The 18th Biennial Conference of International Society for Ecological Modelling

Accuracy assessments of land use change simulation based on Markov-cellular automata model

S.Q. Wang\textsuperscript{a*}, X.Q. Zheng\textsuperscript{a}, X.B. Zang\textsuperscript{b}

\textsuperscript{a}China University of Geosciences (Beijing), 29 Xueyuan Road, Beijing 100083, China
\textsuperscript{b}China University of Mining and Technology, D11 Xueyuan Road, Beijing 100083, China

Abstract

The potential of Markov-cellular automata model has been recognized and used widely in land use changes simulation and forecast, but previous researchers focused on analysis of future simulation result and land use change. Few thoughts have given to the analysis of modal accuracy and uncertainty and previous accuracy assessments have largely been conducted by simply calculating the overall value of selected accuracy metric. To increase the applied assessment value of Markov- cellular automata model, we explored the accuracy of simulation through the calculation of Kappa index for location and quantity. The case study area was Changping District, which is a rapidly growing area of Beijing. We classified the 1988 and 1995 TM image into six land use types in accordance with Chinese standards. Computed the transition suitable atlas based on different range and measurement units of the physical and socioeconomic factors. And then, we integrated these datasets to simulate the 2000 land use change map by Markov- cellular automata model. Lastly, we analyzed the simulation result and assess the accuracy of different cell size and neighbour size through various kappa index. The results of simulation showed that simulation accuracy of small cell size is better than the big cell size although it takes a substantial amount of time to run. However, the prediction accuracy of the model was quite stable with the neighbour sizes from 3*3 to 13*13. These results allowed us to understand what and how the factor including cell size and neighbour size affected simulation accuracy and how to select the best cell size for simulation.

© 2011 Published by Elsevier B.V. Selection and/or peer-review under responsibility of School of Environment, Beijing Normal University. Open access under CC BY-NC-ND license.

Keywords: Accuracy assessment, land use change simulation, Markov-cellular automata model;

* Corresponding author. Tel.: 86-10-82321807; fax: 86-10-82321807.
E-mail address: w_sq_2002@163.com.
1. Introduction

Land-use system is a natural, social and economic complex spatial dynamics with open, non-linear, comprehensive, random, hierarchical, dynamic and uncertainty of complex characteristics. Dynamic evolution of urban land use is highly complex affected by the natural, social, economic, cultural, political and legal and other factors. At present, urban land use changed dramatically under the common influence of physical geography and human environmental system, affecting all aspects of sustainable urban development, the problems brought to mankind.

Cellular Automata (CA) is a kind of discrete grid dynamic model, whose time, space and state are all discrete, and spatial interaction and causality on time are completely specified in terms of a local relation. CA is suitable to study complex spatial-temporal geographic system, especial for urban land use, and it has been an important tool and research focus for urban land use/cover change modeling. Although the potential of the Markov-cellular automata model has been recognized and used widely in land use changes simulation and forecast, most studies remain in the simulation results of the analysis, warning, application and cellular sensitivity analysis, lack of the modeling the process of modeling and simulation results lack the necessary inspection and uncertainty analysis. With the CA model more complex, integration and assimilation of remote sensing and geospatial data, quantitative of social and cultural drivers, these factors make the model simulation uncertainty. Without a lot of verification and evaluation for simulation results, it will hinder the development and improvement of the model for resources and environmental management. The Government departments and the public can not effectively access reliable predictive information and decision-makers and the public affected by a large number of redundant, mixed, ambiguous and misleading information. Neglect and lack of spatially explicit simulation model of space will not confirm the model simulation results of assurance[1].

