Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/090225

An Integrated Approach based on Markov Chain and Cellular Automata to Simulation of Urban Land Use Changes

*Jiejun Huang**, *Yanyan Wu, Ting Gao, Yunjun Zhan and Wei Cui*

School of Resource and Environmental Engineering, Wuhan University of Technology, Wuhan 430070, China

Received: 27 May 2014, Revised: 26 Aug. 2014, Accepted: 28 Aug. 2014
Published online: 1 Mar. 2015

Abstract: Land use change is one of the most important scientific research themes in the field of global environmental change. Due to the presence of uncertainty and randomness in the real world, it is difficult to simulate land use change exactly. To address the spatial uncertainty and temporal randomness in land use change, we propose a model for simulating land use change based on Markov chain and cellular automata (CA), and describes its application to the simulation of land use changes in the city of Wuhan, China. To simulate the urban land use change, the transition rules of the model were first set by globally restrained conditions, locally restrained conditions and a random variable. And then land use patterns and changes were obtained from classified Landsat TM images. A spatial-temporal transition matrix was constructed from the classified images and was applied to the proposed model for simulating land use changes in the city of Wuhan. The experiment results show the validity and feasibility of the Markov-CA-based model for simulating urban land use change.

Keywords: Land use change, Markov chain, cellular automata, simulation models

1 Introduction

Land use and cover change (LUCC) is a hot topic in global change research and has important scientific value for sustainable urban development. To utilize available land resources effectively, GIS, RS and artificial intelligence technologies are being employed to investigate the characteristics and causes of LUCC across a range of spatial and temporal scales. Monitoring and simulation of land use changes at multiple temporal and spatial scales is crucial for understanding, managing and optimizing of land use. What can we do to effectively utilize and protect land resources? How can we better characterize attributes and dynamics of land use and cover? And how will the patterns and characteristics of land use and cover change be in the future? Many scientists and organizations have studied these questions over the last decades. "International Geosphere-Biosphere Program" (IGBP) and the "International Human Dimensions Programme on Environmental Change" (IHDP) proposed the "land use/land cover change" (LUCC) research project jointly in 1995 and made land use/cover change the forefront of global change research [7]. Mahmood et al. [8] have made several recommendations for detecting land use and cover change (LUCC) and understanding its impacts on climate change. Cellular Automata approach has been used to simulate urban expansion and land use evolution [6,17]. Tan et al. [14] investigated and evaluated the impact of Land surface temperature (LST) with respect to land use changes in Penang Island, Malaysia. Pontius et al. [10] presented an assessment model that predicts land use/cover changes among land categories between two points. To overcome the inadequacies of Cellular Automata, nonlinear coupling Cellular Automata supported vector machines, fuzzy set theory and fuzzy logic are used to improve the original Cellular Automata model [1]. Kleynhans et al. [4] have proposed a method for detecting land cover change using time-series NDVI data derived from 500-m MODIS satellite images. Ettabaa et al. [2] have established a CA model based on a multi-agent system to detect and simulate land use change in the northern part of Tunisia. Lewis [5] developed a joint econometric-simulation framework is for forecasting

* Corresponding author e-mail: hjj@whut.edu.cn

detailed spatial pattern of land-use and ecosystem change. Quang Bao Le, et al. [11,12] used dynamic multi-agent land use simulation systems to study the impact of long-term intervention policies on the improvement of the environment and social economic benefits. Guo et al. [16] used the Chaos Genetic Algorithm to improve the accuracy of monitoring land use changes. Ying Pan et al. [15] explored the influence of scale on a CA-based land use change monitoring model, and found that changes in the CA cell size, neighborhood size and shape all have a certain effect on the modeling results. Huang et al. [3] studied spatio-temporal changes of land use in Wuhan City based on Remote Sensing and Geographical Information System. Schweitzer et al. [13] developed a generic modeling software package for simulation of terrestrial environments (SITE) that can customize simulation models for regional land-use dynamics. The latest simulation models of land use change include logistic regression model, stochastic model and cellular automata model [18,19,20].

