

# Chapter 12

## Modeling Urban Land Use Change: Integrating Remote Sensing with Socioeconomic Data

Junmei Tang

**Abstract** Rapid urban development has stimulated the progress in predicting and evaluating urban landscape evolution. As a result of rapid socioeconomic development, the land use pattern of Houston, TX, has undergone significant changes over the past 30 years. It is essential to simulate urbanization processes in Houston to examine where and to what extent landscape change has occurred and further to understand how and why the change can occur. This research developed two cellular automata (CA) models based on the same remote sensing data source: one was based on the classification from Landsat images and another one incorporated the socioeconomic data with the same classification results. The predicted results from these two models suggested that the incorporation of socioeconomic data improved the accuracy in human-intervened landscapes, such as residential and industrial/commercial area. More socioeconomic data and finer data sources were needed to improve the CA model to predict the heterogeneous pattern within urban areas.

**Keywords** Urban land use change • CA model • Socioeconomic data • Remote sensing

### 12.1 Introduction

Rapid urbanization in the past 50 years, triggered by the population growth and migration from rural to urban and suburban areas, presents one of the greatest challenges in environmental, economic, social, political, and cultural research (Antrop 2004; Tang et al. 2012; Tayyebi et al. 2012). The total urban population is 82 % with an estimated 1.2 % annual increasing rate from 2010 to 2015 in the United States (US Census 2011). The motivation to model urban landscape dynamics arises from the process of examining where and to what extent landscape change has occurred, and furthermore, the need to understand how and why the changes can

---

J. Tang (✉)

Geography and Environmental Systems Department, University of Maryland,  
Baltimore County, Baltimore, MD 21250, USA

e-mail: [junmei@umbc.edu](mailto:junmei@umbc.edu)

occur (Weng 2002; Yang and Lo 2002). One of the greatest challenges in designing effective urban models is that their performances are often limited by the inadequate digital data source over time as well as the consideration of external driver such as socioeconomic development and human disturbance (Pickett et al. 1997; McIntyre et al. 2000).

Remote sensing data, with the ability to provide large-scale data sources such as historical maps or urban land use maps, has been used as an effective tool in quantitatively measuring urban landscape and modeling urbanization at a relatively large spatial scale (Herold et al. 2003; Tang 2011). Images from satellite sensors provide a large amount of cost-effective multispectral and multi-temporal data to monitor landscape changes and estimate biophysical characteristics of land surfaces (Weng 2002). Many researchers have proposed the routine to combine remote sensing with GIS in urban growth models (Tang 2011; Tayyebi et al. 2013). Significant progress in acquiring remotely sensed data in a higher spatial resolution and developing the spatial geographic process model has widened our research on the process, driving forces, and impacts of the urbanization.

The cellular automata (CA) model, introduced by Tobler in 1979, is one of the most powerful spatial dynamics techniques used to simulate complex urban systems (Batty and Xie 1994). The CA model allows researchers to view the city as a self-organizing system in which the basic land parcels are developed into various land use types. Cecchini and Viola (1990) applied simple decision rules in the CA model to predict the complex, large-scale structure in the urban growth process. Wu (1998) combined the multicriteria evaluation (MCE) and GIS into the CA model to define the transition rules in a visualized environment. Shafizadeh-Moghadam and Helbich (2013) used AHP (analytical hierarchy process) to determine the weight in a Markov chains-cellular automata urban growth model.

The advantages of the CA model in simulating urban spatial process and dynamics (Hillier and Hanson 1984; White and Engelen 1993) have been widely documented because the theoretical abstraction of the CA model and the practical constraints in the real world can be easily related (Batty and Xie 1994; Clarke and Hoppen 1997; Wu and Martin 2002). The model begins from a homogeneous cell-based grid and adjusts itself through the transition rule derived from its local spatiotemporal neighborhood. This makes the CA model suitable to simulate complex and hierarchical structures since more unknown, immeasurable spatiotemporal variables can be incorporated and manipulated in this model. Another advantage in CA simulation is the ability of the model to incorporate proper parameters or weights to model the alternative socioeconomic states in the model development (Clarke and Gaydos 1998; Li and Yeh 2000). With better computer techniques, the CA model is also able to explore more complex human behavior through defining different transition rules (Li and Yeh 2000; Wu and Martin 2002). However, the tension between the simple local transition rule in CA models and the complex, unpredicted social changes in urban landscapes still remains.

In this context, this chapter attempted to develop a spatial-explicitly CA model to simulate urban growth patterns using the classification result from Landsat images and another one incorporated the socioeconomic data with the same classification

results. Two CA models were compared to test how the socioeconomic data could improve the urban model simulation in Houston during the last 30 years. Specially, the following research questions were addressed: How the socioeconomic data could be incorporated with remote sensing in the urban growth model? Does the socioeconomic data improve the model? In which classes does this model improve?

## 12.2 Urban Model Review and Socioeconomic Data in the Model

With the availability of spatial data on a large scale, various sophisticated models, especially after the late 1990s, were developed such as UrbanSim model (Waddell 2002), Markov chain model (Stewart 1994), LUCAS model (Berry et al. 1996), CLUE model (De Kong et al. 1999), area-based model (Lichtenberg 1985; Tayyebi et al. 2011, 2013); CA model (Batty and Xie 1994), Land Transformation Model (Pijanowski et al. 1997, 2014), and agent-based model (Liebrand et al. 1998). The detailed review of these spatial explicit models is listed in Table 12.1.