The CA-Markov land use change model can be applicable for spatial land use simulation and land-cover reconstructions. Using satellite-derived land use/cover maps, Courage et al.[2] combined the Markov-cellular automata model to simulate land use changes in the Musana and Masembura CAs in Zimbabwe. Although this study represented an important contribution to land use modeling by the integration of biophysical and socioeconomic data into a spatially explicit Markov_cellular automata land use simulation model in a landscape, these simulation result lack of model validation and accuracy assessment. Some researchers in China used this robust approach to simulate the land use change and analysis the structure the land use in the future [3,4]. These studies generally used the map with fixed spatial resolution. Other authors have focused on the effects of the different parameters included in the model, such as neighborhood type and size, cell size, and both[5-7]. However, as a grid format, cellular automata lead to the existence of cell size sensitivity and would also be affected by cell size variation. These results showed that CA model responses are different depending on the spatial resolution, neighborhood size and type. Visual inspection of the cross-classification maps at all spatial resolution for rectangular and circular neighborhood types leads to the conclusion that different neighborhood types generate different simulated patterns for the same neighborhood sizes. And response relationship between cell size, neighborhood size and the simulation accuracy does not exist directly or absolutely [5].

The validity of the model results have been evaluated by comparing the KAPPA index of agreement for each category, spatial patterns of land use type and fractal parameter. KAPPA index can give summary static of agreement in terms of the proportion of the total number of pixels, but KAPPA index is too simple to express the spatial patterns and spatial distribution. Landscape ecology is one discipline where the ability to characterize spatial patterns is recognized to be essential, and many of the techniques routinely used in that field have been assembled in the FRAGSTATS software package[8].

Accuracy of analysis and simulation as land use/cover change an important part of the model has become the standard evaluation model and an important tool of simulation results for policy makers.
However, the problems of accuracy assessment still exist as following: single accuracy analysis method. Kappa Index, ROC curve, error analysis of multi-resolution and landscape index are the most widely used method for precision assessment, but the calibrate method is simpler. Kappa Index is limited to supply the overall accuracy analysis, due to spatial location, number and chance factors caused by the same rate and changes. Accuracy of the results of the landscape index can not be described quantitatively, lack of reference to uncertainty analysis of the image. In the conventional accuracy assessment methods, reference map is generally considered error-free. Direct use of land-use maps as the accuracy assessment data source, interpretation of aerial image or remote sensing image of the results will not be 100% correct, hence the confusion matrix method for accuracy testing is difficult to ensure the accuracy of simulation.

In the paper the study area was Changping District, which is a rapidly growing area of Beijing. We reclassified the 1988 and 1995 TM image into six land use types in accordance with Chinese standards. Computed the transition suitable atlas based on different range and measurement units of the physical and socioeconomic factors. And then, we integrated these datasets to simulate the 2000 land use change map by Markov- cellular automata model. Lastly, we analyze the simulation result and assess the accuracy of different cell size through various kappa index.

2. Study area

The case study area was Changping District located in the northwest of Beijing, which is a rapidly growing area of Beijing, China. The Latitude and longitude range is E116°50′17″~116°29′49″ and N40°2′18″~40°23′13″. Changping District covers an area of 1352 km2, of which is 60% of the mountain, 40% for the plain. The altitude varies from 800 m to 1000 m in the mountain, 30 m to 100 m in the plain.

Changping District has multiple land use types and the main types of land use are cropland, forest land and construction land. From 2004, land use changed dramatically under the common influence of physical geography and human environmental system. For example, paddy land gradually reduced or even disappear, ecological land use of closed forest land, artificial pasture, natural reserve, forest park increased drastically. At the same time mining area and isolated industrial area, rural settlement area increased slowly. The land use maps were the land use map of 1988 and 1995 from interpretation for TM remote sensing. In accordance with Chinese standards these data is classified into six categories mainly including cultivated land, woodland, meadow, water, settlement area and others. The accuracy of image reclassify was about 85%. To get the suitability maps for each land cover, the rural settlements, roads, rivers, transport and other relevant statistical data also be used in the study. Suitability maps for each land cover establish the inherent suitability of each pixel for each land cover type.

