

A New Method to Model Neighborhood Interaction in Cellular Automata-Based Urban Geosimulation

Yaolong Zhao^{1,2} and Yuji Murayama²

¹ Faculty of Land Resource Engineering, Kunming University of Science and Technology, Kunming, Yunnan 650093, P.R. China
yaolongzhao@gmail.com

² Graduate School of Life and Environmental Sciences, University of Tsukuba, Tennodai 1-1-1, Tsukuba, 305-8572 Japan
mura@atm.geo.tsukuba.ac.jp

Abstract. Local spatial interaction (i.e. neighborhood interaction) between land-use types is an important component in Cellular Automata -based urban geosimulation models. Herein a new method based on the integration of Tobler's First Law of Geography with Reilly's gravity model and coupled with logistical regression approach is proposed to model and calibrate the neighborhood interaction. This method is embedded into a constrained CA model to simulate the spatial process of urban growth in the Tokyo metropolitan area. The results indicate that this method captures the main characteristics of neighborhood interactions in the spatial process of urban growth. Further, this method provides an alternative and extensive approach to present local spatial interactions for "bottom-up" urban models.

Keywords: Geosimulation, cellular automata, urban growth, GIS.

1 Introduction

One of the most important developments in Geographic Information Science (GIScience) is the expansion of theories, models, and technologies to effectively discern and interpret spatiotemporal patterns, relationships, and interactions among features, activities, processes, and events in geographic domains. In current era, as rapid changes of urban land-use all over the world have greatly impacted on local [1-3] and global environmental changes [4, 5], the issue of modeling the spatial process of urban growth to better understand the mechanism and consequences of urbanization and explore the extent of future urban land-use change has attracted sweeping attentions of scientists with background in different disciplines ranging from anthropology to mathematical programming. This issue also has enriched the theory and technology of simulation model of geographic phenomena.

An important component in Cellular Automata (CA)-based urban geosimulation models is the local spatial interaction between neighborhood land-use types. The neighborhood interaction is often addressed based on the notion that urban

development can be conceived as a self-organizing system in which natural constraints and institutional controls (land-use policies) temper the way in which local decision-making processes produce macroscopic urban form. Different processes can explain the importance of neighborhood interaction. At large scale, simple mechanisms for economic interaction between locations were provided by the central place theory [6] that describes the uniform pattern of towns and cities in space as a function of the distance that consumers in the surrounding region travel to the nearest facilities. Spatial interaction between the location of facilities, residential areas and industries has been given more attention in the work of Krugman [7, 8]. The spatial interaction is explained by a number of factors that either cause concentration of urban functions (centripetal forces: economies of scale, localized knowledge spill-over, thick labor markets) and others that lead to a spatial spread of urban functions (centrifugal forces: congestion, land rents, factor immobility etc.).

In keeping with the spirit of simplicity, neighborhood interaction in the application of CA on urban geosimulation models most often adopt either the Von Neumann 3×3 (or 5×5) or the Moore 3×3 neighborhood [9-11]. For most physical systems, these are clearly the most appropriate definitions since such systems typically have only local causation. However, in the case of human systems like cities, the idea of locality may be much larger, since people and institutions are aware of their surroundings in a wider space [12]. Thus it is desirable to define a neighborhood large enough (i.e. extended neighborhood) to capture the operational range of the local processes being modeled by CA. White and Engelen (1993) firstly proposed this kind of configuration of neighborhood for exploring the relationship of CA-based model of urban form evolution [13]. In 1997, White et al. calibrated the neighborhood effect by means of a trial and error approach for geosimulation of Cincinnati city [14]. In 2004, this research group proposed automatic calibration procedure for this kind of neighborhood effect [15].

The objective of this study is to improve the methodology of existing CA models by proposing a theoretical framework to model and calibrate the neighborhood effect. This framework aims to assist modelers in the implementation and quantification of neighborhood interaction in urban geosimulation models.