In terms of the methods to represent the model object, there are vector-based models and grid-based models (Herold 2004), and both of them have been used to incorporate socioeconomic data. Vector-based models use the thematic map as the input data for the model, and the spatial objects are usually defined as homogeneous land units. *UrbanSim* is one of land use simulation models for the growth government, regional land use, and transportation planning in the states of Hawaii, Oregon, and Utah (Waddell 2002). Within the context of urban infrastructure and governmental policy, *UrbanSim* represents zonal structure in the urban area to monitor the socioeconomic-related behaviors of households, business, and land developers. Theoretically, *UrbanSim* is an object-oriented model. *What if* model (Klosterman 1999) begins with uniform analysis zones or homogeneous land units generated from the GIS software. Through applying the governmental policies and land use demands, this model derives the aggregating value of the regional condition on the land units. *What if* model projects future land use patterns by balancing the supply, demand, and land sustainable at different locations. *Area-based* model is a vector-based model used in resource assessments to predict the availability of farm and forest land. Transformed from the regional model (Palmquist 1989), *area-based* model allocates the proportions of a given land use to predefine land use categories using Lichtenberg's (1985) acreage allocation method (Tayyebi et al. 2011, 2013).

Another vector-based model is *Markov* model which predicts future landscape patterns based on the spatial transition probability. Although Markov model is a typical spatial transition model, early Markovian analysis is a descriptive tool to predict land use change on a local or regional scale (Bell 1974; Bourne 1976; Arsanjani et al. 2013). Actually, the Markov model is not a strict vector-based model; it is based on the statistical results from the thematic map. Lopez et al. (2001) used Markov chain to simulate the relationships among a set of urban and

**Table 12.1** Detailed comparison of urban models

	Model name	Developer	Purpose	Variables	Strengths	Weakness
Vector based	UrbanSim	Waddell (2002)	Predict land use, physical development, the movement, and location of businesses and households	Socioeconomic, environmental parameters	Simulate the interaction between urban activities and the natural environment	Deterministic model
	What if	Klosterman (1999)	Determine what will happen in land use patterns in the future if policy choices are made	Natural conditions, urban infrastructure	Fully operational model in adapting particular data sets and policy concerns	Lack of a firm theoretical basis
	Area based	Lichtenberg (1985) and Tayyebi et al. (2011, 2013)	Project the proportions of land use using hedonic rent theory and acreage allocation model	Based on Palmquist's hedonic rent theory and acreage allocation model (Lichtenberg 1985)	Easy to incorporate available socioeconomic data, such as age, income, population, and rent	Long-term prediction is not good
	Markov	Bell (1974) and Tang et al. (2007)	Characterize the land use/land cover change	Multi-temporal change, transition probabilities	Mathematically compact, easy to implement	Does not account for spatial context

Grid based	CA	Batty and Xie (1994) and Tang (2011)	Model the spatial structure of urban land use over time	Landscape maps, environmental factors in neighborhood	The spatial factor and ecological aspects are easily to be incorporated	Face challenge in incorporating human decision making
	LUCAS	Berry et al. (1996)	Simulate the landscape change with socioeconomic information and its environmental impacts	Socioeconomic variables, such as transportation network, ownership, population density, etc.	Flexible and interactive computing environment	The patch size is no sense due to the pixel-based method
	LTM	Pijanowski et al. (1997, 2014)	Analyze the land use change and predict land use pattern	Landscape maps, social, political, and environmental factors	Can be applied to multiple scales using a moving scalable window metric	Suppose all the variables constant
	CLUE	De Kong et al. (1999)	Predict future land use	Biophysical drivers and human drivers	Covers a wide variety of biophysical and human factors and ranges multiple spatiotemporal scales	No social or political factors
Grid/vector	Agent based	Liebrand et al. (1998)	Represent a wide variety of entities and its activities	A simulated environment, entities under human decision making	Flexibility and successful in replicating human decision	Difficult to develop and control

social variables in predicting land use/cover change in the urban fringe of Morelia city, Mexico. Weng (2002) demonstrated that the integration of satellite remote sensing and GIS techniques into the stochastic urban modeling was an effective approach for analyzing the direction, rate, and spatial pattern of landscape change in Zhujiang Delta of China. Tang et al. (2007) improved the Markov chain model by incorporating a modified genetic algorithm in the urban boundary expansion for urban simulation. Mathematically, most vector-based models rely on some static equations, and this characteristic provides the potential in integrating the statistical information into the model entities. The major drawbacks of such models are the poor handling in dynamic entities and poor representation of external variables, e.g., the spatial information and socioeconomic factors.

The models developed on grid have more advantage in solving these problems than the vector ones. Land-Use Change Analysis System (*LUCAS*) is a grid-based model which integrates socioeconomic and ecological variables in the multilayered, gridded maps (Berry et al. 1996). This model consists of three subject modules: socioeconomics, which derives the transition probability from the function of socioeconomic driving variables; landscape change, which predicts the landscape maps from the socioeconomic module; and environmental impacts, which estimates the impacts of selected environmental variables from the landscape maps from second modules. Land Transformation Model (*LTM*) (Pijanowski et al. 1997, 2014) applied the spatial rules to land use transitions for each location in the processed spatial layer or grid. It is easy to quantify the contribution of different spatial variables because of its grid format. In order to aggregate the land use change and change drivers, this model adopted the similar method with the Conversion of Land Use and its Effects (*CLUE*) model (De Kong et al. 1999). Both of them apply the variable values in grid format to create a series of future land use patterns over the time. *Cellular automata* model has been proposed and developed to simulate the urban land use model by incorporating various socioeconomic variables, such as dynamic transportation model (Aljoufie et al. 2013) and dynamic population density (Van Vliet et al. 2012).

*Agent-based* model (Liebrand et al. 1998) is a complex behavior model which used both vector data and raster data. Usually, the raster data is the agents' environment, and the agents, in turn, act on the simulated environment. This model can be applied to a wide variety of simulations, including moving cars, animals, people, or even organizations. The socioeconomic variable, as both agents' status and driving forces, was incorporated into the model to simulate individual activities (An et al. 2005). This model is difficult to develop and control since we need to incorporate the "individual agent" information and predict its potential behaviors.

