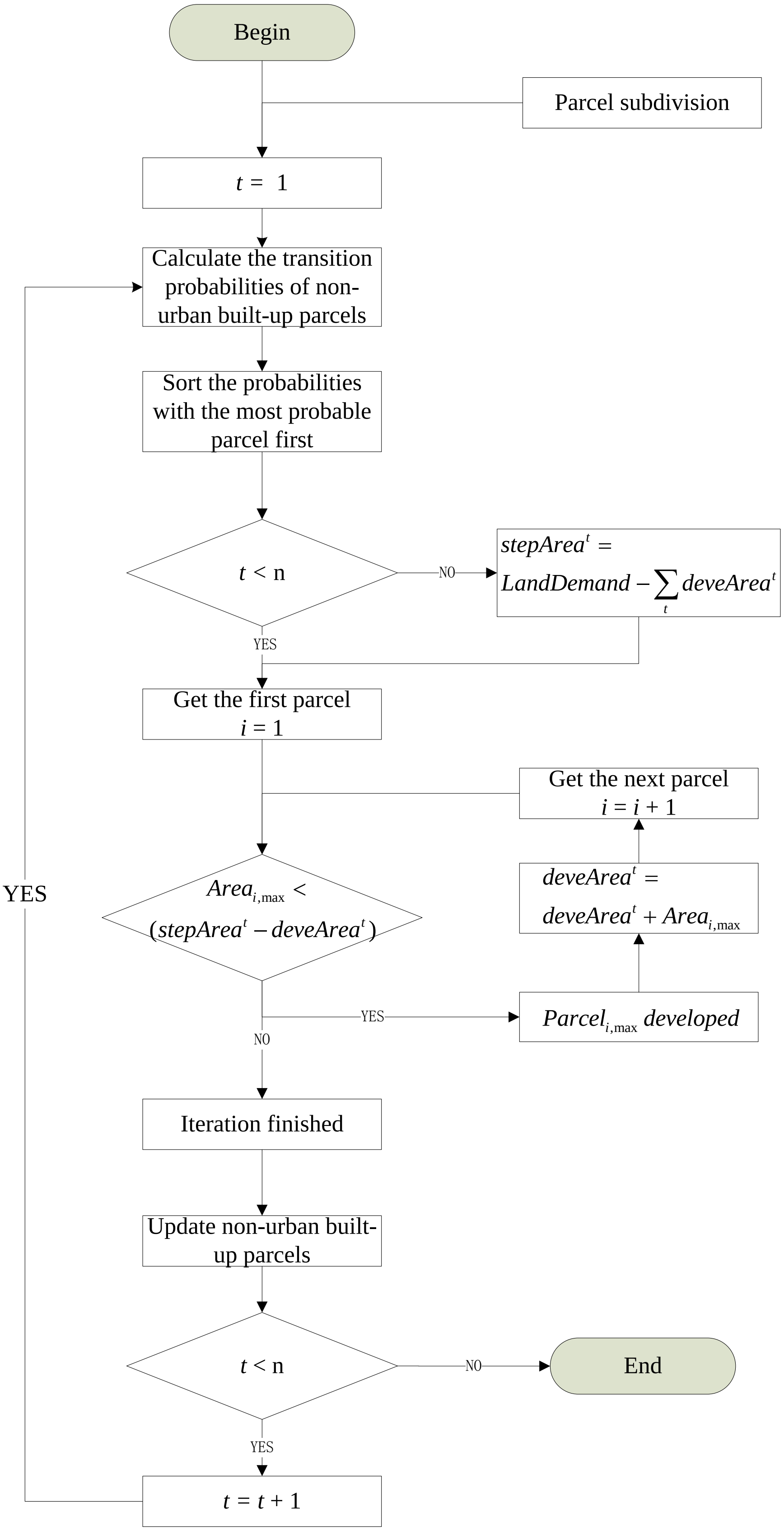
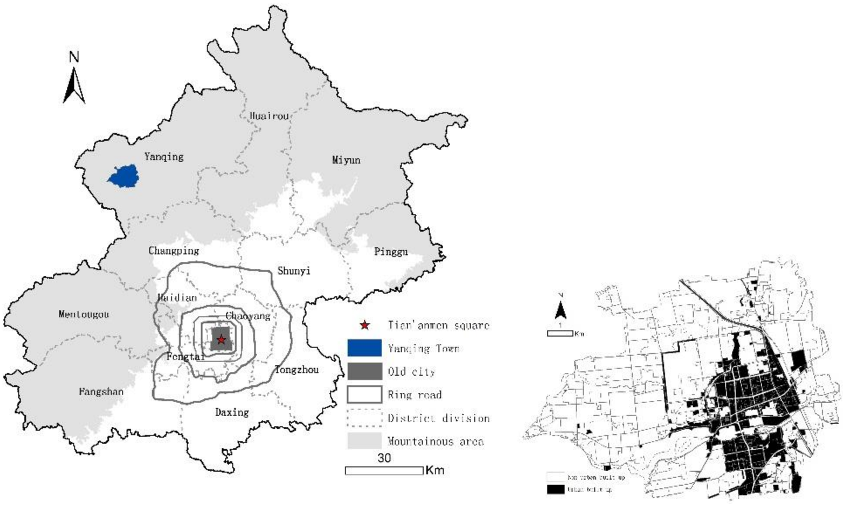