Fig. 1. (a) Beijing map; (b) Changping map
3. Simulation methods

Markov process is a kind of special random moving from one state to another state at each time step. Markov is appropriate to forecast the trend of land use system on the basis of the transition probability matrix, which records the probability that each land cover category will change to every other category. Markov analysis is a convenient tool for modeling land use change when changes and processes in the landscape are difficult to describe but it provides no sense of geography. As a kind of discrete grid dynamic model, time, space and state of Cellular Automata (CA) are all discrete, and spatial interaction and causality on time are completely specified in terms of a local relation. The cellular automata is suitable to study complex spatial-temporal geographic system, especially for regional land use, and it has been an important tool and research focus for regional land use/cover change modeling.

According to Markov model and CA model character, two models can be combined to obtain the trend and spatial structure of different land use. In land-use grid map, each pixel is a cellular, which state express land use type. With the support of GIS software, using the transformation matrix and the conditional probability image to determine the transformation of cellular state, the trend of land-use pattern changes can be simulated.

Simulation of land use change was undertaken using IDRISI32, which combined Cellular Automata and Markov Chain that adds an element of spatial contiguity as well as knowledge of the likely spatial distribution of transitions to Markov chain analysis. The implementation process using IDRISI was as follows:

- Computer transition probability matrix
  
  Land use distribution data of 1988 was used as the starting point for change simulation. Computing the conversion area between the various land use types as Markov state transition probability matrix elements. Using the Markov model to predict the change based purely on the state of land use in 1995 and on land use change in the preceding 8 years between 1988 and 1995.

- Produce the suitability maps
  
  According the developing character of Changping District the study select three criterion: the need of sustainable development, the law and statute of country and local government, city planning. The factors for suitability maps: slope, soil class, distance from settlements, roads and water and their influence on the distribution of land-cover classes were defined. We produce the suitability maps for each land cover according to such criteria to roads, water or existing land cover and establish the inherent suitability of each pixel for each land cover type. Fig 2 showed the different suitability maps of six land use category.

- Select CA filter
  
  A neighborhood filter was used for every modeled land cover type to emphasize the impact of corresponding suitability maps in the areas where that land use type already exits.

- Select simulate time and compute
  
  Based on the above results of transition probability matrix during 1988-1995 and the suitability maps of 1995, we select 5 as the CA cycles to predict the spatial distribution of land use in 2000.
Fig. 2. suitability maps of (a) cultivated land; (b) woodland; (c) meadow; (d) water; (e) settlement area; (g) others

Fig. 3. (a) 2000 simulated land use map (30 m resolution); (b) 2000 reality land use map

4. Accuracy assessment of simulation

Kappa index can be used as a measure of agreement between model predictions and reality or to determine if the values contained in an error matrix represent a result significantly better than random. In addition, the various Kappa statistics are excellent for comparing a map of “reality” versus some “alternative” map. The map of reality is any map that is considered to have high accuracy. Pontius[9] gave the many of the statistics allowing to separate quantification error versus location error, such as Kappa for no information, Kappa for location and for quantity. The formulas for the summary statistics are:

- Kappa for no information = \(K_{no} = (M_0M_L-N_0N_L)/(P_0P_L-N_0N_L)\)
- Kappa for location = \(K_{location} = (M_0M_L-M_0N_L)/(M_0P_L-M_0N_L)\)
- Kappa for quantity = \(K_{quantity} = (M_0M_L-N_0M_L)/(P_0M_L-N_0M_L)\)
- Kappa standard = \(K_{standard} = (M_0M_L-M_0N_L)/(P_0P_L-M_0N_L)\)
- Value of Perfect Information of Location = \(V_{PL} = M_0P_L-M_0M_L\)
- Value of Perfect Information of Quantity = \(V_{PQ} = P_0M_L-M_0M_L\)

This Kappa index will help us to know how these kappa statistics change with various resolution and cell size.
Table 1  Mathematical expressions for the quantity and location agreement between the reality map and a simulated map