Generally, a reliable urban growth model should have the following capabilities: (1) providing an appropriate theoretical and technical framework for urban growth; (2) understanding and describing the historical dynamics of urban structures; and (3) exploring and incorporating different economic and social parameters to monitor the urban growth.

### 12.3 Study Area and Data Preparation

The eastern metropolitan area of Houston, Texas, covering an area of 1,200 km<sup>2</sup>, was chosen as the study site (Fig. 12.1). Houston is situated in the northern portion of the Gulf coastal plain, a 60 by 80 km-wide swath along the Texas Gulf Coast, 80 km from the Gulf of Mexico (Moser 1998). This area has experienced rapid urban development since the 1930s after the discovery of oil (Tang et al. 2008) in nearby oil fields. These discoveries made it the largest city in Texas as of 1930 and the fourth largest city in United States since 1990 (Texas State Historical Association 2002). Although the government tried to diversify its economy (Key to the city 2001), the city's unchallenged role as an international center of oil technology, headquarters for a number of the world's largest energy companies, and a strong refining and petrochemical manufacturing base should shore up the local economy

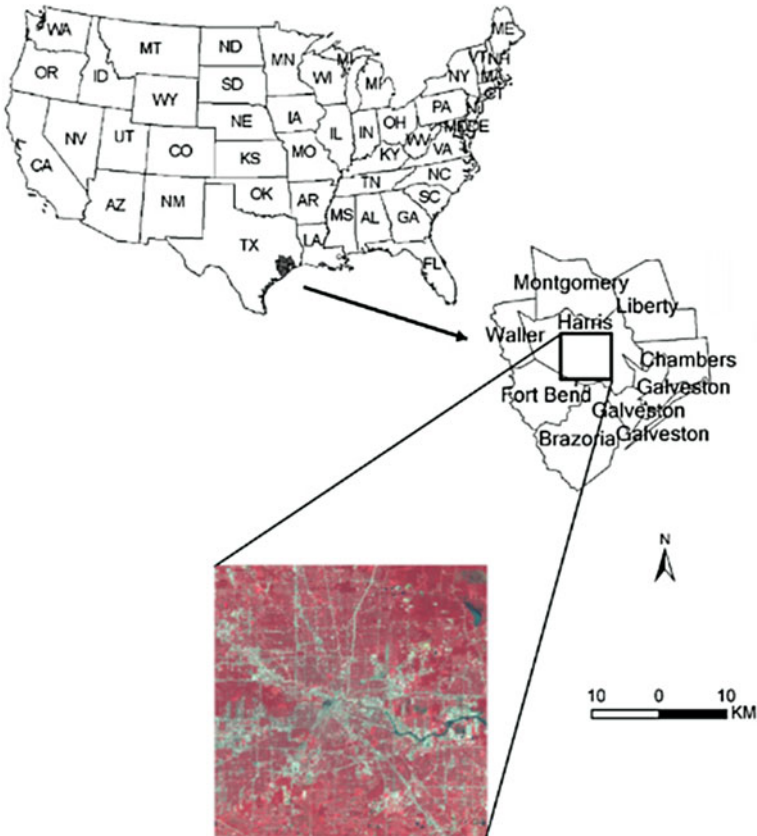
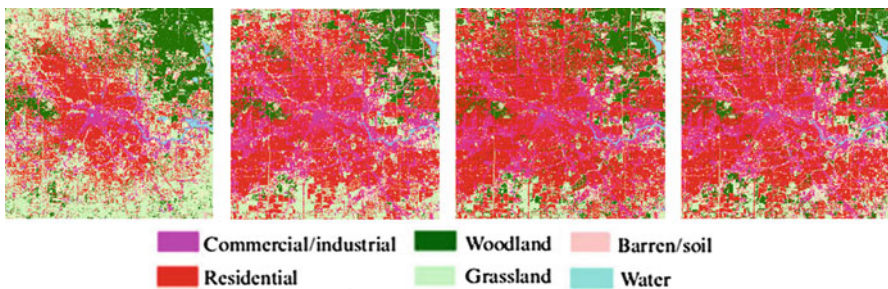


Fig. 12.1 Study site, Houston, Texas, in the United States

of Houston in the near future. The representative land use/land cover classes in this selected region include residential area, commercial/industrial area, transportation, woodland, grassland, and barren/soil.

The satellite Landsat MSS/TM images were collected from 1970s to 2010 in this study. All these images were georeferenced to the Universal Transverse Mercator projection using ENVI. The convention Maximum Likelihood Classification was adopted to obtain four classified landscape maps with six landscape classes for each map. We chose two set of samples around 600 pixels for training samples and test samples, respectively. The selection of separate of training and test samples was guided by the characteristic of each class at different years. The overall accuracy assessment of classified maps was 92 % (1979), 94 % (1990), 96 % (2000), and 95 % (2010). Figure 12.2 shows the detailed proportion of each land use type as shown in Table 12.2.

In order to represent the rapid socioeconomic development in the Houston area, four major socioeconomic variables were collected: population density, house density, road density, and distance to highways (Van Vliet et al. 2012; Aljoufie et al. 2013). These four variables were collected at census block level from the official website of the U.S. Census Bureau (US Census Bureau 2010).