2 Modeling Neighborhood Interaction

Tobler's First Law of Geography (FLG), "Everything is related to everything else, but near things are more related than distant things", is the fundamental theory in this framework. This law, firstly proposed in 1970 [16], has brought strong controversy in geography domain. In 2003, a panel on this law was organized in AAG meeting in New Orleans. Five famous geographers presented their comments in this panel and these comments were published in a forum of *Annals of the Association of American Geographers* in 2004. Some professors agreed with Tobler, others not. However, all of them accepted the actual geography phenomena illustrated by the FLG. The divarication existed on the word "law". Goodchild especially discussed the validity and usefulness of the FLG in GIScience and geography [17]. Here, the controversy of whether phenomena can be expressed as "law" was discarded, and the local knowledge expressed in Tobler's FLG was accepted. It is assumed that the effect of cell states in

the neighborhood area of developable cell accord with the rule of distance decay described in the FLG.

The expression of Tobler’s FLG is very qualitative, and a distance decay function is needed for representing the law. Herein the idea of Reilly’s law of retail gravitation [18] was adopted, which states that ‘A city will attract retail trade from a town in its surrounding territory, in direct proportion to the population size of the city and in inverse proportion to the square of the distance from the city’. Fig. 1 shows one of the extended neighborhood patterns of one developable cell *i*, which is defined as all cells within a radius of eight cells, an area containing 196 cells. It is assumed that in cellular environment all the cells in the neighborhood contribute to the conversion of developable cell *i*. The contribution of one cell is associated with the state of itself and the distance to the developable cell *i*. It can be express as follows:

$$f_{kh} = G_{kh} \frac{A_j}{d_{ji}^2} \tag{1}$$

Where,

- f_{kh} : contribution of one cell *j* with land-use *k* in the neighborhood to the conversion of the developable cell *i* to land-use *h* for next stage,
- A_j : area of the cell *j*,
- d_{ji} : the Euclidean distance between the cell *j* in the neighborhood area and the developable cell *i*, and
- G_{kh} : constant of the effect of land-use *k* on the transition to land-use *h*. + stands for positive, – repulsive.

Fig.2 indicates the scheme of the impact gradient using this function. Notice that this is a modificatory Reilly’s function and no unit problem exist in this function.

Then the aggregated effect of the cells in the neighborhood can be expressed as:

$$F_{kh} = G_{kh} \sum_{j=1}^m \frac{A_j}{d_{ji}^2} I_{kj} \tag{2}$$

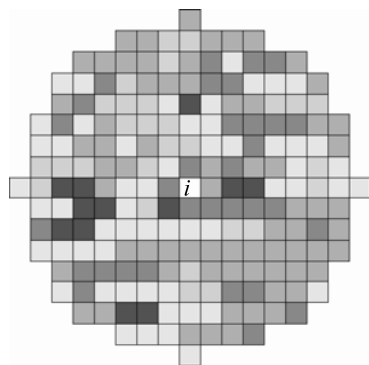


Fig. 1. An extended neighborhood pattern

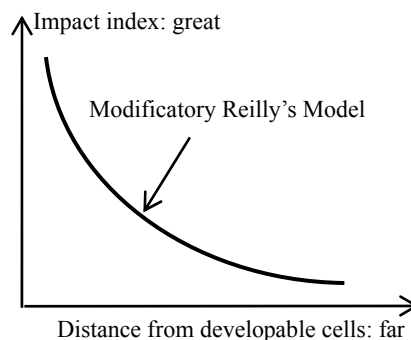


Fig. 2. Scheme of the impact gradient

Where,

m : total number of the cells in the neighborhood, and

I_{kj} : index of cells. $I_{kj}=1$, if the state of cell j is equal to k ; $I_{kj}=0$, otherwise.

For one land-use type of one cell, there are just two results of transition: change or no change. Therefore, logical regression approach was selected to calculate the probabilities of the transition of cell i under neighborhood effect. The contribution of neighborhood effect to the probability of conversion to land-use h of a cell (N_i) is described as a function of a set of aggregated effect of different land-use types:

$$\text{Log}\left(\frac{N_{ih}}{1-N_{ih}}\right) = \beta_{oi} + \sum_k \beta_{ikh} F_{ikh} = \beta_{oi} + \sum_k \beta_{ikh} G_{kh} \sum_m \frac{A_m}{d_{mi}^2} I_{mk} \quad (3)$$

As G_{kh} is a constant, let:

$$\beta'_{oi} = \beta_{oi}, \quad \beta'_{ikh} = \beta_{ikh} G_{kh}$$

Then:

$$\text{Log}\left(\frac{N_{ih}}{1-N_{ih}}\right) = \beta'_{oi} + \sum_k \beta'_{ikh} \sum_m \frac{A_m}{d_{mi}^2} I_{mk} \quad (4)$$

Where, β'_{oi} and β'_{ikh} are the coefficients which should be calibrated.