However, these studies could not simulate the uncertainty and randomness in land use change comprehensively. To address the spatial uncertainty and temporal randomness in land use change, this paper proposes a novel simulation model that employs cellular automata to simulate spatial evolution and employs Markov chain to simulate randomness respectively.

This paper is organized as follows. Section 2 describes the procedure for implementing the Markov-CA based simulation model. Section 3 explains how to define transition rules for implementing the Markov-CA based simulation model. Section 4 evaluates the integrated model via a simulation experiment in the study area. Finally, section 5 draws conclusions for this study.

2 The procedure for implementing the Markov-CA Model

The land use change simulation model presented in this paper is an integration of the Markov method and a CA model. A typical CA system consists of four components: cells, states, neighborhood and transition rules. Transition rules are the real engines of changes in a CA. These transition rules control the transformation from one cell state to another one over a specific period of time, based on the states of the neighborhood cells. The land use change simulation procedure includes definition of spatial objects, setting of restriction conditions, determination of model parameters, setting of cellular transition rules, and validation of the model, as described below.

2.1 Defining spatial objects

To define the spatial objects, the simulation objects are first divided into regular grids. Each grid can be considered as a cell. For urban land, the cellular space consists of all the grids with in the study area. Each cell may appear in one of the following states: vegetated land, agriculture land, developed land, water or undeveloped land. The expanded Moore model is adopted here to realize neighborhood extension. The transition rule, the core of the Markov-CA model, is a function of changes in cellular objects, and is the drive of land use pattern transformation.

2.2 Setting restriction conditions for objects

The state changes of cells from one cell state to another one are restricted by two major restriction conditions: the spatial value of cells and change threshold of cell state. Spatial value influence urban land use change, but threshold values control the type change of land use. Not only is urban development influenced by spatial factors such as traffic, hydrology, and terrain, but also is influenced by local environment. For example, many cities were developed by extending traffic lines along river valleys.

2.3 Determining model parameters

In the Markov-CA-based simulation model, each parameter is considered as a factor that influences urban land use change. Each factor is set to a weight value according to some regression methods. The weight value can be set to represent the degree to which the corresponding factor impacts the state of land use. Land use change may be influenced by some spatial parameters such as size and width of the study area, transition direction and rate, etc. Furthermore, land use change is influenced by some temporal parameters such as its attribute, amount and spatial location. In the timespace orthogonal coordinate system, the proposed Markov-CA-based model can simulate and predict the land use change trends and characteristics.

2.4 Setting transition rules

On the basis of the three steps above, a standard cellular model is built. The last key step of Markov-CA-based model is to set the transition rule that is a function of cellular object changes, and determine the transition conditions of land use pattern. To define land use change transition rules in the Markov-CA model, global restriction conditions, local restriction conditions and random variables are introduced. The global restriction conditions limit the global change trend of land use. Furthermore, to better simulate the spatial change trend of urban land use, the Markov model needs to define the change probabilities of the state of land use by setting the thresholds for changing the types of land use and random factors.

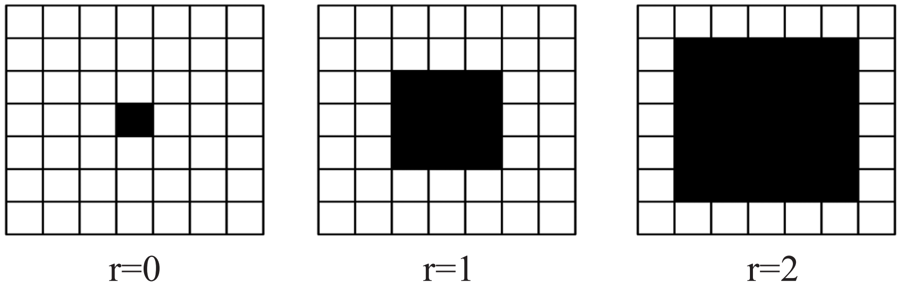


Fig. 1: Neighborhood Radius Diagram

2.5 Implementing and applying the model

Once built according to the methods described in section 2.1-2.4, the Markov-CA-based model can be applied to simulate spatial and temporal change of urban land use for the study area, by employing the parameters and transition rules, set according to the real information of the study area. Furthermore, the Markov-CA-based model can predict the trend of land use in the future after validation of the ability of the proposed model for simulating the extent and spatial characteristics of land use change.