<table>
<thead>
<tr>
<th>Ability to keep location</th>
<th>Ability to keep quantity</th>
<th>Perfect, $P_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfect, $P_0$</td>
<td>$\sum_{j \in n!} \min(</td>
<td>\delta_j(0,j)</td>
</tr>
<tr>
<td>Medium, $M_G$</td>
<td>$(1/J) + k_{\text{mean}} \cdot (N_G - 1/J)$</td>
<td>$P_e$</td>
</tr>
<tr>
<td>No, $N_i$</td>
<td>$1/J$</td>
<td>$\sum_{j \in n!} \delta_j(1,j)$</td>
</tr>
</tbody>
</table>

We rely on the Kappa statistic to quantify the validation of land use change model. First, we change the neighbor size(filter) to assess the impact of the neighbor. And then, we change the cell size(resolution) to validation of the CA model.

Fig. 4. Kappa for quantity and location( a. neighbour size; b. cell size)

From Fig 4 (a) it showed that prediction accuracy of the model was quite stable with less than 1% change in its accuracy between the filter of 3*3 to 13*13. Furthermore the trend of change of kappa index for no information, location and standard declined with the increase of filter size. From Fig 4 (b) we can see that the performance of the model was very good at the resolution ranged between 30 m and 120 m, where calculated Kappa index about 85%. It suggested that the best resolution to model the land use change in Changping is 30 m, which is fit to the remote sensing resolution of TM and spent more time to simulate the land use change. From the 240 m and coarser, the simulation accuracy started to decrease quite significantly, the prediction accuracy was only 81.4% and 76.8% respectively at 500 m and 1000 m. And trend of the different Kappa index change is same to the trend of the filter change.

Fig. 5. Components of agreement and disagreement at different (a)neighbour size; (b) resolution

Because the percent agreement due to chance tends to increases with resolution increase, to examine how the various Kappa index of agreement change with resolution is import. From Fig 5(a) the agreement due to chance is steady about 14.29% with the change of neighbor size. The Fig 5(b) showed
percentage of agreement and resolution for the agreement between the simulated map and reality map at different resolutions. The agreement due to chance of resolution had opposite change increasing from 14% to 16% with the resolution for 30 m and 1000 m. Fig 5(b) showed that agreement due to quantity and location tended to decline slightly as the resolution become coarser. However, the disagreement due to quantity and location tended to increase from 1.6% to 4.7% as one moved from the finest resolution to the coarsest resolution.

5. Conclusion

At different cell size and spatial resolution we found that the accuracy in the simulation of land use change in the Changping area is uncertain. Quantitative measurements for model volition, such as the variations of the kappa index of agreement, are very useful to independently account for accuracy in quantity and location of simulated land use change[1]. From the analysis of simulation accuracy we found cell size impacted less on the simulation results. Spatial resolution impacted more on the simulation results and the fine resolution is fit to simulation but spend more time. And sometime finest resolution images can’t obtain because of limited image resolution and can not be able to capture the morphology and spatial pattern of urban area.

From recent researchers’ study we can see that many researchers focused on the analysis of the effects of the different parameters included in the model, such as neighborhood type and size and resolution. However, the impact of these parameters remains uncertain and other factors such as driven element of land use change, the stochastic component remain unstudied. Further research in the accuracy assessment of simulation would involve the new validation methods, impact of stochastic component, and impact of remote sensing reclassify accuracy. In addition, the map of reality is any map that is considered to have high accuracy or absolutely correct, which is not rational to obtain the highest accuracy. In accuracy assessment of CA simulation, the “reality” map must be considered the accuracy it really obtained.

Acknowledgements

This work was supported by a grant from the Fundamental Research Funds for the Central Universities (NO.2011YYL070, NO.2011YYL067) and. the Special fund for scientific research on public causes of Ministry of Land and Resources (NO.20101018-3)

References