3 Calibration of Neighborhood Interaction

3.1 Study Area and Data Set

The Tokyo metropolitan area was identified as study area to confirm the approach as this area possesses abundant dataset of land-use. The dataset is named ‘‘Detailed Digital Information (10m grid land-use) Metropolitan Area’’ of Tokyo which was released in 1998 by Geographic Survey Institute of Japan. The data in 1984 and 1989 was used to calibrate the neighborhood effect.

The dataset was designed 10 cell states of land-use and was aggregated into 100×100m grid from 10×10m in majority rule. Land-use of water represents fixed features in the model, that is, this feature are assumed not to change and which therefore do not participate in the dynamics in order to protect the life environment. Two types of land-use of forest & wasteland, and cropland take passive features that participate in the land-use dynamics. However, the dynamics are not driven by an exogenous demand for land. They appear or disappear in response to land being taken or abandoned by the active functions. The active functions are four land-use categories which are forced by exogenously generated demands for land to CA in response to the growth of the urbanized area: vacant, industrial, residential, and commercial land. Other three types of land-use of road, public land, and special land show passively active features which are dynamics in the model. Active and passively active types of land-use are assumed to participate in the neighborhood interaction.

3.2 Calibration of the Neighborhood Interaction

As one of statistical analysis techniques, logistical regression has to consider the problem of spatial statistics like spatial dependence and spatial sampling [19, 20] in the

calibration procedure. The integration of both systematic and random sampling method was adopted to eliminate spatial dependence effect. Firstly, land-use changes were detected from the data set in 1984 and 1989. Systematic sampling was implemented and approximately half cells of the changes of every one of four active land-use types were remained. Then random selection of not changed cells was carried out to create nearly 1:1 ratio for changed cells and not changed cells. Its total size was 27,070 cells. The result of calibration is shown in Table 1.

Table 1. Result of calibration of neighborhood interaction

factors and test	Active land-use types			
	Vacant land	Industrial land	Residential land	Commercial land
Total size (cells) of sampling	11034	1732	11596	2708
$\beta'_{Vacant\ land, h}$	1.147	*0.091	0.190	0.158
$\beta'_{Industrial\ land, h}$	0.334	1.446	0.262	0.457
$\beta'_{Residential\ land, h}$	0.103	**	0.562	0.209
$\beta'_{Commercial\ land, h}$	0.348	0.727	0.181	1.821
$\beta'_{Road, h}$	0.199	**	0.421	0.561
$\beta'_{Public\ land, h}$	0.198	**	0.199	0.224
Constant β'_0	-2.428	-1.988	-2.830	-2.763
Test				
PCP (%)	84.3	87.6	83.6	86.3
ROC	0.924	0.937	0.905	0.937

PCP: Percentage Correctly Predicted (0-100%); ROC: Relative Operating Characteristic (0-1). *: significant at $p < 0.05$; **: non-statistically significant; others significant at $p < 0.001$.

Table 1 illustrates that all the values of PCP of four active land-use types are more than 80% and all of the values of ROC more than 0.9, thus showing goodness of fit of this approach. The results of the test also indicate the existence of neighborhood interaction in urban land-use changes.

4 Simulation and Results

Constrained CA model was put forward to confirm this method. In this model, transition potentials for each cell are calculated as follows:

$${}^tP_{ik} = (1 + {}^{t-1}N_{ik})(1 + {}^{t-1}S_{ik})(1 + {}^{t-1}Z_{ik})(1 + {}^{t-1}A_{ik}){}^{t-1}v. \quad (5)$$

Where, ${}^tP_{ik}$ = the potential of the cell i for land use k in time t ; ${}^{t-1}N_{ik}$ = the neighborhood effect on the cell i for land-use k at time $t-1$, which equals the value of N_{ih} in equation 4; ${}^{t-1}S_{ik}$ = the intrinsic suitability of the cell i for land-use k at time $t-1$; ${}^{t-1}Z_{ik}$ = the zoning status of the cell i for land-use k at time $t-1$; ${}^{t-1}A_{ik}$ the accessibility of the cell i to transportation for land-use k at time $t-1$; ${}^{t-1}v$ is the scalable random perturbation term at time $t-1$. Four active land-use types were changed by an intervention that is exogenous to the CA model from regional systems.

