Research paper

Modeling the urban landscape dynamics in a megalopolitan cluster area by incorporating a gravitational field model with cellular automata

Chunyang He\textsuperscript{a,c}, Yuanyuan Zhao\textsuperscript{b}, Jie Tian\textsuperscript{c}, Peijun Shi\textsuperscript{a}

\textsuperscript{a} State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China
\textsuperscript{b} College of Resources Science \& Technology, Beijing Normal University, Beijing 100875, China
\textsuperscript{c} Department of International Development, Community, and Environment, Clark University, 950 Main Street, Worcester, MA 01610, United States

HIGHLIGHTS

- This paper proposes a new megalopolis landscape dynamic model by combining a gravitational field model with a CA model.
- The model proposed produced more accurate simulation results than the CA model, which did not account for urban flows.
- The model proposed is of great value for locating ‘hotspots’ of future urbanization in a megalopolis cluster area.
- The model proposed can provide valuable information for regional landscape planning and environmental management in a megalopolis cluster area.

ARTICLE INFO

Article history:
Received 23 April 2012
Received in revised form 13 January 2013
Accepted 14 January 2013

Keywords:
Beijing–Tianjin–Tangshan megalopolitan cluster area
Cellular automata model
Gravitational field model
Megalopolitan cluster area
Urban flow
Urban landscape dynamics

ABSTRACT

The effective modeling of the urban landscape dynamics in a megalopolitan cluster area (MCA) is essential to understanding its spatial evolution process. However, existing urban landscape dynamic models based on cellular automata (CA) are limited in that they do not consider urban flows (e.g., flows of people, material, and information) between the different cities/towns in an MCA. This paper proposes a new megalopolitan landscape dynamic model (MLDM) that is better suited for simulating the urban landscapes in an MCA by combining a gravitational field model (GFM) with a CA model. The GFM was used to model the influence of inter-city urban flows and to refine the transition rules of the CA model. The MLDM was applied to simulate the urban landscape in the MCA of Beijing–Tianjin–Tangshan, and produced more accurate simulation results than the CA model that did not account for urban flows. The MLDM-based prediction of future landscapes suggested that urbanization will continue in the region through 2020, especially in a few ‘hotspot’ areas. Close attention should be paid to these areas for strategic regional planning and environmental protection in this heartland of China.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The term megalopolitan cluster area (MCA) refers to a region comprising a considerable number of cities clustered around the regional economic core of one or two super-large cities. Although these cities can significantly differ one from another, they are closely interconnected and have strong interactions due to the highly developed regional transportation and information networks (Lang & Knox, 2009; Yao, Chen, & Zhu, 2006). MCAs are often the most vigorously developing urban areas around the world and attract a great deal of attention (Gu, Hu, Zhang, Wang, & Guo, 2011; Tian, Jiang, Yang, & Zhang, 2011a; Vicino, Hanlon, & Short, 2007). For example, the MCA of Washington–Baltimore, Boston, Philadelphia, and New York covers only 1.4% of US territory but is home to 17.3% of the US population (Vicino et al., 2007). In China, the three major MCAs (Yangtze, Pearl River Delta and Beijing–Tianjin–Tangshan) comprise only 2.2% of the country’s territory but are home to 12% of the Chinese population and contribute 34% of China’s gross domestic product (GDP) (NSBC, 2010). These MCAs have become highly urbanized in the past decades, and the local natural landscapes have been removed or considerably changed. This urbanization process has resulted in enormous ecological and environmental issues, such as increased surface temperature and decreased biodiversity (Chen, Li, Zheng, Guan, & Liu, 2011; Foley, DeFries, & Asner, 2005; Grimm et al., 2008; Jenerette & Wu, 2001; Nikanorov, Khoruzhaya, & Mironova, 2011; Valiela & Martinetto, 2007; Yang, Hou, & Chen, 2011). Spatial process models are important tools for better understanding the driving forces to urban landscape dynamics and evaluating their ecological and environmental impacts (Valbuena, Verburg, Bregt, & Ligtzenberg, 2010). There has been...
a great need for models that can effectively simulate the urban landscape dynamics of MCAs to support the urban/regional planning and the sustainable development in these areas (Berling-Wolf & Wu, 2004a; Wu & David, 2002; Wu, 2010).