3 Methods

The core of the Markov-CA model is definition transition rules to control the change of cell states during simulation. Because transition rules may directly influence the validity of the result, we need to consider many factors. In this study the transition rules are defined by global restriction conditions, local restriction conditions and a random variable. The global restriction conditions include all spatial factors, such as distances to railways, roads and regional centers; the local restriction conditions include the neighborhood radii of the cells; and the random variable is a supplementary value introduced to reflect the complexity of urban land evolution better.

3.1 Global restriction conditions

In simulation process of urban land use, the cell with a bigger change probability means that it is more suitable to be developed for use. The suitability of a land for urban development is determined by a series of factors, such as natural factors, socioeconomic indices and government policies. In the proposed Markov-CA-based simulation model, the restriction conditions only include traffic conditions and terrain. The global restriction condition is defined by equation (1).

$$p_{ij}^{t} = \frac{1}{1+\exp(-r_{ij}^{t})} \tag{1}$$

Here $r_{ij}^{t} = a_0 + a_1 \cdot x_1^t + a_2 \cdot x_2^t + \ldots + a_n \cdot x_n^t$, x_1^t, x_2^t, $\ldots$, x_n^t represent spatial variables affecting urban development, such as the distance to centers of industrial zones, the distance to railways, the distance to roads and current land use type. a_0, a_1, a_2, $\ldots$, a_n represent weights of spatial variables.

Defining weights of the spatial variables is an essential step of setting transition rules. The weights can be obtained from a regression method. The probability of land use change is influenced by many factors. A line regression model can be used to estimate the probability (P) of land use change ($0 \le P \le 1$), and is expressed as

$$\hat{p} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \tag{2}$$

$$\hat{p} = \mathrm{logic}(P) = \ln(P/1-P) \tag{3}$$

$$\begin{aligned} P &= \frac{\exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n)}{1+\exp(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n)} \\ &= \frac{1}{1+\exp[-(\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n)]} \end{aligned} \tag{4}$$

Here x_1, x_2, $\ldots$, x_n are a series of spatial distance variables, and α, β_1, β_2, $\ldots$, β_n are the weights of the distance variables. The error of the model should conform to a binomial distribution.

3.2 Local restriction conditions

To define the local restriction conditions, the land use types of the cells surrounding the current cell and the proportion of each type are first determined. The influence radius of a cell can be defined according to the land use type of this cell. In geographical space, the degree of influence of an object usually decreases as the distance to this object increases. If the radius is set too big, the neighborhood will not have significant influence on the center cell. Therefore, a proper influence radius must be selected for the proposed Markov-CA-based simulation model (see Fig. 1).

The change probability of pixel i can be written as

$$p(i) = \frac{\sum_n con(i)}{n-1} \quad (5)$$

Here "i" represents the types of land use, including developed land, undeveloped land, agricultural land and vegetated land, $con(i)$ represents the number of the cells classified as the type within the neighborhood, $p(i)$ is the probability of that cell becoming developed land, and n is the number of all cells within the influence radius.

3.3 Setting the random variable

The simulation results are more accurate when a proper random variable is introduced into the proposed model. The random variable can be defined as

$$R = 1 + (-\ln \gamma)^{\alpha} \quad (6)$$

Here γ is a random value between 0 to 1, and α is a parameter that controls the value of the random variable.

3.4 Setting transition rules

Based on two images captured at different times, we can obtain a state change probability pattern for land use. The thresholds for transition rules can be set based on the probability image and the real land use information. Once the transition rules are ready, the land use change pattern can be simulated. Because urban land simulation process can be influenced by all the above three restriction conditions, the rules are designed as:

$$P_t = p_{ij}^t \times p(i) \times R \quad (7)$$

Here P_t is the total probability, p_{ij}^t is the global restriction value of the cell, $p(i)$ is the restriction value for the cell neighborhood, and R represents the random factor during urban development.