**Fig. 12.2** Satellite images and classification results from MLC method on October 1979, December 1990, November 2000, and October 2010

**Table 12.2** The proportion of each land use type from 1979 to 2010 in Houston

	Area (km <sup>2</sup> )			
	1979	1990	2000	2010
Houston				
Residential	312.47	479.78	564.43	572.43
Industrial/commercial	93.50	228.58	198.21	197.07
Grassland	511.34	287.53	235.89	178.63
Woodland	209.68	140.32	184.70	179.43
Barren/soil	70.13	81.98	32.72	87.13
Water	31.73	9.67	11.85	13.11



## 12.4 Methodology

A cellular automata model was developed to investigate the scenarios of future urban land transformations in Houston. This model started on a 30-m grid and the transition rules were applied to all cells at the same time, and the entire grid was updated at the annual iteration. The transition rules were defined as the difference between the center cell and eight neighbors within  $3 \times 3$  Moore's neighborhood. To determine the state of a cell in a certain time period, the simulation function was written as:

$$S_{i,j}^{t+1} = a_N \times N_{i,j}^t + a_M \times M_{i,j} + a_{SE} \times SE_{i,j} \quad (12.1)$$

where  $N_{i,j}^t$  denotes the diffusion factor regarding its neighborhoods,  $M_{i,j}$  denotes the Markov transition probabilities,  $SE_{i,j}$  denotes the socioeconomic status of each single cell and its neighborhoods;  $a$  represents the coefficients for these variables.

For a self-organizing CA model, the diffusion factor, Markov transition rules, and socioeconomic status were defined as:

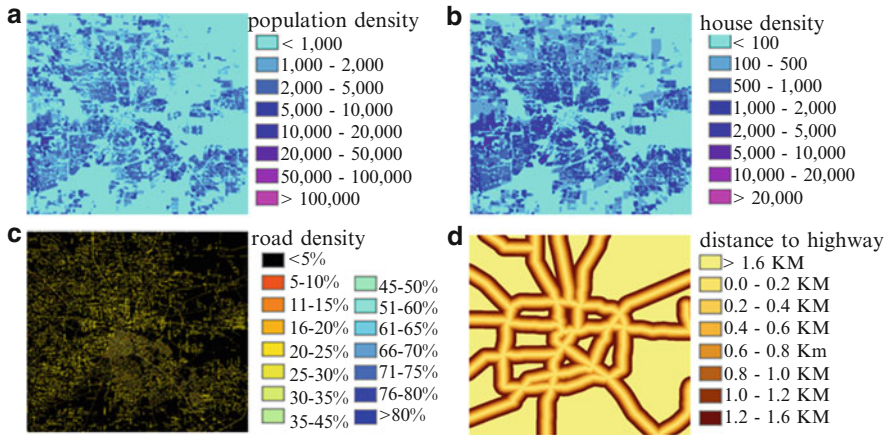
$$N_{i,j} = \frac{n_{i,j}}{\sum n_{i,j}} \quad (12.2)$$

$$M_{i,j} = \sum_{k=1}^k \frac{N(i,j) / \sum_{i=1}^m m}{k * m_k} \quad (12.3)$$

$$SE_{i,j} = \frac{\sum_{n=1}^n \left( \frac{d_{i,j}^n - \min(d_{i,j}^n)}{\max(d_{i,j}^n) - \min(d_{i,j}^n)} \right)}{n} \quad (12.4)$$

where  $n_{i,j}$  is the total number of class  $i$  surrounding the observed class  $j$ ,  $N(i,j)$  is the observed landscape amount changing from class  $i$  to class  $j$  during total  $m$  years at  $k$  internal steps, and  $d_{i,j}^n$  is the different value in the selected four socioeconomic variables between the observed center cell and its  $n$  neighbors (Fig. 12.3).

Although the socioeconomic data were collected at the last year of simulation, the difference of socioeconomic values between the observed cell and the neighbors was used to determine the socioeconomic factors. Obviously, different socioeconomic variables have different impact weights to the urban land use/land cover change. In order to find the weightiness of each socioeconomic variable, 20 experts in the field of socioeconomic and land use change were invited to assign weights to each variable using the index ranging from 0 to 10 to represent the weight from the highest impact to the lowest impact. The average value of these ratings was shown in Table 12.3.



**Fig. 12.3** The visualization of the socioeconomic value in Houston (a) Population density; (b) House density; (c) Road density; and (d) Distance to highway

**Table 12.3** The weight of socioeconomic indices

Houston	Population density	Road density	Distance to highway	House density
Barren/soil	3.67	3.50	3.40	4.00
Industrial/commercial	8.80	8.17	8.42	7.33
Grassland	3.18	2.75	3.36	4.90
Residential	9.46	8.08	6.25	8.92
Transportation	7.66	9.00	8.95	6.82
Woodland	2.82	2.58	3.18	4.70

A critical issue in the CA model is the provision of proper methods to calibrate the CA model to find appropriate coefficients for the diffusion factor, Markov transition rules, and socioeconomic status (Hagen-Zanker and Lajoie 2008; Van Vliet et al. 2011). To calibrate the model, we used the classified Landsat TM image as empirical maps on the following dates: November 5, 1984; July 20, 1990; October 6, 1999; and November 9, 2000. We randomly selected an encoded weight number (ranging from 1 to 10) for each factors, run the CA model using these weight number, and compared the cells simulated in the CA model with the cells located in the empirical maps to choose the weight number with the highest fitness. The CA model was run at yearly intervals to represent one combination until the next calibration year. These steps were repeated until the year of the last calibration map.

For the validation, the model’s simulation output was compared to the empirical map, occurring in the same simulated year (Pontius et al. 2004; Pontius and Cheuk 2006) through visual inspection and quantitative evaluation. In this research, we adopted the classified map in October 31, 2011, as an empirical map and overlaid it with the predicted map to generate a black-and-white error image. Meanwhile, an error matrix was built up with the user’s and producer’s accuracy for each class as well as the overall accuracy and Kappa for the entire landscape.