At present, the simulation of urban landscape dynamics mainly relies on four types of models (Berling-Wolf & Wu, 2004a; Parker, Manson, Janssen, Hoffmann, & Deadman, 2003). The first type describes urban structure and morphology in a qualitative way; most models of this type were developed before the 1940s, including concentric zone model, sector model, and multiple nuclei model (He, Okada, Zhang, Shi, & Li, 2008). The second type is built upon Newton's theory of gravity, and focuses on depicting the spatial interactions between different entities. Typical examples include the gravity model (Foot, 1981) and the Lowry model (Harris, 1985). The third type uses differential equations to represent the spatial process; the models of this type were initially proposed in the 1970s with the system dynamic models being representative (Gilbert & Troitzsch, 1999). The last type includes the discrete dynamic models that have flourished since the 1970s; the best known examples include cellular automata (CA) (Santé, García, Miranda, & Crecente, 2010) and multi-agent system (Parker et al., 2003) models. CA models have been widely applied to simulate urban landscape dynamics in the past two decades (Berling-Wolf & Wu, 2004b; He et al., 2006a; He, Okada, Zhang, Shi, & Zhang, 2006b; He et al., 2008). Numerous works have successfully demonstrated the capability of CA models in representing the complex process of urban landscape evolution (Barredo, Kasanko, McCormick, & Lavalle, 2003; Fang, Gertner, Sun, & Anderson, 2005; Li & Yeh, 2002; Mitsova, Shuster, & Wang, 2011; Sui & Zeng, 2001). However, a CA model typically assumes that the regional pattern change is an aggregate of the local changes, and simulates the evolution of an urban ecosystem based merely on local interactions (Qi et al., 2004). This approach focuses on micro individual cells and often neglects the links and interactions between cells and the macro spatial patterns (Qi et al., 2004). In addition, the ‘bottom-up’ construction of traditional CA models usually prevents them from fully capturing the macro-scale socioeconomic driving forces of urban landscape dynamics (Ward, Murray, & Phinn, 2000; White & Engelen, 1997; White & Engelen, 2000). Therefore, researchers have attempted to extend and enhance traditional CA models by incorporating socioeconomic models. For instance, White and Engelen (2000) linked a CA model with a regional development model to simulate the evolution of urban landscapes for the entire country of the Netherlands. Wu and David (2002) integrated hierarchical theory with a CA model to simulate the urban landscape dynamics in the metropolitan area of Phoenix, AZ, USA. He et al. (2008) incorporated a potential model with a CA model to simulate the landscape dynamics in the Beijing metropolitan area. Most recently, Kuan (2011) coupled a CA model with an artificial neural network model to simulate the urban landscape changes in the Beijing–Tianjin–Tangshan region, China. These researches have significantly enhanced the ability of CA models to accurately simulate landscape dynamics in an urban area.

However, most of the existing models are focused on one single city and not suitable for modeling the urban landscape dynamics in an MCA. In order to solve this problem, Li and Liu (2006) proposed an extended CA model in which the transition rules were determined by using a case-based reasoning (CBR) technique. The idea is to build a case library based on the geographical data including land use/cover, and then determine the state of a cell state according to the proximity between the cell and the cases in the library calculated. The advantages of this model include not having to define the transition rules and being possibly able to simulate for complex situations. The disadvantages include: (1) the model accuracy is highly sensitive to the representativeness of the case data; (2) the model, once established, may not be effectively applied to a different study area; and (3) the physical meaning of the model is not so clear.

Urban flow refers to the frequent interflows of people, material, information, capital, and technologies among a cluster of cities/towns (Zhu & Yu, 2002). There exists a variety of transportation networks including highways, expressways, waterways and airways within and between the urban areas in an MCA. The socio-economics of the cities/urban areas of different sizes strongly interact with each other, collectively driving the regional development of an MCA. Therefore, urban flows often play a significant role in the evolution of an MCA (Gu et al., 2011; Limtanakool, Schwanen, & Dijst, 2007; Seto & Fragkias, 2005). Urban flow intensity characterizes the strength of the influence from the centralization and decentralization of the cities in an MCA. Zhu and Yu (2002) defined the concept of urban flow intensity and used it in analyzing the urban flows in the HuNingHang megalopolitan cluster. Cao and Wang (2007) analyzed the intensity of the urban flows in a city cluster of Northeast China. These studies have suggested that urban flow intensity is an effective indicator of the strength of the socioeconomic interactions between the member cities/towns in a megalopolitan cluster area (Derudder & Taylor, 2005; Yao et al., 2006).