The total probability is then normalized to a value within (0, 1) and compared with the threshold $P_{threshold}$ When $P_t < P_{threshold}$, the land use is changed to other land, including vegetated land, agriculture land and undeveloped land, as shown as follows.

$$\begin{cases} P_t \geq P_{threshold},\ i \rightarrow \text{Developed land} \\ P_t < P_{threshold},\ i \rightarrow \text{Others land} \end{cases} \quad (8)$$

The transition rules are defined as follows: under the local restriction conditions, the state change probability for each cell can be calculated using the number of corresponding cells in the preset neighborhood (Eq. 7). When the global restriction condition and random factor remain unchanged, the value of P_i' have the largest probability to change to vegetated land, undeveloped land or agriculture land.

$$\begin{cases} \text{If } P_i' \geq P_{i\ threshold},\ \text{change } \rightarrow i \\ \text{If } P_i' < P_{i\ threshold},\ \text{unchanged} \end{cases} \quad (9)$$

Where i represents one of three land use types: agriculture land, vegetated land and undeveloped land.

3.5 Training and development the proposed model

Based on the classified images, we can get information about land use change probabilities. Table 1 shows the probabilities of land use change in Wuhan city from 1999 to 2002.

As shown in the Table 1, the change probabilities of different types of land use are obviously different. For the developed land, the state change probability values, $P\{agricultural, undeveloped, developed, vegetated\}$, are $\{0.0934,\ 0.0796,\ 0.7594,\ 0.0676\}$. These values are the probabilities of developed land converted to the other three land use types. Whereas, for water area the state change probabilities are very limited, we can think that the water cells did not experience any change. In other words, those cells will still be "water" in the next state.

In the selected Markov image, the probability values of the cells do not constitute a continuous range. In the actual land development process, characteristics of land use changes vary with locations. Because land use change is spatial and temporal, combination of Markov model and CA is an appropriate approach for simulating and predicting this process. For example, spatial factors such as distances to agriculture land, vegetated land, undeveloped land and developed land are used to control the extent and rate of urban land expansion. The direction and width of urban land expansion can be determined by adjusting the threshold values. Three distance factors, distances to railways, roads and regional centers, are applied to the Markov probability images of 1999 and 2002. The simulation parameters are calibrated by testing and analyzing the simulated and actual land use images. The calibration process ends when the simulated amount of each land use type is equal to the real value. Then the local P_{ij} is calculated using the neighborhood radius. Once the random factor R is determined, the total probability can be calculated using equation (7). The simulation process is shown in Fig. 2. The left diagram 'weight P_b calculation' shows the steps of calculating the standardized weight values of P_b in the right diagram.

4 Results and discussion

This section describes the simulation experiment we conducted to verify the proposed model. The study area is

Table 1: Transition Probability Patterns of Land Use from 1999 to 2002

Year		2002			
1999		Agricultural Land	Undeveloped Land	Developed Land	Vegetated Land
	Agricultural Land	0.6585	0.0287	0.3111	0.2016
	Undeveloped Land	0.127	0.5097	0.2507	0.0926
	Developed Land	0.0934	0.0796	0.7594	0.0676
	Vegetated Land	0.1316	0.0378	0.1904	0.6807