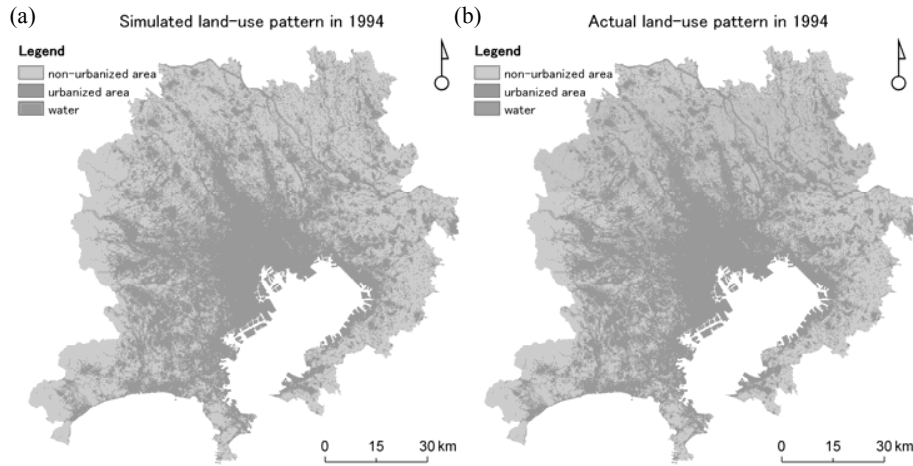


Fig. 3. Land-use maps in the study area: (a) simulation in 1994; (b) reality in 1994

In this schema the neighborhood factor N makes city works like a non-linear system as suitability, accessibility and land-use zoning status are relatively stable in certain period.

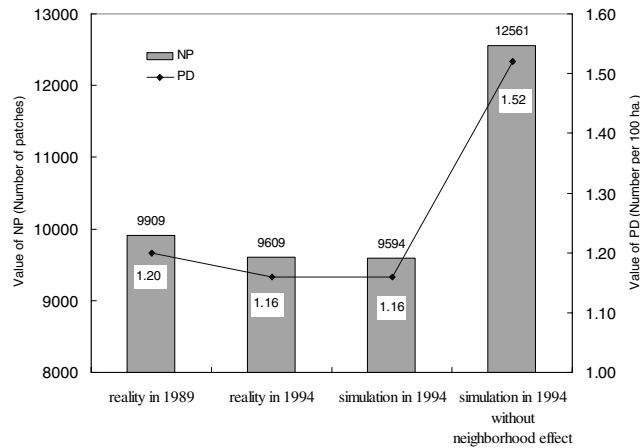
In a learning stage, ${}^{t-1}S_{ik}$, ${}^{t-1}Z_{ik}$, and ${}^{t-1}A_{ik}$ also were calibrated through the dataset in 1984 and 1989. The calibrated model was used to simulate the spatial process of urban growth of the Tokyo metropolitan area from 1989 to 1994. Land-use maps in simulation and reality are shown in Fig. 3. Note that in order to make the maps clearer, land-use types of forest & wasteland and cropland have been grouped into non-urbanized area and other land-use types except water into urbanized area.

A good CA-based model produces results which have all the patterned complexity of the real system [12]. Comparison of the simulated result with the actual data in terms of fractal dimension and spatial metrics was carried out towards the test of this proposed approach. These two indices are excellent for presenting the pattern of complex system like city [21]. Table 2 shows the assessment result of the simulated urbanized area in terms of fractal dimension. Areas, the distance from which to Tokyo station is more than 50km, were omitted considering the effect of boundary in this table. In order to understand the change of urbanized area structure and confirm the ability of this model in capturing the change of structure, fractal dimension of urbanized area in 1989 also is shown. Table 2 indicates that urbanized area shows bifractal structure in the study area. Urbanized area had grown more greatly in the second radius zone with 16-50km than in the first one with 0-16km. Model well captures this characteristic.

Fig.4 shows the assessment result of simulation in terms of spatial metrics. Two metrics were used for the assessment: NP (number of patches) and PD (patch density). Values of NP and PD had declined from 1989 to 1994, thus indicating the characteristics of compact growth or conglomeration of the existing urbanized area in the study area. Simulated urbanized area presents the same characteristics. However, if we take out the component of neighborhood interaction from the model, NP and PD turbidly increased. This confirms the utility of the proposed approach in modeling neighborhood interaction.