The gravitational field model (GFM) was first proposed by Lagrange in 1773 to measure the gravitational influence of a planet from a distance (Stewart, 1947). In the 1940s, Stewart adopted the GFM to represent the influence of population as a function of distance and applied it for the first time in socioeconomic research (Brown, 1982; Gardiner, Martin, & Tyler, 2011; Stewart, 1941, 1942, 1947). In the 1980s, Friedman (1986) treated an MCA as an urban field in which the spatial influence of one city on another is inversely related to the distance between them. Liang (2009) employed the GFM to calculate the local/regional impact of 670 Chinese cities with non-agricultural populations exceeding 80,000. Wu, Zhang, Jin, & Deng (2009) used a GFM to study the spatial interactions among 18 cities in the coastal area of eastern China during 1991–2002, by analyzing the economic, population, and transportation data. Most recently, Wang, Deng, Liu, & Wang (2011) used a GFM to investigate the urban expansion in central China from 1990–2007. These studies have shown that GFMs are effective in representing the spatial interactions among cities. However, a GFM alone is inadequate for representing the complex spatial processes of the urban landscape evolution in an MCA because it cannot well capture the micro-scale dynamics.

This paper presents a new megalopolitan landscape dynamic model (MLDM) that can better model the urban landscape dynamics in MCAs. The model basically integrates a GFM with a CA model. In the MLDM, the demand for urban land in an MCA is first estimated based on the socioeconomic data. Then, a GFM is employed to predict the urban flows within the MCA. Next, the GFM is incorporated in defining the transition rules of a CA model to more accurately map the potential new urban cells that are to meet the urban land demand of the MCA. A case study was conducted by applying the model to simulate the urban landscape dynamics in the Beijing–Tianjin–Tangshan megalopolitan cluster area (BTT-MCA) of China for the time period of 1990–2009 and to predict the ‘hotspot’ areas that are highly likely to experience rapid urbanization from 2009 to 2020.

2. Description of the MLDM

The objective of the MLDM is to consider not only the evolution of individual land cells, but also urban flows between the different cities/towns in an MCA. Landscape dynamics in an MCA can be regarded as a self-organizing process influenced by broad-scale environmental, geographical, and socioeconomic factors (Barredo...
The MLDM assumes that the evolution of the urban landscape in an MCA is shaped by the regional demand for urban lands and the potential conversion of a non-urban land cell to an urban cell is co-determined by local neighborhood factors (from a CA model perspective) and the inter-city urban flows within the MCA. The model therefore integrates a GFM with a CA model so that the influence of urban flows can be modeled by the former while that of local neighborhood factors can be modeled by the latter (Fig. 1).

The MLDM consists of three parts: (1) the estimation of demand for urban land in an MCA; (2) a GFM-based definition of the influence of inter-city urban flows on the urban landscape evolution; and (3) the prediction of future urban land cells by considering the urban flow influence defined in (2) in framing the transition rules of the CA model (Fig. 2).

2.1. Estimating the demand for urban land

At present, several non-spatial models are available for estimating the demand for urban land in an MCA during a given time period, including the linear programming model (Wang, Yu, & Huang, 2004), the regression analysis models (López, Bocco, & Mendoza, 2001; He et al., 2008), and the system dynamic models (He et al., 2006a, b). The regression analysis models are merited for the relatively simple calculations and fewer parameters required. López et al. (2001) suggested that the population residing in an urban area is highly correlated with its size. Therefore, historical data on the size of an urban area can be regressed on its population and the regression model can be further used to predict the future demand for urban land based on population growth data. Please refer to section 3.4 for further details.

2.2. Defining the influence of urban flows on landscape evolution

As discussed by Liang (2009) and Wang et al. (2011), an MCA can be regarded as a field of urban flow intensity, which decreases over distance from the city centers. We adopted this concept and used the GFM to quantify the influence of urban flows on the landscape evolution in an MCA. According to Liang (2009), the influence of urban flows on a non-urban cell \((x, y)\) at time \(t\) \((i_{Ix,y})\) can be expressed as follows:

\[
i_{Ix,y} = \frac{F_i}{D_{y,y,x,y}}
\]

where \(i_{Ix,y}\) is the intensity of the urban flow from city \(i\) at cell \((x, y)\); \(D_{y,y,x,y}\) is the Euclidean distance from the city center \((x_i, y)\) to the cell \((x, y)\); and \(F_i\) is the urban outflow from city \(i\). As suggested by Zhu and Yu (2002) and Yao et al. (2006), \(F_i\) can be calculated as follows:

\[
F_i = N_i \times E_i
\]
where \( N_i \) stands for the internal function (the ability of city \( i \) in an MCA to support itself), which can be fairly well represented by the city's GDP per employee. \( E_i \) denotes the external function (the ability of city \( i \) to support other cities in the MCA), which can be represented by totaling the external functions of all the economic sectors of city \( i \) as follows (Yao et al., 2006; Zhu & Yu, 2002):