Begin
Parcel subdivision
$t = 1$
Calculate the transition probabilities of non-urban built-up parcels
Sort the probabilities with the most probable parcel first
$t < n$
NO
$stepArea^t =$
$LandDemand - \sum_t deveArea^t$
YES
Get the first parcel
$i = 1$
Get the next parcel
$i = i + 1$
$Area_{i,\max} <$
$(stepArea^t - deveArea^t)$
$deveArea^t =$
$deveArea^t + Area_{i,\max}$
YES
$Parcel_{i,\max}$ developed
NO
Iteration finished
Update non-urban built-up parcels
$t < n$
NO
End
YES
$t = t + 1$
YES

Fig. 3. The V-BUDEM simulation process

3 Model application

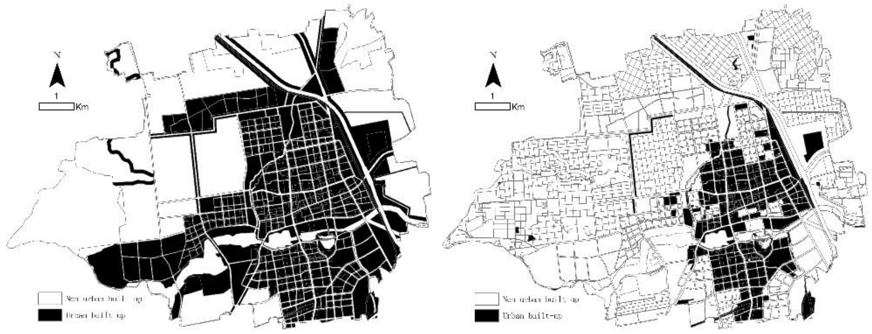
3.1 Study area

The whole study area of V-BUDEM is the Beijing Metropolitan Area (BMA) (Figure 4A), but in this paper, Yanqing Town (Figure 3) was chosen as a case area to illustrate the model's applicability. The BMA has an area of 16,410 km^2. It has experienced rapid urbanization in terms of GDP and population growth since the Reform and Opening Policy of 1978, established by Chinese central government. There are 16 districts under BMA jurisdiction. In 2013, the total GDP was 1 950.06 billion Yuan, with an average of 93.213 thousand Yuan per capita, and the total residential population was 21.148 million (Beijing Municipal Bureau of Statistics, 2014). In 2010, the area of the BMA urban built-up land was 1758 km^2, about 3.5 times than that in 1976, which was 495 km^2. Yanqing (Figure 4B), a small town located in the northwestern of the BMA, has a distance of about 74 km from Beijing's city center. In 2011, the residential population is 37 thousand (Yanqing Government, 2012). In 2010. Yanqing has an area of 67.36 km^2 (including 12.22 km^2 urban built-up land), and consists of 1121 parcels (including 604 urban built-up parcels).



A B

Fig. 4. The Beijing Metropolitan Area and Yanqing Town



A B

Fig. 5. Yanqing's land use pattern in 2010 and land use plan in 2020

3.2 Data

The input spatial data is classified into six types, namely LANDUSE, PLANNING, CONSTRAINT, LANDRESOURCE, LOCATION, and BOUNDARY. All the spatial data was converted into Personal Geodatabase Feature Class, using the same coordinate and projection system. Explanations of these data are shown as follows.