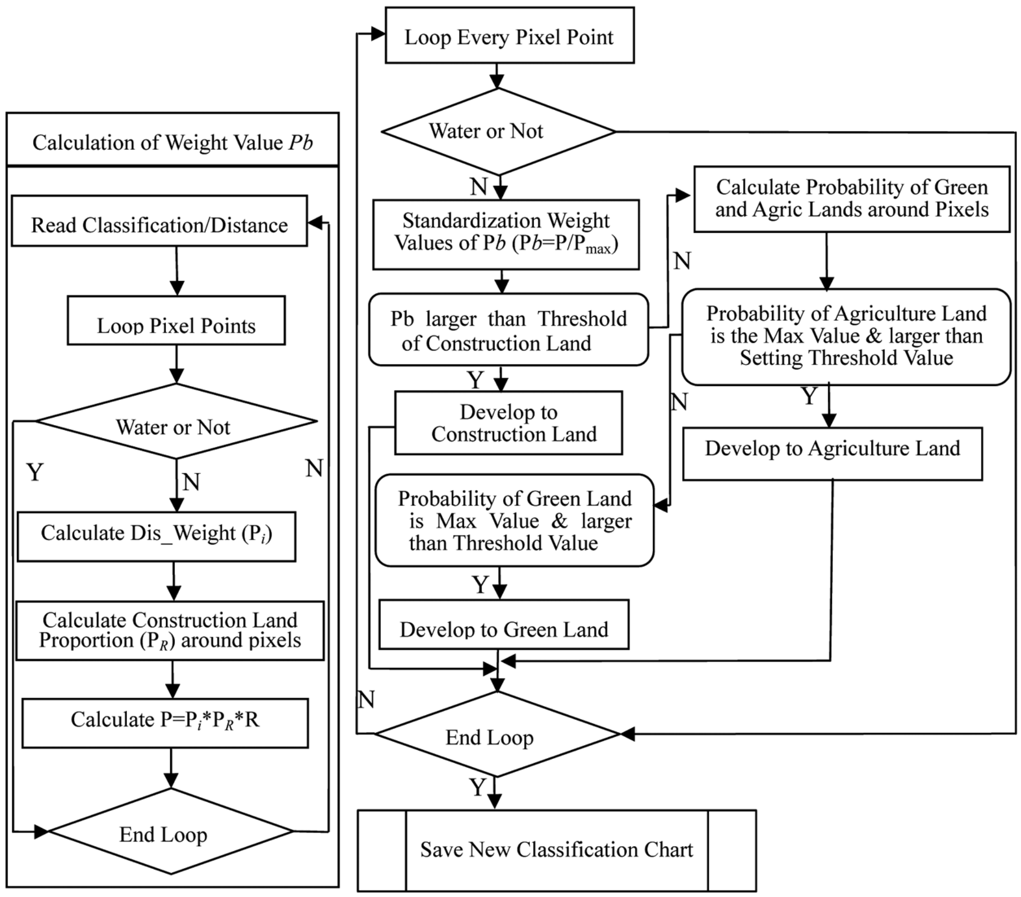


Fig. 2: Flow of urban land use change simulation

located in Wuhan City. Wuhan lies in the eastern Jianghan Plain at the intersection Yangtze and Han rivers. Lakes are densely covered in Wuhan, the network of rivers is vertical and horizontal and this city has been always called a city with One Thousand Lakes.

Theoretically, the number of land use types should increase or remain stable over time. But the change rate shows that developed land has been increasing constantly. When urbanization achieves a certain degree, urban land will no longer expand, and the rate of land use change will decrease.

The simulation parameters are determined through multiple tests and statistical analyses. After 18 times of simulation, the resulting image of 1999 became extremely similar to the actual image of 2002. These 18 times of land use change occurred in about three years, meaning that six times of land use change in the model represent one year's land use change in the real world. Based on the estimated parameters and transition rules, the simulated land use image of Wuhan city in 2005 is obtained by running the proposed model (see Fig. 3).

Then the classified image of 2005 is employed as the validation dataset to check the accuracy of simulation results. The error rates of simulated agriculture land, undeveloped land, developed land and vegetated land are 5.57%, 9.80%, 6.56% and17.70%, respectively. The largest error rate occur to the simulated vegetated land, this is because the government artificially increased the