\[
E_i = \sum_{j=1}^{m} E_{ij}
\]  

where \( E_{ij} \) is the output function of a specific economic sector \( j \) of city \( i \), and \( m \) is the number of sectors in city \( i \). \( E_{ij} \) is calculated as follows because it is effective and the data required to calculate for it are usually available:

\[
E_{ij} = G_{ij} - G_i \times \frac{G_j}{G}
\]  

where \( G_{ij} \) is the number of employees in economic sector \( j \) of city \( i \), \( G_i \) is the number of employees in all the economic sectors of city \( i \), \( G \) is the total number of employees of the MCA, and \( G \) is the total number of employees of the MCA. If \( E_{ij} \leq 0 \), the economic sector \( j \) of city \( i \) has no output function and \( E_{ij} \) is given a value of 0. If \( E_{ij} > 0 \), it means that the economic sector \( j \) of city \( i \) has an output function for the other cities in the MCA because the proportion of employees in the sector \( j \) of city \( i \) is larger than that of the entire MCA. In other words, the sector \( j \) is more centralized in city \( i \) than the other cities/towns in the MCA and therefore has an output function to support the other cities in the MCA. A larger \( E_{ij} \) value indicates a stronger output function for the economic sector \( j \) of city \( i \) (Yao et al., 2006).

When a land cell is influenced by the urban flows from cities of similar sizes, the maximum influence is accounted (Wang et al., 2011). Thus, for a non-urban cell \((x, y)\), the urban flow influence \(I_{k,x,y}\) from a number \(n\) of cities with a comparable size \(k\), can be expressed as:

\[
I_{k,x,y} = \max(I_{1,x,y}, I_{2,x,y}, \ldots, I_{k,x,y})
\]  

### 2.3. Calculating the probability of land change

According to Barredo et al. (2003), He et al. (2008) and White and Engelen (2000), the urban landscape evolution in an MCA can be understood as a complex process determined simultaneously by driving forces, resistance forces, and random perturbation. In the MLDM, the driving forces primarily include: (1) land suitability, (2) the contextual effect of a local neighborhood, and (3) the macro-level urban flows from the different cities in an MCA.

The probability \(P_{x,y}\) for cell \((x, y)\) to be converted into urban use at time \(t\) can be expressed as:

\[
P_{x,y} = f(D_{x,y}, R_{x,y}, V_{x,y})
\]  

where \(D_{x,y}\), \(R_{x,y}\), and \(V_{x,y}\) represent the driving forces, resistance forces, and random perturbation influencing the land change, respectively. \(D_{x,y}\) can be further expressed as:

\[
D_{x,y} = f(S_{x,y}, N_{x,y}, I_{x,y})
\]  

where \(S_{x,y}\) stands for the suitability, and \(N_{x,y}\) and \(I_{x,y}\) represent the neighborhood effect and the urban flow intensity, respectively. The resistance force \(R_{x,y}\) can be considered a function of two variables as follows:

\[
R_{x,y} = f(J_{x,y}, E_{x,y})
\]  

where \(J_{x,y}\) is the continuity attribute for cell \((x, y)\) to maintain its current land use or cover state \(z\) at time \(t\). \(E_{x,y}\) is the ecological constraint (e.g., located in an reservation area) that prevents a cell \((x, y)\) from being converted to urban use. The random perturbation term \(V_{x,y}\) can be constructed as:

\[
V_{x,y} = 1 + [-\ln(rand)]^a
\]  

where \(rand\) \((0 < rand < 1)\) is a random variable and \(a\) is an adjustable parameter representing the magnitude of the random perturbation. To keep the calculation manageable while incorporating urban flows into the transition rule of a CA model, the probability \(P_{x,y}\) for a non-urban cell to be converted into an urban cell can then be expressed as:

\[
P_{x,y} = \left( \sum_{i=1}^{m-n} W_i \times tS_{i,x,y} + W_{m-n+1} \times tN_{x,y} + \sum_{i=m-n+2}^{m-1} W_i \times tI_{i,x,y} - W_m \times tJ_{x,y} \right) \times \prod_{r=1}^{tEC_{x,y}} tV_{x,y}
\]  