(1) LANDUSE: In this test, we only used Yanqing's land use pattern in 2010 (corresponding to existing urban form, shown in Figure 5A). the data is from the database of Beijing Institute of City Planning (BICP). Originally, there were seven land use types, including urban built-up land, rural built-up land, transport land, agriculture land, forest land, wetland land, and vacant land. Then, they were integrated into three land use types, including urban built-up land, rural built-up, and other land. The variable *isagri* and *isrural* (Figure 6A-B) were derived from the LANDUSE dataset.

(2) PLANNING: In this test, Yanqing's planned urban form, shown in Figure 5B, is from the Beijing Urban Master Plan (2004-2020). The variable *planning* (Figure 6E) is derived from the PLANNING dataset with 1 for urban built-up land and 0 for non-urban built-up land. In 2020, there are 30.07 km^2 for urban built-up land, with 975 urban built-up parcels.

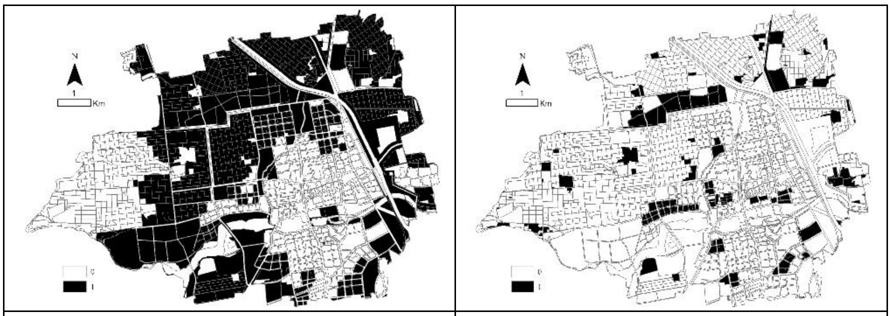
(3) CONSTRAINT: The data, from the Beijing Limited Construction Zone Plan (Long et al. 2006), reflects urban development constraints derived from 110 spatial layers of natural resource protection and hazard prevention according to current laws, legislations, and standards of China. Beijing was zoned into the

forbidden built-up areas, constrained built-up areas, and suitable built-up areas. If the value of variable *con_f* (Figure 6C) is 1, the parcel is located in forbidden built-up areas.

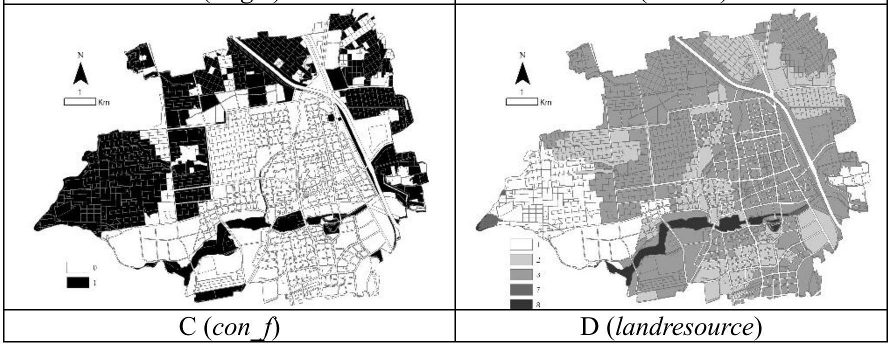
(4) LANDRESOURCE: The data, from the BICP's database, indicates the suitability for the agricultural use and was classified into eight types, ranging from 1 to 8. The variable *landresource* (Figure 6D) is derived from LANDRESOURCE with the same value.

(5) LOCATION: The data, including the distributions of centers at various levels, road network, and river resource distribution, are from BICP's database. The location variable, e.g. *d_tam*, *d_vcity*, *d_city*, *d_vtown*, and *d_town*, indicates the minimum distance to city-level or town-level centers for various administrative divisions. The location variable *d_road* indicates the minimum distance to the road network and *d_river* as the minimum distance to the rivers. The location spatial variables were derived from LOCATION using the Distance/ Straight Line command in ArcGIS 10.2, and shown in Figure 6G-M.

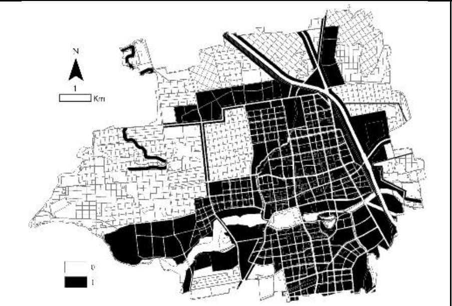
(6) BOUNDARY: The data, from BICP's database, includes various boundaries, such as administrative, ring road, eco-zoning, and watershed boundaries. In Yanqing test, the variable *d_bdtown* (Figure 6F) which means the minimum distance to town boundaries, is considered.



