Computers, Environment and Urban Systems 35 (2011) 289–296

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/compenvurbsys

An analysis of the effect of the stochastic component of urban cellular automata models

Andrés Manuel García, Inés Santé *, Rafael Crecente, David Miranda

Land Laboratory, Department of Agricultural and Forestry Engineering, University of Santiago de Compostela, Escuela Politécnica Superior, Campus universitario s/n, 27002 Lugo, Spain

ARTICLE INFO

Article history:
Received 25 May 2010
Received in revised form 29 November 2010
Accepted 30 November 2010
Available online 8 January 2011

Keywords:
Urban cellular automata
Stochastic cellular automata
Urban simulation
Monte Carlo method

ABSTRACT

Urban cellular automata models have proved useful tools in urban growth prediction because of their simplicity and their ability to reproduce complex emergent dynamics. Complex emergent dynamic systems involve processes that are difficult to predict, in which randomness plays a key role. In view of the fact that randomness is particularly relevant to complex processes, the aim of this paper is to analyze the sensitivity of the results of urban cellular automata models to the different methods used to incorporate the stochastic component in the models. The urban growth patterns obtained using different stochastic components are analyzed and compared using a number of spatial metrics. The results show that the differences observed in the simulated patterns are sufficiently relevant to justify the need for this type of analysis, which allows for the selection of the stochastic component that best suits the dynamics of the area.

1. Introduction

Cellular automata (CA) were first developed by S. Ulam and J. Von Neumann in the 1950s as discrete dynamic systems in which local interactions among components generated global changes in space and time. White (1998) defined a CA as "*a discrete cell space, together with a set of possible cell states and a set of transition rules that determine the state of each cell as a function of the states of all cells within a defined cell-space neighbourhood of the cell*". CA provided a useful tool for the study of complex systems, insofar as CA models allow for the generation of macro-scale complex patterns from simple, micro-scale rules. Early urban CA models were implemented as abstract models for the simulation of urban development, and were aimed at testing hypotheses derived from urban theories (e.g. Cecchini, 1996; Itami, 1988; Phipps & Langlois, 1997; Portugali & Benenson, 1995; Wu & Webster, 1998). These theoretical approaches served as the basis for the design of CA models aimed at simulating real-world urban development processes (e.g. Clarke, Hoppen, & Gaydos, 1997; White & Engelen, 1997; Xie, 1996).

Due to the complex characteristics of urban systems like emergency, path dependency or self-organization, urban CA results may be very sensitive to variations in their parameters (Manson, 2007; Messina et al., 2008). Other authors have focused on this issue by analyzing the effects of the different parameters included in the model, such as neighborhood type and size (Kocabas & Dragicevic, 2006), cell size (Dietzel & Clarke, 2004; Jantz & Goetz, 2005; Samat, 2006), both (Menard & Marceau, 2005), land-use classes (Dietzel & Clarke, 2006) or temporal resolution (Liu & Andersson, 2004). However, the impact of other components of CA models, among which the stochastic component, remains almost unstudied. Yeh and Li (2006) dealt with the effects of the stochastic perturbation in the predictability of the models, but the influence of the different methods used to introduce randomness in the results of the model and in the generated urban patterns has not been studied.

The incorporation of a stochastic component in urban CA models responds to the need to model the uncertainty associated with urban processes. Urban growth presents some unpredictable features that cannot be explained by deterministic variables (Yeh & Li, 2006). Accordingly, most urban CA models incorporate stochastic parameters to produce more realistic simulations. According to White and Engelen (1993), cities reflect social processes. Because social and biological processes occur in variable environments, their ability to evolve in order to adapt to the medium becomes essential. Without such ability, these processes cannot survive. Evolvability requires a system to be at the transition point between order and chaos, such that the system is not just chaotic or ordered, but complex. White and Engelen (1993) provide the example of the genes that compose the gene pool of a population. If such genes did not mutate, i.e., if the genes in the gene pool of a population remained in a constant state of equilibrium, the population would not be able to adapt to environmental changes. Conversely, if many

* Corresponding author. Tel.: +34 982252231x23324; fax: +34 982285926.
E-mail addresses: andresmanuel.garcia@usc.es (A.M. García), ines.sante@usc.es (I. Santé), rafael.crecente@usc.es (R. Crecente), david.miranda@usc.es (D. Miranda).

0198-9715/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2010.11.001

mutations occurred, the gene pool would deteriorate and the population would eventually disappear.

In urban CA models, 'mutations' in urban growth processes are simulated by introducing some randomness. The degree of randomness introduced in the models can be adjusted in different ways in order to obtain an amount of mutations that makes the system evolve between order and chaos, i.e., that makes the system complex. Therefore, the stochastic component is critical to address the uncertainties in complex geographical phenomena (Li & Yeh, 2004).

This paper analyzes the effects of the two most widely used approaches for considering 'mutations' in CA models on the generated urban patterns. The results of the analysis provide the information required to assess the advantages and shortcomings of each approach.