### 12.5 Results and Discussion

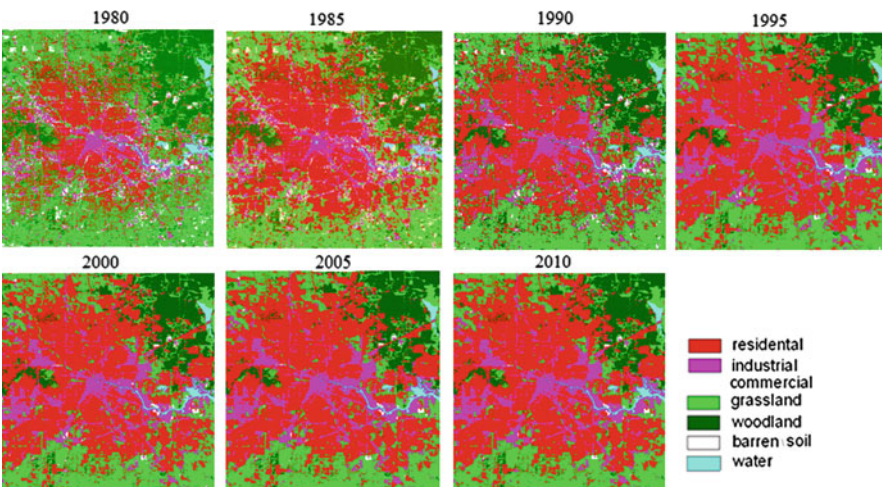
Since our model was based on actual observation from the last 30 years in Houston, the temporal transition probability matrix is calculated by accumulating the periods from 1979 to 2010. We first calculated the yearly transition matrix between each two subsequent maps between 1979–1990, 1990–2000, and 2000–2010 and then calculated the yearly transition matrix between 1979 and 2010 using Eq. 12.3. The yearly transition probability matrix from 1979 to 2010 is shown in Table 12.4.

Using the yearly transition probability matrices in Table 12.4, we parameterized the Markov transition probability and socioeconomic variable on the census block level into the CA model. Two CA models were built up, one with the socioeconomic variables and another one without. Figures 12.4 and 12.5 show the initial state and simulated pattern of Houston with the socioeconomic variables and without the socioeconomic variables, respectively.

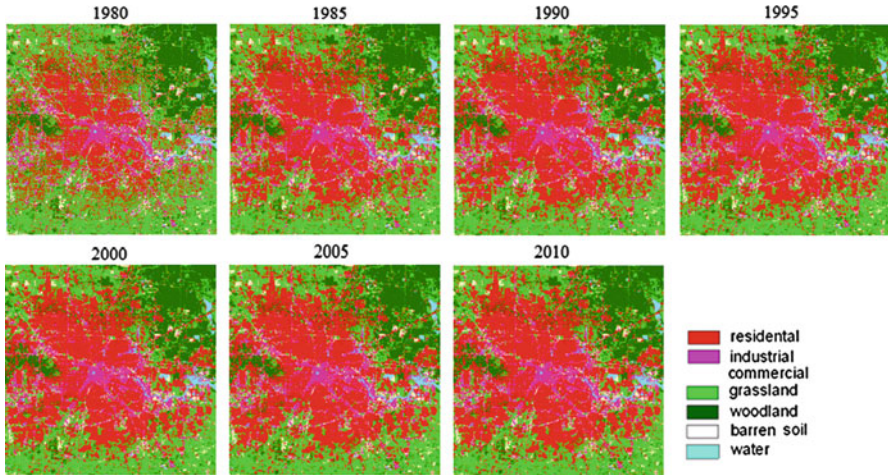
The simulated results from two models have similar pattern in general urban sprawl pattern: fast shrinkage in grassland and woodland and clear outward expansion in residential or industrial/commercial area. This growth pattern could

**Table 12.4** Yearly transition probability (%) matrix from 1979 to 2010

Houston (1979–2010)	Residential	Industrial/commercial	Grassland	Woodland	Barren/soil
Residential	98.07	0.82	0.56	0.42	0.12
Industrial/commercial	1.24	98.01	0.23	0.03	0.4
Grassland	1.85	0.35	97.13	0.53	0.12
Woodland	0.09	0.18	1.19	98.25	0.26
Barren/soil	2.26	2.27	1.26	0.19	94.02



**Fig. 12.4** The simulated landscape pattern of Houston with the socioeconomic factors



**Fig. 12.5** The simulated landscape pattern of Houston without the socioeconomic factors

be observed in the southeastern and northeastern city with a large amount of new residential and industrial/commercial area being built in the last 30 years. Different from other large cities in the United States, Houston did not adopt city zoning laws in its urban planning. Lacking city zoning has led to an abundance of urban sprawl in Houston, resulting in a relatively large metropolitan area and low population density. Land developers inspired the spread of Houston when they built suburbs such as Pasadena (1892), Houston Heights (1892), Deer Park (1892), Bellaire (1911), West University Place (1919), and River Oaks (1922–24).

Although the simulated results from two models have the similar sprawl pattern, the model with the parameterized socioeconomic variables had a better correspondence with the “abrupt” expansion in residential and industrial/commercial area. From Fig. 12.6, we could find that the “abrupt” expansion were simulated well in the model with the socioeconomic data as the larger predicted area in these human-related landscapes in the year 2010. This “abrupt” expansion was caused by the rapid economic development, population growth, and road construction in Houston. The simulated pattern by the model without socioeconomic factors was much tardier, especially in simulating the rapid growth in suburban area. The differences between these two models indicate that the CA spatial model could simulate the urban evolution behaviors with incorporating enough driving factors.