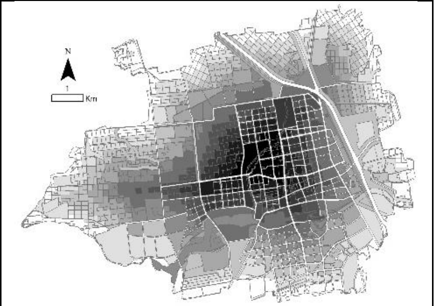
A (*isagri*) B (*isrural*)



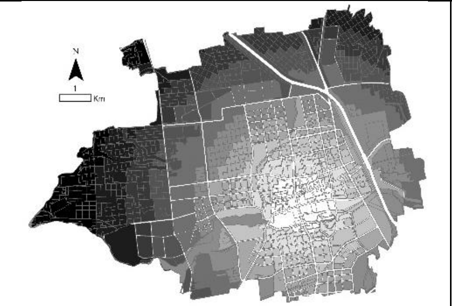
C (*con_f*) D (*landresource*)



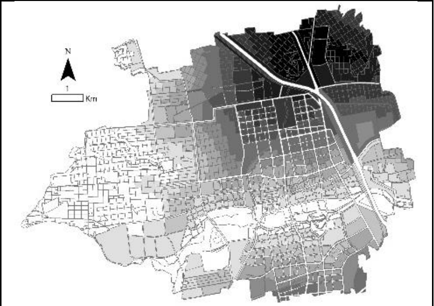
E (*planning*)



F (*d_bdtown*)

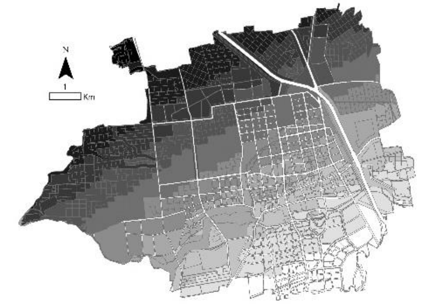


G (*d_city*)



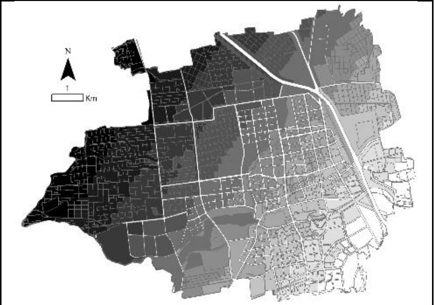
H (*d_river*)

I (*d_road*)

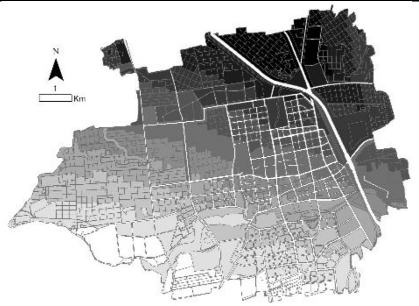


J (*d_tam*)

K (*d_town*)



L (*d_vcity*)



M (*d_vtown*)

Fig. 6. The spatial distributions of static constraints data

3.3 Yanqing2020 simulation

In this subsection, we simulate urban growth scenario "Yanqing2020", from 2010 to 2020. Before the simulation, existing parcels were subdivided according to existing and planned urban forms using our parcel subdivision framework. Figure 6 shows the corresponding subdivided existing urban form. After subdivision, there are 9.76 km^2 for urban built-up land, a little smaller than that before subdivision (12.22 km^2). There are 2201 parcels in total, with 617 urban built-up parcels, which is a little more than that before subdivision (604 parcels).

Secondly, time step was set as 5 times with a total of ten years. According to the planned urban form in 2020, there are 30.07 km^2 (namely $LandDemand$) for urban built-up land. Assuming the area which would be developed in each time step is the same, $stepArea$ is 6.01 km^2. In addition, α was set as 3, and k was set as 20 empirically.