where \(\sum_{i=1}^{m-n} W_i \times tS_{i,x,y}\) is the suitability for the conversion at time \(t\). \(tS_{i,x,y}\) is a standardized suitability score \([0, 100]\) based on the consideration of a number \((i = 1, \ldots, m - n)\) of factors. \(m\) is the total number of weights and \(m - n\) is the number of suitability factors. \(W_i\) denotes the weight associated with each factor. \(tN_{x,y}\) represents the neighborhood effect for cell \((x, y)\) at time \(t\) and \(W_{m-n+1}\) is its associated weight. The term \(\sum_{i=m-n+2}^{m-1} W_i \times tI_{i,x,y}\) quantifies the overall urban flow influence from all the cities/towns in an MCA, in which \(tI_{i,x,y}\) is the urban flow influence from the cities of a comparable size \(k\) and \(W_i\) is the corresponding weight. \(tJ_{x,y}\) is the continuity attribute of cell \((x, y)\) and \(W_m\) is its weight. These weights \((W_1, W_2, \ldots, W_m)\) are used to represent the relative importance of the factors that drive or resist urban landscape changes. \(\prod_{r=1}^{tEC_{x,y}} tV_{x,y}\) will have a value of 0 if cell \((x, y)\) is preserved due to constraint \(r\).

Neighborhood size is essential in running a CA model, as it should well define the extent of interactions between land uses and the scale of the land dynamics of an ecosystem (Caruso, Rounsevell, & Cojocaru, 2005). According to Barredo et al. (2003), He et al. (2008) and White and Engelen (2000), a neighborhood in the MLDM is a circular area around the cell of interest with a radius of five cells. This definition of neighborhood is believed to be appropriate for capturing the local urban cells’ agglomeration effect (Chen, Gong, He, Luo, & Tamural, 2002). The neighborhood effect \(N_{x,y}\) can be quantified as:

\[
N_{x,y} = A \times \sum_{c} U_c
\]  

where \(U_c\) indicates whether a cell at the distance \(c\) from cell \((x, y)\) is urban. \(U_c\) will have a value of 1 if it is urban and 0 if not. \(A\) is a scalar used to linearly standardize \(N_{x,y}\) into the range of \([0, 100]\).

Once the probability \(P_{x,y}\) for cell \((x, y)\) to be converted into urban use from a type \(z\) is determined, the evolution of the urban landscape can be simulated with respect to the total demand for urban land. To simplify the computation of the model, we can generally classify the various land use types into urban and non-urban, and only model the conversion processes from one generalized type to the other. Competition for space among different types of land use does not need to be considered because almost all the consumption of non-urban lands in our study area is for urban development. The simulation of the urban landscape dynamics in an MCA follows a repetitive computational workflow. \(P_{x,y}\) is calculated for each cell, and the cell with the highest \(P_{x,y}\) value is labeled as a future urban cell with high confidence. If the estimated total demand for
urban land in the simulated period cannot be satisfied by changing that cell into urban use, another iteration begins. The loop continues until the total area of urban land reaches the demand.

3. Applying the MLDM to simulate the urban landscape dynamics in the BTT-MCA, China

3.1. Study area and data

The study area of BTT-MCA is located in the North China Plain with a latitude range of 38° 28’ N to 41° 05’ N and a longitude range of 115° 25’ E to 119° 53’ E. It has a total spatial area of about 55,774.5 km² and accounts for 0.5% of the total area of China. The BTT-MCA consists of 43 cities and towns, including Beijing, Tianjin, Tangshan, Qinhuangdao, and Langfang (Fig. 3). Beijing is the capital of China and the political and cultural center of the country as well. Tianjin is the third largest city and the largest port city in North China. Beijing and Tianjin are under the direct jurisdiction of the central government, and are regarded as the ‘dual core’ of the BTT-MCA (Tan, Li, Xie, & Lu, 2005). In 2009, the urban population and the GDP of the BTT-MCA were 20.68 million and 2250 billion Chinese Yuan, accounting for 6.61% and 3.33% of the totals for Mainland China, respectively (NSBC, 2010). Since China implemented its open-door policy in the 1970s, the BTT-MCA has experienced tremendous urbanization (Kuang, 2011; Tan et al., 2005).

A series of land use/cover maps (1990, 2000, and 2009) for the BTT-MCA were used in this study. All of these maps were produced from the classification of Landsat TM/ETM+ images of good quality and have six land use/cover classes (urban land, cropland, grassland, forest, water, and others). Socioeconomic data (e.g., population, GDP, and number of employees in each economic sector) from the National Statistics Bureau of China (NSBC) were also collected and included in our modeling. A digital elevation model was obtained from the International Scientific Data Service Platform (http://datamirror.csdb.cn/dem/files/ys.jsp, accessed 02.02.12) to represent the regional topography, and a number of geographical information system (GIS) data layers were also collected from the National Administration of Surveying, Mapping, and Geoinformation of China, including administrative boundaries, rivers, road networks, city centers, airports, and coastal ports. All the GIS data were georegistered to the Albers Conical Equal Area projection and the raster datasets were resampled to have the same cell size of 300 m (1228 columns, 969 rows), which was sufficient to capture the detailed information about urban landscape.