In order to display the error in the predicted map, we compared our predicted results with the empirical maps. The differential map was shown in Fig. 12.7. White pixels in the figure represented the area predicted correctly, while dark pixels represented the incorrect prediction. Generally, the residential areas were best predicted and most of the errors were found in the suburban area, which were mostly grassland and barren/soil landscapes. The woodland was predicted better than other natural landscapes, which might be caused by the large forest reserved area in northeastern Houston in the Sheldon Lake State Park and Dwight D. Eisenhower Park.

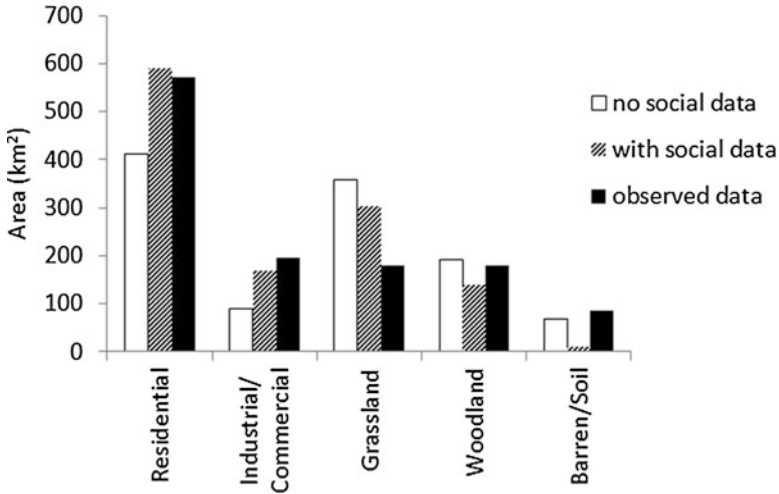


Fig. 12.6 The estimated results from two models in 2010

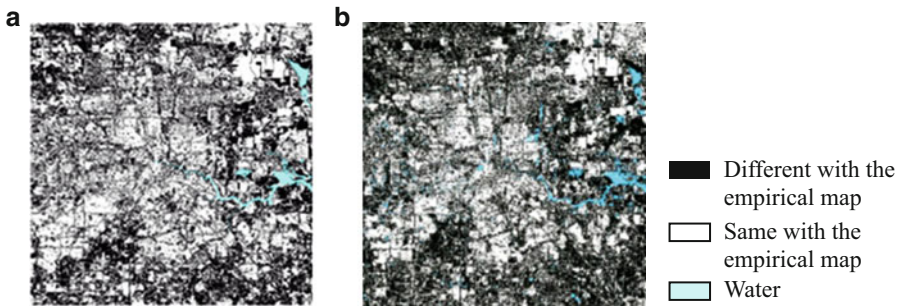


Fig. 12.7 The differential map between the predicted map and empirical map (a) with socioeconomic factors and (b) without socioeconomic factors

In Fig. 12.7, the predicted result with the socioeconomic data (Fig. 12.7a) was better than the one without the socioeconomic data (Fig. 12.7b) with more white pixels. This could be confirmed in the southwestern Houston, such as Gulfton, Sharpstown, and Bellaire, and southeastern Houston between Deer Park and Pasadena. The incorrect predictions were always also found in the industrial/commercial area in the Southern and Northeastern Houston, such as Missouri City and Jersey village. It was easy to understand since the chosen socioeconomic data, especially the population density and house density, were better to represent the residential area instead of industrial/commercial area.

Further validation of models between the simulated one and predicted one was analyzed through the confusion matrix (Table 12.5). This table showed the comparison results between the simulated result and empirical maps as the value of

**Table 12.5** Confusion matrix and the model validation for two models

Empirical map										
Predicted result										
<b>A: the predicted model with socioeconomic factor</b>										
Houston	Residential	Industrial/commercial	Grassland	Woodland	Barren/soil	User's accuracy (%)	Producer's accuracy (%)			
Residential	1,756,004	257,388	230,903	201,696	175,892	66.97	77.32			
Industrial/commercial	123,890	467,825	34,540	7,576	110,447	62.86	54.56			
Grassland	320,672	109,304	518,946	209,275	185,157	38.63	59.60			
Woodland	68,802	15,960	84,838	419,194	28,620	67.90	50.02			
Barren/soil	1,688	6,949	1,544	357	7,258	40.78	1.43			
<b>Overall accuracy (%): 59.30; Kappa: 0.43</b>										
<b>B: the predicted model without socioeconomic factor</b>										
Houston	Residential	Industrial/commercial	Grassland	Woodland	Barren/soil	User's accuracy (%)	Producer's accuracy (%)			
Residential	1,095,900	219,890	210,650	150,980	158,990	59.03	53.91			
Industrial/commercial	71,801	229,830	48,414	861	48,414	55.87	27.90			
Grassland	585,370	162,880	467,240	186,970	195,860	29.23	54.18			
Woodland	189,290	44,897	140,520	431,390	52,279	46.97	51.63			
Barren/soil	90,344	116,409	33,504	5,363	63,262	20.48	11.74			
<b>Overall accuracy (%): 45.56; Kappa: 0.27</b>										

user's accuracy and producer's accuracy represented the accuracy for each class and the overall accuracy and Kappa represented the accuracy for the entire landscape. In both models, the best predicted class was the residential area (with 66.97 %/59.03 % user's accuracy and 77.32 %/53.91 % producer's accuracy) and the worst prediction class was barren/soil (with 40.78 %/20.48 % and 1.43 %/11.74 %). The barren/soil class, although had the least area in the study area, were easy to be confused with other classes such as industrial/commercial or residential area. The incorporation socioeconomic data into the model improved the simulation on the residential or industrial/commercial classes which made the barren/soil having the least accuracy.