Table 2. Assessment of the simulated urbanized area in terms of fractal dimension

Fractal dimension in different radius zones	Reality in 1989	Reality in 1994	Simulation in 1994
In 0-16km radius	1.94	1.95	1.95
In 16-50km radius	1.45	1.48	1.48

**Fig. 4.** Comparison of the significance of spatial metrics of urbanized area between reality and simulation in the Tokyo metropolitan area

5 Concluding Remarks

This study developed a novel approach for modeling and calibrating neighborhood interaction in CA-based urban geosimulation. The proposed method provides a theoretical framework for presenting neighborhood effect of CA. The results of simulation using the Tokyo metropolitan area as a case study indicates that urban geosimulation model which embeds this method well captures the main characteristics of spatial process of urban growth. The results also confirmed the utility of this method for presenting dynamics of complex system.

This approach can be used not only in regular grid cells, but also in irregular cells, like in vector structure, as it considers the area of cells and the distance decay. To discuss this issue would be a value extension to the current study.

References

1. Lin, G.C.S., Ho, S.P.S.: China's land resources and land-use change: insights from the 1996 land survey. *Land Use Policy* 20 (2003) 87-107
2. McKinney, M.L.: Urbanization as a major cause of biotic homogenization. *Biological Conservation* 127 (2006) 247-260
3. Paul, M.J., Meyer, J.L.: Streams in the urban landscape. *Annual Review of Ecology Systematics* 32 (2001) 333-365

4. Grimm, N.B., Grove, J.M., Pickett, S.T.A., Redman, C.L.: Integrated approaches to long-term studies of urban ecological systems. *Bioscience* 50 (2000) 571-584
5. Lambin, E.F., Turner, B.L., Geist, H.J., Agbola, S.B., Angelsen, A., Bruce, J.W., Coomes, O.T., Dirzo, R., Fischer, G., Folke, C., George, P.S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E.F., Mortimore, M., Ramakrishnan, P.S., Richards, J.F., Skanes, H., Steffen, W., Stone, G.D., Svedin, U., veldkamp, T.A., Vogel, C., Xu, J.: The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change* 11 (2001) 261-269
6. Christaller, W.: *Central Places of Southern Germany* (Edition 1966). Prentice Hall, London (1933)
7. Fujita, M., Krugman, P., Mori, T.: On an evolution of hierarchical urban systems. *European Economic Review* 43 (1999) 209-251
8. Krugman, P.: The role of geography in development. *International Regional Science Review* 22 (1999) 142-161
9. Batty, M.: Urban evolution on the desktop: simulation with the use of extended cellular automata. *Environment and Planning A* 30 (1998) 1943-1967
10. Wu, F.: SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. *International Journal of Geographical Information Science* 12 (1998) 63-82
11. Yeh, A.G.O., Xia, L.: A constrained CA model for the simulation and planning of sustainable urban forms by using GIS. *Environment and Planning B* 28 (2001) 733-753
12. White, R., Engelen, G.: High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems* 24 (2000) 383-400
13. White, R., Engelen, G.: Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environment and Planning A* 25 (1993) 1175-1199
14. White, R., Engelen, G., Uljee, I.: The use of constrained cellular automata for high-resolution modeling of urban land-use dynamics. *Environment and Planning B* 24 (1997) 323-343
15. Straatman, B., White, R., Engelen, G.: Towards an automatic calibration procedure for constrained cellular automata. *Computers, Environment and Urban Systems* 28 (2004) 149-170
16. Tobler, W.: A computer movie simulating urban growth in the Detroit region. *Geographical Analysis* 46 (1970) 234-240
17. Goodchild, M.F.: The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers* 94 (2004) 300-303
18. Reilly, W.J.: *The Law of Retail Gravitation*. Knickerbocker, New York (1931)
19. Irwin, E.G., Geoghegan, J.: Theory, data, methods: developing spatially explicit economic models of land use change. *Agriculture, Ecosystems & Environment* 85 (2001) 7-23
20. Cheng, J., Masser, I.: Urban growth pattern modeling: a case study of Wuhan city, PR China. *Landscape and Urban Planning* 62 (2003) 199-217
21. Barredo, J.I., Demicheli, L.: Urban sustainability in developing countries' megacities: modelling and predicting future urban growth in Lagos. *Cities* 20 (2003) 297-310