Thirdly, parameters of constraint variables should be estimated according to different historic phases. Considering there are only 1121 parcels in Yanqing, thereby the parameter set estimation may not significant and convincing. Additionally, the objective of this test is to show the feasibility, especially of the part of parcel subdivision framework in the V-BUDEM model, with no attention to identify suitable and accurate parameter at this stage. Assuming the constraint influence of each x_k is the same between Yanqing and the whole BMA, we adopted the parameter set (excluding the weight of neighbor variable x_n), with the whole accuracy of 96.8%, directly from Long et al. (2009), see Table 2. In the logistic regression for its identification, dependent variables x_k (excluding neighbor variable) were calibrated via algebra operation on the BMA's existing urban form in 2006 and planned urban form in 2020. Each independent variable is significant at the acceptable level. Detailed explanations can be found in Long et al. (2009).

Table 2. Policy parameter set for 2006-2020

Name	Coefficient	Name	coefficient
isrural	6.886 21***	*d_river*	-0.000 52***
Isagri	6.971 87***	*d_road*	-0.000 96***
d_tam	-0.000 10***	*d_bdtown*	-0.000 27***
d_vcity	-0.000 03***	*planning*	8.770 71***
d_city	-0.000 10***	*con_f*	-0.200 97*
d_vtown	-0.000 28***	*landresource*	-0.093 55**
d_town	-0.000 11***		

Note: ***p (significance) = 0.001; **p = 0.05; *p = 0.5

To identify neighbor parameter x_n, the neighbor distance, which may affect simulation result greatly, should be determined at first. We set it as 200 m empirically. In the process of identifying x_n via MonoLoop method, x_n is sampled from 0 to 5, with an interval of 0.5 empirically. According to the simulated results of Kappa index, shown in Figure, x_n was set as 2.5, with a Kappa value of 81.16. The weights x_k, obtained by the logistic regression, x_n from MonoLoop , and with other parameters were then input into the established transition rules to simulate urban growth scenario Yanqing2020. The simulated urban form was shown in Figure 7.