One disadvantage of incorporating socioeconomic data into the model was the overestimation of residential area in which led to a relative underestimate in the industrial landscape as well as other natural landscapes such as woodland and grassland. This might be improved as more and more socioeconomic data were incorporated as driving forces in the model. The analysis of the model validation showed that the appropriate ancillary parameters were necessary for the CA model to derive a solid result. In fact, the value of the simulation approach lied in its exploratory nature which enabled the improvement of models with additional variables later. Meanwhile, the CA model had an "aggregate" function to smooth the heterogeneous pattern within the urban and suburban area. One solution to solve this problem was to incorporate better data source into the model, such as higher spatial resolution images or sub-pixel classifications, to improve the accuracy of CA models.

## 12.6 Conclusion

The spatiotemporal CA model of urban landscape patterns using multi-temporal TM and MSS imagery enabled us to characterize the internal structure of landscapes and monitor the landscape dynamics for Houston. Moreover, we also explored the potential of socioeconomic variables to detect how human forces affect the urban spatial pattern.

The CA model, coupled with the Markov transition probability, has indicated the capability of trend projection for the landscape change. This spatiotemporal model provided not only the quantitative description of change in the past but also the direction and magnitude of change in the future. However, based on the experimental results and exploratory analysis, several limitations still exist within the current study:

- Since the modeling process involves the usages of data from multiple sources, the accuracy of prediction result will be closely related to the individual accuracy with each type of data, especially different remote sensing data sources. The development of a robust method to incorporate data in different spatial resolution was still an interesting issue.
- Although the Markov transition probability was calculated on the census block level, it was stationary and unable to accommodate the unpredictable influence

variables, such as the climate, policy, and human disturbance. In addition, the pace of landscape change was usually kept on changing over the entire period.

- In this research, we supposed the relationship between socioeconomic factors, neighborhood effect, and Markov transition probability was linear and deterministic during the calibration. Finding an exact dynamic coefficient between them was still an intricate study in the urban modeling.

Currently, it was not fully conclusive that the CA model based on socioeconomic data was inferior to the one without socioeconomic data, especially for the natural landscapes. It was still necessary to find more sophisticated methods applying to a series of varied landscape to verify this new model.

Most urban landscapes have been influenced by human disturbance, resulting in a heterogeneous mosaic of natural and human-managed patches that vary in size, shape, and arrangements (Turner 1989). The landscape responses to human disturbances are important, however, difficult to be estimated because the landscape-level simulation involved numerous challenging experiments and hypotheses in the development of models (Vaz et al. 2012). These hypotheses are always assumed to make the process model easier to be manipulated, leading to a more homogenous pattern in the predicted result. Thus, it is necessary to relate the homogenous analysis in the model prediction with the heterogeneous analysis in the quantitative landscape method for a comprehensive understanding of the urbanization process. In conclusion, this urban studies show that by incorporating more spatial algorithms into the prediction of landscape change, more accurate long-term landscape changes can be reproduced in the future.

## References

- Aljoufie M, Zuidgeest M, Brussel M, Van Vliet J, Van Maarseveen M (2013) A cellular-automata based land use and transport interaction model applied to Jeddah, Saudi Arabia. *Landscape Urban Plan* 112:89–99
- An L, Linderman M, Qi J, Shortridge A, Liu J (2005) Exploring complexity in a human-environment system: an agent-based spatial model for multidisciplinary and multiscale integration. *Ann Assoc Am Geogr* 95(1):54–79
- Antrop M (2004) Landscape change and the urbanization process in Europe. *Landscape Urban Plan* 67:9–26
- Arsanjani JJ, Helbich M, Kainz W, Darvishi A (2013) Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *Int J Appl Earth Obs* 21:265–275
- Batty M, Xie Y (1994) From cells to cities. *Environ Plan B* 21:31–48
- Bell EJ (1974) Markov analysis of land use change: an application of stochastic processes to remotely sensed data. *Socio Econ Plan Sci* 8:311–316
- Berry MW, Flamm RO, Hazen BC et al (1996) Lucas: a system for modeling land-use change. *IEEE Comput Sci Eng* 3(1):24–35
- Bourne LS (1976) Monitoring change and evaluation the impact of planning policy on urban structure: a Markov chain experiment. *Plan Can* 16:5–14
- Cecchini A, Viola F (1990) Eine Stadtbausimulation. *Wissenschaftliche Zeitschrift der Hochschule für Architektur und Bauwesen* 36(1):159–162