Fig. 3. The study area.
of the GDP and urban population in the Beijing–Tianjin–Tangshan megalopolitan cluster area in 2009.}

3.2. Calculating the influence of urban flows between cities in the BTT-MCA, China

Based on their socioeconomic characteristics in 2009, the 43 cities/towns in the BTT-MCA were classified into four categories: (1) the primary core city of Beijing, (2) the secondary core city of Tianjin, (3) the major cities of Tangshan, Langfang, and Qinhuangdao, and (4) the other smaller cities and towns (Fig. 4). The influence of the urban flows from the cities in each category was calculated from the GIS data using Eq. (1). The influence values were subsequently standardized into the range of [0, 100] using the following formula and the resultant maps are displayed in Fig. 5.

\[
l_k = \frac{I_l - I_{\min}}{I_{\max} - I_{\min}} \times 100
\]

where \(l_k\) is the standardized influence of an urban flow influence, \(I_{\min}\) is the raw minimum influence value while \(I_{\max}\) is the raw maximum.

3.3. Model calibration and urban landscape simulation from 1990 to 2009

Calibration and validation are critical for the performance of CA models because it largely depends on the appropriateness of the transition rules, which typically involve several important parameters (Wu, 2002; Straatman, White, & Engelen, 2004). The calibration of the MLDM is usually conducted based on reliable historical data, which hopefully facilitates the effective simulation of future urban landscape dynamics (Wu, 2002). However, the calibration and validation of urban CA models are in fact rather challenging due to the complexity of urban landscape evolution processes and the numerous combinations of parameters involved (Verburg, De Nijs, Ritsema van Eck, Visser, & De Jong, 2004). Progress has been made for urban CA models with the development of two types of calibration methods (Li & Yeh, 2002). One of them is based on mathematical and/or statistical analysis, such as logistic regression (Wu, 2002) and artificial neural networks (Kuang, 2011). The other type are the trial and error approaches, including visual tests (Ward et al., 2000), landscape metrics comparisons (Berling-Wolf & Wu, 2004b), and computer simulation comparisons (He et al., 2008). As suggested by Chen et al. (2002) and He et al. (2008), an adaptive Monte Carlo approach was chosen to automatically calibrate the parameters in the MLDM. This approach is mainly driven by data and avoids subjective determination of the weights used in the transition rules of the CA model. Specifically, a Monte Carlo sampling process was used to search for the optimal combinations of weights that resulted in the most accurate simulation results. Since all the weights in a transition rule should total to 1, the calibration process can be expressed as follows:

\[
\text{Objective function} \quad \max A(w_1, w_2, \cdots, w_m)
\]

where \(w_k\) is the weight for factor \(k\) and \(A\) is a measure of agreement between the simulation result and the reality. We chose the kappa index for \(A\) because it is widely recognized as an effective measure of accuracy for remote sensing classifications. As suggested by Chen et al. (2002) and He et al. (2008), up to 500 iterations were conducted to achieve the best performance. Although the simulation was computationally intensive, it was manageable with the strong computing power of our workstations.

The 2000 and 2009 urban land use maps were used to calibrate the MLDM by treating the 1990 map as the starting point. Thirteen of the most important impact factors were incorporated in the MLDM, including the influence of the urban flows from all the four categories of cities, slope, distances to highways, railways, expressways, coastlines, airports, harbors, and city centers, neighborhood effect, and continuity attribute. Since it is understandable that the influence of urban flows was dynamic, and therefore it was modeled for 1990, 2000 and 2009 in order to simulate the urban landscape change during 1990–2000, 2000–2009 and 2009–2020. Because these factors were measured in different units, they were all scaled into the range of [0, 100]. Urban landscape dynamics were simulated for the periods of 1990–2000 and 2000–2009, with 500 different sets of weights for the 13 impact factors. The simulated urban landscapes for 2000 and 2009 based on each set of weights were compared with the actual urban landscapes in these two years. The most accurate simulation results achieved a kappa index value of 0.81 and 0.82 for 2000 and 2009, respectively.