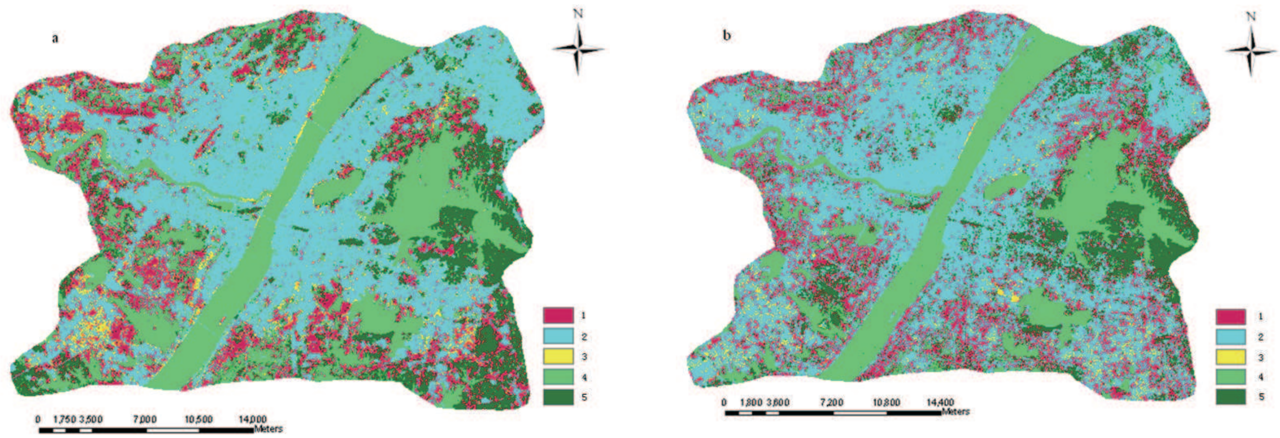


Fig. 3: Comparison between the simulation result (a) and the actual image (b) of 2005 (1 represents agriculture land, 2 represents developed land, 3 represents undeveloped land, 4 represents water and 5 represents vegetated land)

vegetation cover rate from 2002 to 2005 by converting many land use units to vegetated land. The large error in simulated undeveloped land comes from the intrinsic dynamic nature of this type of land. Within the simulated process of urbanization, large tracts of undeveloped land is changed to industrial, residential and vegetated land. However, parts of deserted industrial land and wilderness around the lakes remain undeveloped. Overall, the simulated land use pattern is very close to the actual land use in Wuhan city with an average accuracy of 87.75%. This level of simulation accuracy shows the efficiency of the proposed modeling approach for simulating of land use change.

5 Conclusions

Land use change is a complicated process. In this study, we analyzed land use patterns in Wuhan city using satellite images from 1999 to 2005, and presented a Markov based cellular automata (CA) model for simulating and predicting urban land use change. Several Landsat TM images were used to generate land use maps and to acquire the actual land-use change patterns. A spatio-temporal transition matrix was constructed from the classified images and applied in the proposed model to simulate land use changes in the study area. The experiment results showed the validity and feasibility of the Markov-CA model for land use change simulation. Not only could the propose land use change model help us understand the complexity of the components of spatial systems, but also provide theories and reference for land planning and land resource management.

In future study, we will strengthen cellular component expansion with intelligent algorithms such as support vector machine and ant colony algorithm. Socioeconomic factors and urban system evolution theories should also be considered to add into the model to improve the simulation accuracy. In addition, the potential impact of land use changes on human activities and the dynamic mechanism of urban land development need to be further studied.

Acknowledgment

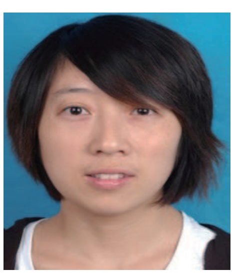
We gratefully acknowledge the helpful comments and suggestions offered by the anonymous reviewers. The work presented in this paper was supported by the National Natural Science Foundation of China (No. 41071104, 41171319), the Youth Chenguang Project of Science and Technology of Wuhan City of China (No. 200950431203), the Fundamental Research Funds for the Central Universities (WUT: 2014-II-007, 2014-IV-141)