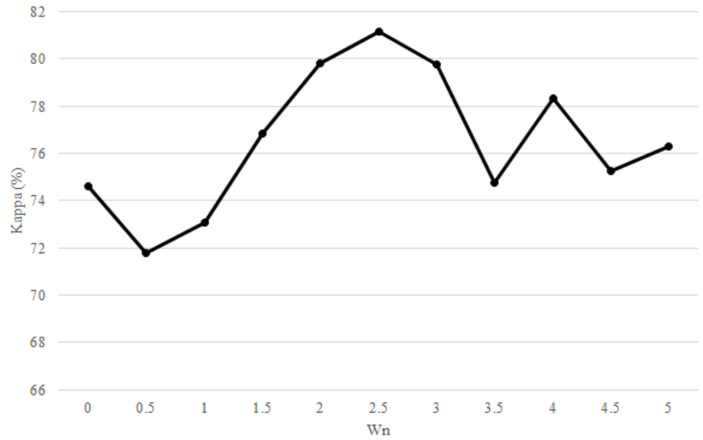


Fig. 7. Kappa values of the MonoLoop procedure

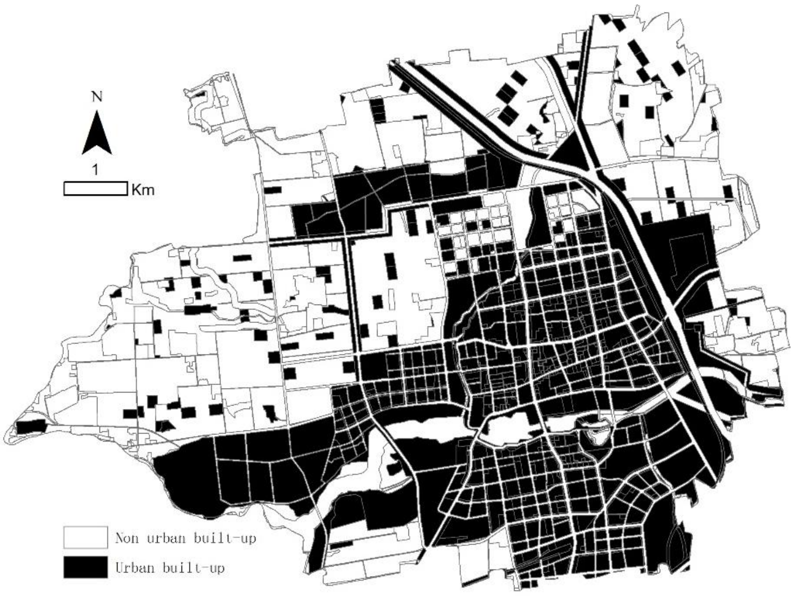


Fig. 8. Simulated urban form in 2020

The Kappa value is 81.16, showing the simulated urban form was fitted well with the planned urban form. In the simulated urban form in 2020, there are 1517 parcels in total. Among them, there are 1066 urban built-up parcels, with a total area of 30.05 km^2, slightly smaller than predicted LandDemand 30.07 km^2. Specifically, 903 urban built-up parcels are located in the planned urban areas, while 163 urban built-up parcels located in the planned non-urban areas. When it comes to the parcels adopting the boundaries of the planned urban built-up parcels, most of them are developed into urban built-up land, except some mainly located in middle-northern part of Yanqing Town. The fact shows it's reasonable to adopt the boundaries of the planned urban built-up parcels when subdividing existing parcels. Additionally, there are some parcels, adopting existing parcel boundaries in 2010 or subdivided from large ones, have been developed into urban built-up land, which would be unlikely to happen if we adopt planned non-urban built-up parcel boundaries, which usually have large size thereby not suitable to be regarded as the basic spatial units when simulating. The simulated result shows the vector-based CA method is feasible and our parcel subdivision framework is reasonable for simulating urban growth.

4 Conclusion and discussion

BUDEM (Beijing Urban Development Model) is a raster-based CA model for supporting city planning and corresponding policies evaluation in Beijing (Long et al. 2009). The model has been proved to be capable of analyzing historical urban growth mechanisms and predicting future urban growth in Beijing. In this paper, we improved the initial raster-based BUDEM model into a vector-based one, V-BUDEM. In the V-BUDEM model, urban space consists of irregular land parcels, and a parcel's neighborhood is defined as all parcels surrounding it within a certain distance. The model adopted a constrained CA method, and considered four constraints: self-status, government, locational, and neighborhood constraints. Fourteen constraint variables were selected for simulating future scenarios. Additionally, a framework of parcel subdivision was adopted in this model to subdivide existing parcels. At this stage, we assume parcel subdivision would be achieved before starting the urban growth simulation, rather than consider a dynamic subdivision process during the simulation. Specifically, the parcel subdivision framework consists of four steps: (1) subdividing existing parcels according to road distribution in the urban plan; (2) subdividing existing parcels according to the boundaries of planned urban built-up parcels; (3) adopting the boundaries of the existing parcels, which belong to planned non-urban built-up parcels; and (4) subdividing existing large non-urban built-up parcels using an automated parcel subdivision tool developed by us. If some small parcels subdivided from an existing large non-urban built-up one in step (4) are not developed after simulation, they will be merged, and the original boundary will be adopted again. In this paper, after the description of the V-BUDEM, We tested the model to simulate urban growth scenario of Yanqing Town, located in northwestern Beijing, from 2010 to 2020. The Kappa value is 81.16, showing the simulated urban form was fitted well with the planned urban form. Most of the parcels adopting planned urban built-up parcel boundaries were developed into urban built-up land. And some parcels, adopting existing parcel boundaries in 2010 or subdivided from large ones, have been developed into urban built-up land, which would be unlikely to happen if we adopt planned non-urban built-up parcel boundaries directly. The simulated result shows the vector-based CA method is feasible and our parcel subdivision framework is reasonable for simulating urban growth.

The main contributions of this study are as follows: (1) the model adopts a vector-based CA method using land parcels to represent urban space, composing the landscape a user would perceive as meaningful, and can simulate urban growth in a way more close to real world situation; (2) the model integrates a process of parcel subdivision, and the proposed parcel subdivision framework comprehensively considered the impacts of existing parcel boundaries in the existing urban form and planned parcel boundaries in the planned urban form, and especially developed a straight-forward and automatic parcel subdivision tool, included in the framework as the fourth step, to partition existing large parcels; (3) compared with other urban models, V-BUDEM is developed specifically for urban planning applications, e.g., it can be used to identify policies required for implementing the planned urban form desired by planners and decision makers.