- Clarke KC, Hoppen S (1997) A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environ Plan B* 24:247–261
- Clarke KC, Gaydos LJ (1998) Loose-coupling a cellular automata model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *Int J Geogr Inf Sci* 12(7):699–714
- De Kong GHJ, Verburg PH, Veldkamp A, Fresco LO (1999) Multi-scale modelling of land use change dynamics in Ecuador. *Agr Syst* 61:77–93
- Hagen-Zanker A, Lajoie G (2008) Neutral models of landscape change as benchmarks in the assessment of model performance. *Landsc Urban Plan* 86:284–296
- Herold M (2004) Remote sensing and spatial metrics for mapping and modeling of urban structures and growth dynamics. Ph.D. dissertation, University of California-Santa Barbara
- Herold M, Goldstein NC, Clarke KC (2003) The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote Sens Environ* 86(3):286–302
- Hillier B, Hanson J (1984) *The social logic of space*. Cambridge University Press, Cambridge
- Key to the city (2001) Houston, Harris County, Texas. Available at: <http://www.uscitiesonline.com/txcountyhouston.htm>. Cited 15 Jan 2014
- Klosterman RE (1999) The what if? Collaborative planning support system. *Environ Plan B* 26(3):393–407
- Li X, Yeh AGO (2000) Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *Int J Geogr Inf Sci* 14(2):131–152
- Lichtenberg ER (1985) The role of land quality in agricultural diversification. Ph.D. dissertation, University of California-Berkeley
- Liebrand WBG, Nowak A, Hegselmann R (1998) *Computer modeling of social process*. Sage Publications, London
- Lopez E, Bocco G, Mendoza M, Duhau E (2001) Predicting land cover and land use change in the urban fringe: a case in Morelia city Mexico. *Landsc Urban Plan* 55(4):271–285
- Mcintyre NE, Knowles-Yanez K, Hope D (2000) Urban ecology as an interdisciplinary field: differences in the use of “urban” between the social and natural sciences. *Urban Ecosyst* 4:5–24
- Moser DC (1998) Diane Moser Properties. Available online at: <http://www.texasbest.com/houston/geograph.html>. Cited 15 Jan 2014
- Palmquist RB (1989) Land as a differentiated factor of production: a Hedonic model and its implications for welfare measurement. *Land Econ* 65(1):23–28
- Pickett STA, Burch WR, Dalton SE, Foresman TW, Grove JM, Rowntree R (1997) A conceptual framework for the study of human ecosystems in urban areas. *Urban Ecosyst* 1:185–199
- Pijanowski BC, Long DT, Sage SH, Cooper WE (1997) A land transformation model: conceptual elements, spatial object class hierarchies, GIS command syntax and an application to Michigan’s Saginaw Bay Watershed. Land use modeling workshop, Sioux Fall, South Dakota, June 3–5, 1997.
- Pijanowski BC, Tayyebi A, Doucette J et al (2014) A big data urban growth simulation at a national scale: configuring the GIS and neural network based Land Transformation Model to run in a high performance computing environment. *Environ Model Softw* 51:250–268
- Pontius RG Jr, Cheuk ML (2006) A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. *Int J Geogr Inf Sci* 20(1):1–30
- Pontius RG Jr, Shusas E, McEachren M (2004) Detecting important categorical land changes while accounting for persistence. *Agric Ecosyst Environ* 101:251–268
- Shafizadeh-Moghadam H, Helbich M (2013) Spatiotemporal urbanization process in the megacity of Mumbai, India: a Markov chains-cellular automata urban growth model. *Appl Geogr* 40:140–149
- Stewart W (1994) *Introduction to the numerical solution of Markov chains*. Princeton University Press, Princeton
- Tang J (2011) Modeling urban landscape dynamics using sub-pixel fractions and fuzzy cellular automata. *Environ Plann B* 38:903–920
- Tang J, Wang L, Yao Z (2007) Spatio-temporal urban landscape change analysis using Markov chain and modified genetic algorithm. *Int J Remote Sens* 28(15):3255–3271

- Tang J, Wang L, Yao Z (2008) Analyses of urban landscape dynamics using multi-temporal satellite images: a comparison of two petroleum-oriented cities. *Landsc Urban Plan* 87(4):269–278
- Tang J, Chen F, Schwartz SS (2012) Assessing spatiotemporal variations of greenness in the Baltimore-Washington corridor area. *Landsc Urban Plan* 105:296–306
- Tayyebi A, Pijanowski BC, Tayyebi AH (2011) An urban growth boundary model using neural networks, GIS and radial parameterization: an application to Tehran, Iran. *Landsc Urban Plan* 100(1):35–44
- Tayyebi A, Pekin BK, Pijanowski BC et al (2012) Hierarchical modeling of urban growth across the conterminous USA: developing meso-scale quantity drivers for the land transformation model. *J Land Use Sci* 8(4):422–442
- Tayyebi A, Perry PC, Tayyebi AH (2013) Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster-vector routines with remote sensing and GIS. *Int J Geogr Inf Sci* 28:639–659. doi:[10.1080/13658816.2013.845892](https://doi.org/10.1080/13658816.2013.845892)
- Texas State Historical Association (2002) Handbook of Texas, online. Available at: <http://www.tsha.utexas.edu/handbook/online/articles/view/HH/hdh3.html>. Cited 26 Dec 2009
- Tobler W (1979) Cellular geography. In: Gale G, Olsson S (eds) *Philosophy in geography*. Reidel, Dordrecht, pp 379–386
- Turner MG (1989) Landscape ecology: the effects of pattern on process. *Annu Rev Ecol Syst* 20:171–197
- US Census (2010) Population and household. Available at: <http://www.census.gov>. Cited 20 Sep 2013
- U.S. Census (2011) Your gateway to census 2010. Available at: <http://en.wikipedia.org/wiki/Urbanization>. Cited 15 Jan 2014
- Van Vliet J, Bregt AK, Hagen-Zanker A (2011) Revisiting Kappa to account for change in the accuracy assessment of land use change models. *Ecol Model* 222:1367–1375
- Van Vliet J, Hurkens J, White R, Van Delden H (2012) An activity based cellular automaton model to simulate land use dynamics. *Environ Plan B* 39:198–212
- Vaz E, Nijkamp P, Painho M, Caetano M (2012) A multi-scenario forecast of urban change: a study on urban growth in the Algarve. *Landsc Urban Plan* 104(2):201–211
- Waddell P (2002) UrbanSim: modeling urban development for land use, transportation, and environmental planning. *J Am Plan Assoc* 68(3):297–313
- Weng Q (2002) Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modeling. *J Environ Manag* 64(2):273–284
- White R, Engelen G (1993) Cellular automata and fractal urban form: a cellular modeling approach to the evolution of urban land-use patterns. *Environ Plan A* 25(8):1175–1199
- Wu F (1998) Simulating urban encroachment on rural land with fuzzy-logic-controlled cellular automata in a geographical information system. *J Environ Manag* 53(4):293–308
- Wu F, Martin D (2002) Urban expansion simulation of Southeast England using population surface modelling and cellular automata. *Environ Plan A* 34(10):1855–1876
- Yang X, Lo CP (2002) Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. *Int J Remote Sens* 23(9):1775–1798