The set of weights that produced the most accurate simulation results (Fig. 6) are displayed in Table 1. As suggested by Sany et al. (2010), the single use of Kappa index is not enough to assess the urban spatial patterns generated by CA models. So we performed the further accuracy assessment of the simulated results using Moran I index by referring to the relative works of Wu (2002) and Liu, Li, Liu, He, & Ai (2008b). It was found the Moran I values of the simulated urban pattern in 2000 and 2009 are 0.93 and 0.96, respectively, close to the actual urban pattern’s 0.97 and 0.99. Overall, the simulation results by the MLDM matched the actual urban landscapes quite well, although some small-scale details were not effectively captured. To assess their importance to the model, the influence of the urban flows was excluded from consideration in a test simulation for both periods. The highest kappa index values for the two year were both below 0.8, 0.75 and 0.78 for 2000 and 2009, respectively. Meanwhile, their Moran I values are 0.89 in 2000 and 0.93 in 2009, also more different to the actual’s 0.97 and 0.99 than the simulated results with urban flow. The lower
Fig. 5. Influence of urban flows (from the different cities) on urban landscape dynamics within the Beijing–Tianjin–Tangshan megalopolitan cluster area, China, 2009.

Fig. 6. Model calibration and urban landscape simulation in the Beijing–Tianjin–Tangshan megalopolitan cluster area, China (1990–2009).
accuracy indicates the significance of accounting for urban flows in the simulation. The simulations that do not account for urban flows tend to overestimate the roles of neighborhood effect and the transportation networks in driving the urban landscape evolution in an MCA. For example, the MLDM accounting for urban flows suggested that, in 2000, newly converted urban lands were distributed more around Beijing than around the other pre-existing urban areas [Fig. 6(a2)]. This is a much more accurate representation of the actual situation than what was produced by the CA model that did not consider urban flows which showed a relatively uniform distribution of newly converted urban lands around all the pre-existing urban areas [Fig. 6(a1)]. The MLDM accounting for urban flow influence also correctly suggested that the newly converted urban lands were mainly located around Tianjin and the northeast of Beijing in 2009 [Fig. 6(b2)].

3.4. Prediction of urban landscape dynamics from 2009 to 2020

The ability to predict the future urban landscape in the BTT-MCA is very important. In order to predict the future landscape in the study area using the proposed MLDM, we first estimated the future demand for urban land based on population growth data (1990–2009) for the BTT-MCA (Fig. 7). A linear regression model (Eq. (15)) was built to predict urban population on year:

\[ y = 39.17x + 1257.27 \]  

(15)

where \( y \) stands for the urban population and \( x \) stands for a year relative to 1990 (e.g., \( x = 6 \) for 1995). The regression was found to have a very strong predictive power, with an \( r^2 \) of 0.97.

Subsequently, another regression model (Eq. (16)) was developed based on the historical data (1990–2009) on the urban population and the corresponding total area of urban land in the BTT-MCA (Fig. 8):

\[ y = 3.27x - 3171.05 \]  

(16)

where \( y \) stands for the total area of urban land and \( x \) stands for the total urban population. This regression model was also found to have a rather strong predictive power (\( r^2 = 0.99 \)).

The historical and predicted data of urban population, GDP, and urban land area for the BTT-MCA are summarized in Table 2.

Future urban landscapes in the BTT-MCA were simulated for 2009–2020 using the calibrated MLDM and the estimated future demand for urban land. Fig. 9 displays the simulated urban landscapes for 2015 and 2020, which clearly suggest a continuing urbanization in the region. Fig. 10 highlights the areas in the BTT-MCA that are likely to be converted into urban use before 2020, mainly including the northeast of Beijing, the east of Tianjin, and the outskirts of Langfang. Such foreseeable rapid urbanization in the BTT-MCA may create or intensify environmental and ecological problems, such as air pollution, biodiversity reduction, and local/regional climate change (Chen et al., 2002, 2011). Therefore, the highlighted areas deserve closer attention for a more effective management of this heartland of China.
Table 2
Urban population, gross domestic product (GDP), and urban land area in the Beijing–Tianjin–Tangshan megalopolitan cluster area, China (1990–2020).

<table>
<thead>
<tr>
<th>Year</th>
<th>Urban population (10,000s)</th>
<th>GDP (billions of Yuan)</th>
<th>Urban land area (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>1356.83</td>
<td>246.62</td>
<td>1296.09</td>
</tr>
<tr>
<td>2000</td>
<td>1626.1</td>
<td>748.21</td>
<td>2094.75</td>
</tr>
<tr>
<td>2009</td>
<td>2068.32</td>
<td>2250.45</td>
<td>3609.90</td>
</tr>
<tr>
<td>2015</td>
<td>2275.82</td>
<td>3857.78</td>
<td>4270.88</td>
</tr>
<tr>
<td>2020</td>
<td>2471.69</td>
<td>6686.50</td>
<td>4911.38</td>
</tr>
</tbody>
</table>

Fig. 9. Simulated urban landscape dynamics in the Beijing–Tianjin–Tangshan megalopolitan cluster area, China (1990–2020).
4. Conclusion and discussion

The urban landscape dynamics in an MCA are driven by many factors on different scales. Socioeconomic interactions in form of urban flows between the cities and towns in an MCA play an important role in shaping its landscape evolution. Present CA models focusing on a single city often fail to account for the influence of urban flows and are therefore inadequate for accurately simulating the urban landscapes in MCAs. This paper proposed an MLDM that enhances a CA model by accounting the influence of urban flows with the aid of a GFM.