In the future, several works can be done to improve V-BUDEM study. At first, the model application would be expanded from Yanqing Town to the whole Beijing metropolitan area, including identifying urban growth mechanisms in various historical phases, retrieving urban policies needed to implement the planned urban form in 2020, and predicting urban growth scenarios in the future (e.g. 2030) using parameter set identified from historical urban forms. Secondly, the fourth step of our parcel subdivision framework would be improved, so as to be an independent automatic subdivision tool, which can be used to speed up parcel subdivision during the process of planning establishment, just like the tools proposed by other literature (e.g. Wickramasuriya et al. 2011). At present, we are developing this tool. Thirdly, the detailed urban form, showing the distributions of commercial, industrial, and residential parcels, can be established using Planner Agents (Zhang and Long 2013; Long and Zhang 2014). And finally, we would calculating transport energy consumption in the established detailed urban form based on Agenter (Long and Shen 2013) and FEE-MAS model (Long et al. 2013). Comparing the energy consumption among different established forms, a low carbon form, creating the least energy consumption, can be identified, which is valuable to support the construction of low carbon cities.

Acknowledgments **The first author would like to thanks for the financial support from the Beyond the Frontiers Programme of Radboud University Nijmegen.**

References

Alexandridis K, Pijanowski BC, Lei Z (2007) Assessing multiagent parcelization performance in the MABEL simulation model using Monte Carlo replication experiments. Environment and Planning B: Planning and Design 34:223-244

Batty M, Xie Y, Sun Z (1999) Modeling urban dynamics through GIS-based cellular automata. Computers, Environment and Urban Systems 23:205-233

Beijing Municipal Bureau of Statistics (2014) Beijing 2013 National Economic and Social Development Statistics Bulletin. Retrieved from: http://www.bjstats.gov.cn/xwgb/tjgb/ndgb/201402/t20140213_267744.htm

Yanqing Government (2012) Population Statistics. Retrieved from: http://www.bjyq.gov.cn/sy/yqgk/rkzk/

Chen J, Jiang J (2000) An event-based approach to spatio-temporal data modeling in land subdivision systems. GeoInformatica 4:387-402

Chen Q, Mynett AE (2003) Effects of cell size and configuration in cellular automata based prey-predator modelling. Simulation Modelling Practice and Theory 11:609-25

Clarke KC, Gaydos LJ (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 12:699-714

Cohen J (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20:37-46

Congalton R, Mead RA (1983) A quantitative method to test for consistency and correctness in photointerpretation. Photogrammetric Engineering & Remote Sensing 49:69-74

Dahal KR, Chow TE (2014) A GIS toolset for automated partitioning of urban lands. Environmental Modelling & Software 55:222-234

Demetriou D, See L, Stillwell J (2013) A spatial genetic algorithm for automating land partitioning. International Journal of Geographical Information Science 27:2391-2409

Demetriou D (2014) The development of an integrated planning and decision support system (IPDSS) for land consolidation. Springer, Berlin

Easa SM (2008) Unified direct method for land subdivision: circular sides permitted. Journal of Surveying Engineering 134:55-60

ESRI (2013) Creating shapes from graph networks. Retrieved from: http://cehelp.esri.com/help/index.jsp?topic=/com.procedural.cityengine.help/html/manual/is/create/block_params.html

Guan Q, Wang L, Clarke KC (2005) An artificial-neural-network-based constrained CA model for simulating urban growth. Cartography and Geographic Information Science 32:369-380

Hammam Y, Moore A, Whigham P (2007) The dynamic geometry of geographical vector agents. Computers, Environment and Urban Systems 31:502-519

Hu SY, Li DR (2004) Vector cellular automata based geographical entity. Geoinformatics 249-256

Jantz CA, Goetz SJ (2005) Analysis of scale dependencies in an urban land‐use‐change model. International Journal of Geographical Information Science 19:217-241

Jenerette GD, Wu J (2001) Analysis and simulation of land-use change in the central Arizona–Phoenix region, USA. Landscape Ecology 16:611-626

Jjumba A, Dragićević S (2012) High resolution urban land-use change modeling: Agent iCity approach. Applied Spatial Analysis and Policy 5:291-315

Ko DW, He HS, Larsen DR (2006) Simulating private land ownership fragmentation in the Missouri Ozarks, USA. Landscape ecology 21:671-686

Landis JD (1995) Imagining land use futures: applying the California urban futures model. Journal of the American Planning Association 61:438-457

Liu Y (2012) Modelling sustainable urban growth in a rapidly urbanising region using a fuzzy-constrained cellular automata approach. International Journal of Geographical Information Science 26:151-167

Long Y, Han H, Lai S-K, Mao Q (2013a) Urban growth boundaries of the Beijing Metropolitan Area: Comparison of simulation and artwork. Cities 31:337-348

Long Y, Mao Q, Dang A (2009) Beijing urban development model: Urban growth analysis and simulation. Tsinghua Science & Technology 14:782-794