References

[1] S. Dragicevic, Fuzzy Information **1**, 203-207 (2004).
[2] K.S Ettabaa, I.R Farah, M.B Ahmed, Information and Communication Technologies **2**, 3064-3069 (2006).
[3] J.J Huang, Y.Y Wu, Y.X Wu, Y.J Zhan, H. Wu, International Journal of Advancements in Computing Technology **4**, 215-222 (2012).
[4] W. Kleynhans, J.C Olivier,K.J Wessels, B.P Salmon, F. Van Den Bergh, K. Steenkamp, IEEE Geoscience and Remote Sensing Letters **8**, 507-511 (2011).
[5] J. Lewis David, Resource and Energy Economics **32**, 98-116 (2010).
[6] A.-M Li, A.M Lv, C.L Sui, Journal of ZhengZhou Institute of Surveying and Mapping **26**, 165-169 (2009).
[7] B.L Turner, D.S Skole, Land-use and land-cover change science/research plan, IGBP report No.35 and HDP Report No.7. Stochkholm, IGBP. (1995)
[8] R. Mahmood, A. Pielke Sr Roger, G. Hubbard Kenneth, D. Niyogi,G. Bonan, Bulletin of the American Meteorological Society **91**, 37-46 (2010).

[9] O. Okwuashi, P. Nwilo, J. McConchie, E. Eyo, Enhancing a GIS Cellular Automata Model of Land Use Change Using Support Vector Machine In Proceeding(s) of the 17th International Conference on Geoinformatics, Fairfa (2009) pp. 1-5.
[10] Jr. Pontius, G. Robert, Petrova, H. Silvia, Environmental Modelling and Software **25**, 299-309 (2010).
[11] Q.B Le, S.J Park, L.G.V Paul, B.C. Armin, Ecological Informatics. **3**, 135-153 (2008).
[12] Q.B Le, S.J Park, L.G.V Paul, Ecological Informatics **5**, 203-221 (2010).
[13] Schweitzer Christian Priess Joerg A Das Subhashree, Environmental Modelling and Software. **26**, 1052-1055 (2011).
[14] Tan Kok Chooi Lim Hwee San MatJafri Mohd Zubir Abdullah Khiruddin, Malaysia Environmental Earth Sciences **60**, 1509-1521 (2010).
[15] Y. Pan, R. Andreas, Z.R Yu, D. Reiner, Environment and Urban Systems **34**, 400-408 (2010).
[16] Y.Q Guo, Y.B Wu, Z.S Ju, J. Wang, L.Y Zhao, Mathematical and Computer Modelling **51**, 1408-1416 (2010).
[17] Q. Zhang, S.-Y Hu, Y.-G Zu, Y.-Y Shi, Management of Land and Resources **25**, 94-98 (2008).
[18] Rickebusch Sophie Gellrich Mario Lischke Heike Guisan Antoine Zimmermann Niklaus E, Ecological Modelling **209**, 157-168 (2007).
[19] Moreno Niandry Wang Fang, Marceau Danielle J, Environment and Urban Systems **33**, 44-54 (2009).
[20] Van Vliet Jasper, Bregt Arnold K Hagen-Zanker Alex, Ecological Modelling **222**, 1367-1375 (2011).

Jiejun Huang is professor at the faculty of Resource and Environment Engineering at Wuhan University of Technology. He received his PhD degree in Photogrammetry and Remote Sensing at Wuhan University in 2005. His research is currently focused on urban land use change simulation and modeling.



Yanyan Wu is an engineer in Nanjing Guotie Eiectric Co.,Ltd. She received her MSc in Environmental Engineering in Wuhan University of Technology. Her research is currently focused on application of Geographic Information System.

Ting Gao is a MS student in GIS in Geographic Information Science at Wuhan University of Technology, where she received her BSc in GIS in 2012. Her research is currently focused on land use change simulation and GIS application.

Yunjun Zhan is associate professor at the faculty of Resource and Environment Engineering at Wuhan University of Technology. He received his PhD degree in Geodetection and Information Technology in China University of Geosciences in 2008. His research is focused on urban land-use change and simulation based on cellular automat and multi-agent system.

Wei Cui is professor at the faculty of Resource and Environment Engineering at Wuhan University of Technology. He received his PhD degree in Photogrammetry and Remote Sensing at Wuhan University in 2004. His research is currently focused on Geographic Information System and Environmental Information System.