After calibrated, the MLDM was able to simulate the urban landscape dynamics in the BTT-MCA, China, more accurately than the CA model that does not consider urban flows. The kappa index values for the simulated urban landscapes increased from 0.75 to 0.81 in 2000 and 0.78 to 0.82 in 2009, respectively. The simulation results suggested a continuing urban growth in the BTT-MCA and showed that the future urbanization is most likely to take place in the northeast of Beijing, the east of Tianjin, and the outskirts of Langfang. Close attention should be paid to these areas to effectively detect, study and solve the problems associated with such urban expansions in hope for a sustainable development in the region.

The highly accurate simulation of urban landscape dynamics is rather challenging due to the complexity of an MCA. The proposed MLDM certainly needs to be further tested and improved in a number of ways. First of all, further research should be done to determine the sensitivity of the simulation result to the number of iterations in the model calibration. Second, it should be pointed out that the linear transition rules have limitations in capturing the complexity of urban landscape dynamics. More advanced techniques have been adopted to define complex transition rules in the recent years including support vector machine (Yang, Li, & Shi, 2008), kernel-based non-linear technique (Liu, Li, Shi, Wu, & Liu, 2008a), ant intelligence (Liu et al., 2008b), and ant colony optimization (Li, Lao, Liu, & Chen, 2011). These techniques may be used to upgrade the MLDM in our future research in hope to achieve an even higher accuracy. Third, the MLDM currently does not yet consider the behaviors of different stakeholders (e.g., government, enterprises, and individuals) for their influence on the landscape evolution of an MCA, but has the potential to incorporate them by drawing ideas from the newly developed agent-based urban landscape simulation models (Fontaine & Rounsevell, 2009; Li & Liu, 2007; Tian, Ouyang, Quan, & Wu, 2011b; Valbuena et al., 2010). Fourth, although the proposed MLDM indirectly considers the influence of macro-factors by incorporating the socio-economic data and urban planning information (e.g., the distribution of reservation zones), it remains to be a research challenge to fully consider the macro level socioeconomic and political factors due to their complexity and difficulty to be quantified. Lastly, it is a limitation of the present MLDM that more detailed land uses/cover are not differentiated in the urban flow calculation or suitability evaluation. Future research efforts will be spent to upgrade the model by allowing the consideration of more detailed land uses/cover (White & Engelen, 2000).

Nevertheless, the MLDM is of great value to effective urban/regional planning and environmental management (He, Tian, Shi, & Hu, 2011). This study has shown that accounting the influence of urban flows improves the MLDM and makes it more effective; the kappa index values for the simulation results increased from 0.75 to 0.81 in 2000 and 0.78 to 0.82 in 2009, respectively. It is clear that urban flows have significant impact on the landscape evolution and regional development in an MCA. So the interaction and synergic development between cities/towns should be adequately considered in urban and regional planning. Moreover, highways, expressways, urban flows from core cities, and neighborhood effects were found to be relatively stronger forces in shaping in the landscape of an MCA. We can possibly control the evolution of the urban landscape by, for example, strategically placing the new transportation infrastructure, and regulating the growth and clustering of the towns surrounding the major cities of...
Beijing and Tianjin. The simulation result also suggested that the ‘hotspots’ of future urbanization will be located in northeast of Beijing, east of Tianjin, and in the outskirts of Langfang. Agricultural lands and water bodies will most likely be converted into urban use and therefore cause ecological and environmental problems. Our findings can greatly help the urban planners and policy makers by allowing them to foresee the future landscape with a fair accuracy and to take targeted measures proactively.

Acknowledgments

This research was supported by the Natural Science Foundation of China (Grant Nos. 412222003 & 40971059), the National Basic Research Program of China (Grant No. 2010CB950901), and the National High-Tech Research Program of China (Grant No. 2009AA122004). We would like to express our respects and gratitude to the anonymous reviewers and editors for their valuable comments and suggestions on improving the quality of the paper.

References