Long Y, Mao Q, Shen Z (2013b) Urban Form, Transportation Energy Consumption, and Environment Impact Integrated Simulation: A Multi-agent Model. In: Shen Z (ed) Spatial planning and sustainable development. Springer, Berlin, p, 227-247

Long Y, Shen Z (2013) Disaggregating heterogeneous agent attributes and location. Computers, Environment and Urban Systems 42:14-25

Long Y, Shen Z, Mao Q (2012) Retrieving spatial policy parameters from an alternative plan using constrained cellular automata and regionalized sensitivity analysis. Environment and Planning B: Planning and Design 39:586-604

Long Y, Zhang YP (2014) Land use pattern scenario analysis using Planner Agents. Environment and Planning B: Planning and Design forthcoming

Marceau DJ, Wang F, Wijesekara N (2013) Investigating land-use dynamics at the periphery of a fast-growing city with cellular automata at two spatial scales. In Modeling of land-use and ecological dynamics. Springer, Berlin, p, 51-79

Menard A (2008) VecGCA: a vector-based geographic cellular automata model allowing geometric transformations of objects. Environment and Planning B: Planning and Design 35:647-665

Moreno N, Wang F, Marceau DJ (2009) Implementation of a dynamic neighborhood in a land-use vector-based cellular automata model. Computers, Environment and Urban Systems 33:44-54

Moghadam H, Helbich M (2013) Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model. Applied Geography 40:140-149

O Sullivan D (2001a) Graph-cellular automata: a generalised discrete urban and regional model. Environment and Planning B: Planning and Design 28:687-706

O'Sullivan D (2001b) Exploring spatial process dynamics using irregular cellular automaton models. Geographical Analysis 33:1-18

Pinto NN (2010) A cellular automata model based on irregular cells: application to small urban areas. Environment and Planning B: Planning and Design 37:1095-1114

Semboloni F (2000) The growth of an urban cluster into a dynamic self-modifying spatial pattern. Environment and Planning B: Planning and Design 27:549-564

Shafizadeh MH, Helbich M (2013) Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model. Applied Geography 40:140-149

Shen Z, Kawakami M, Kawamura I (2009) Geosimulation model using geographic automata for simulating land-use patterns in urban partitions. Environment and Planning B: Planning and Design 36:802-823

Shi W, Pang MYC (2000) Development of Voronoi-based cellular automata-an integrated dynamic model for Geographical Information Systems. International Journal of Geographical Information Science 14:455-474

Stevens D, Dragicevic S (2007) A GIS-based irregular cellular automata model of land-use change. Environment and Planning B: Planning and Design 34:708-724

Torrens PM, Benenson I (2005) Geographic automata systems. International Journal of Geographical Information Science 19:385-412

Vanegas CA, Aliaga DG, Benes B et al (2009a) Visualization of simulated urban spaces: Inferring parameterized generation of streets, parcels, and aerial imagery. Visualization and Computer Graphics, IEEE Transactions 15:424-435

Vanegas CA, Aliaga DG, Beneš B et al (2009b) Interactive design of urban spaces using geometrical and behavioral modeling. In ACM Transactions on Graphics (TOG). (ed.), Vol. 28, p. 111, ACM

Wakchaure AS (2001) An ArcView tool for simulating land subdivision for build out analysis. Polytechnic Institute and State University, Virginia.

Wang F, Marceau DJ (2013) A Patch‐based Cellular Automaton for Simulating Land‐use Changes at Fine Spatial Resolution. Transactions in GIS 17:828-846

Ward DP, Murray AT, Phinn SR (1999) An optimized cellular automata approach for sustainable urban development in rapidly urbanizing regions. International Journal of Geographical Information Science 7:235-250

White R, Engelen G (2000) High-resolution integrated modelling of the spatial dynamics of urban and regional systems. Computers, Environment and Urban Systems 24:383-400

Wickramasuriya R, Chisholm LA, Puotinen M et al (2011) An automated land subdivision tool for urban and regional planning: Concepts, implementation and testing. Environmental Modelling & Software 26:1675-1684

Wickramasuriya R, Chisholm LA, Puotinen M et al (2013) A method to dynamically subdivide parcels in land use change models. International Journal of Geographical Information Science 27:1497-1513

Zhang YP, Long Y (2013) Land use pattern scenario analysis using Planner Agents: A preliminary study. In 13th International Conference on Computers in Urban Planning and Urban Management, Utrecht, The Netherlands, 2-5 July