First, the most widely used methods for introducing and adjusting randomness are presented. Then, the methodology used in this paper to analyze the effects of the different methods on the simulations is described. In addition, the growth patterns obtained for the case study using the different methods are analyzed based on a number of spatial metrics. Finally, the results of the analysis and the conclusions drawn from such results are presented.

2. Introducing randomness in urban CA models

In regard to the methods used to introduce randomness in urban CA models, there are two principal types of models: (a) models that include a stochastic perturbation, such as the stochastic disturbance term proposed by White and Engelen (1993), in the transition rule (Cheng & Masser, 2004; He, Okada, Zhang, Shi, & Zhang, 2006; White & Engelen, 2000; Yang, Li, & Shi, 2008), and (b) models that introduce randomness when deciding which cells must change state, among which the models based on a Monte Carlo method that compares the transition probability for each cell with a random number, such that the cell will change state only if its probability is higher than the random number (Almeida et al., 2003; Jenerette & Wu, 2001; Li & Liu, 2006; Liu, Li, Liu, He, & Ai, 2008; Wu, 2002). Alternatively, the heuristic approach can be implemented by comparing growth rate and a random number in order to adjust the transitions to the amount of land use conversion observed in real urban systems (Li & Yeh, 2004; Liu et al., 2008).

The stochastic perturbation proposed by White and Engelen (1993) is computed from the following equation:

$$R = 1 + (-\ln(rand))^{\alpha} \tag{1}$$

where *rand* is a random number between 0 and 1, and α is a parameter that controls the size of the stochastic perturbation introduced in the model. High values of α imply that extreme values of *rand* are given more weight. Conversely, if the value of α is low, the extreme values of *rand* are given less weight. Therefore, the stochastic disturbance term will produce a larger or smaller stochastic perturbation in the transition potential for a cell depending on the value assigned to α.

Wu (2002) suggests that including a stochastic perturbation in the transition rule in order to force the transition of the cells with the highest potential for transition introduces a bias in the model, because the cells with lower transition potentials can also change state, but with a lower probability. For this reason, the author proposes the Monte Carlo approach as a more realistic method for selecting the cells that change state. Yet, the Monte Carlo approach has weaknesses, as it does not allow for the control of the degree of randomness or the total amount of simulated growth. To control both factors, Wu (2002) incorporates two equations.

As the ideal site changes with each iteration, the maximum potential value, *max*(P), is recalculated at each iteration using an exponential distance-decay function to transform the probability of site conversion, comparing its value with the probability of the best site:

$$P' = P * \exp[-\delta * (1 - P/\max(P))] \tag{2}$$

where P is the transition potential and δ is a dispersion parameter that controls the shape of the distance-decay function, so that the higher is the value of δ, the steeper the distance-decay gradient (Wu, 2002). Consequently δ has a function similar to the function of α in Eq. (1), i.e. δ controls the degree of randomness introduced in the model (Wu & Martin, 2002, pp. 1861), though in a different way. Eq. (1) scales the stochastic perturbation and therefore determines the degree of stochasticity in the calculation of P. However, in order to adjust the degree of randomness using the Monte Carlo approach, the transition potential must be scaled considering the maximum value of the probability as a benchmark. Therefore, higher values of δ will depress probability away from its maximum score such that greater discrimination between cells is obtained. Accordingly, the cells with higher values will be more differentiated from those with lower values, and there will be less probability that the latter will transition. Consequently, the degree of randomness will be lower.

Once the potentials have been scaled, Eq. (3) is used to control total urban growth:

$$P'' = \frac{P'}{\sum P'} \times N \tag{3}$$

where N is the number of transitions that must occur in each iteration of the model, which is determined exogenously to the model. Once P'' has been obtained, a random number between 0 and 1 is generated, such that the cell changes state if the value of P' is above the random number generated. Otherwise, the cell will not change state.

In view of the key role of randomness in achieving the complexity of urban dynamic systems, and considering the different methods that can be used to introduce randomness in urban CA models, among which the two methods explained above, we have analyzed both methods in order to determine whether significant differences are found in the results of an urban CA model when using either method.

3. Methods

To better analyze the effects of the stochastic component on the output of the model, we used the simplest possible urban CA model. The model proposed by Wu (2002) is simple and easy to calibrate, and allows for the use of the two methods for introducing randomness that are analyzed above. For this reason, we used the model proposed by Wu (2002) as a basis for the analysis.

The model developed by Wu (2002) calculates the transition potential from the following equation:

$$P = \frac{p * n}{8} \tag{4}$$

where n is the number of urban cells in the neighborhood (the model uses a neighborhood of 3×3 cells), which, divided by 8 (the number of cells in the neighborhood excluding the central cell), yields the probability of development for a cell as a function of the neighborhood; and p is the probability of development for a cell, calculated as a function of the considered variables using logistic regression. To keep the model as simple as possible, only three variables were used: slopes, distance to roads and distance to the center of an urban core.