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.

Logistic regression is an statistical method used to determine the probability of a dependent variable to take values of 1 or 0, which correspond in this case to urban and non-urban land uses, respectively, as a function of a number of independent variables (in this case slope, distance to roads and distance to the center of Ribadeo). Such a probability is calculated from the following equation:

$$p(y=1|X)=\frac{\exp(\sum BX)}{1+\exp(\sum BX)b} \quad (5)$$

where p is the probability for a dependent variable to take a value of 1 (urban), X is the vector of independent variables, and B is the vector of the regression parameters.

The stochastic component can be introduced in the model using two approaches: (a) by multiplying the transition potential P for each cell by the stochastic perturbation R, calculated from Eq. (1), and selecting the cells with the highest resulting transition potential P' to be converted to urban; or (b) by applying the Monte Carlo approach using Eqs. (2) and (3).

The comparison of approaches (a) and (b) as described above reveals a problem: if a cell does not have any urban neighbor, the transition potential P, calculated from Eq. (4), is zero. This means that the model will not be able to generate spontaneous growth and, consequently, it will not be able to simulate dispersed patterns. To solve this problem, the following equation is used:

$$P=p\times\left(\frac{n}{8}+1\right) \quad (6)$$

In brief, the probability of development is calculated using Eq. (6). If randomness is introduced in the model using the stochastic perturbation, such a probability of development is multiplied by the value obtained from Eq. (1) to calculate P', and the cells with the highest values of P' are selected to change state. The number of cells selected must be equivalent to the growth expected at each iteration of the model. In this paper, we have considered that urban growth at each iteration equals the land area developed during the period used to calibrate the model, divided by the number of years comprised in such period. If randomness is introduced in the model by using the Monte Carlo approach, P is determined first, and then Eqs. (2) and (3) are used to calculate P''. Once P'' has been calculated, the cells that are converted to urban are selected by comparing the value of P'' with a random number between 0 and 1.

4. Case study

The models described in the previous section were applied to the urban settlement of Ribadeo, NW Spain, and to the four surrounding parishes: Vilaselán, Vilaframil, Piñeira and Obe (Fig. 1). Three land use maps produced in 1978, 1995 and 2007, and derived from aerial photo interpretation were used. The maps of 1978 and 1995 correspond to the only historical aerial photographs available in this area, and therefore to the only source of information useful to obtain land use maps at the scale required in this study. The maps produced in 1978 and 1995 were used to calibrate the model, and the map produced in 2007 was used to compare the simulated urban patterns with the real ones. To obtain the variables considered to calculate the probability of development (slope, distance to roads, and distance to Ribadeo town center), we used a digital terrain model and a number of road maps available, which were corrected and updated using aerial photographs.

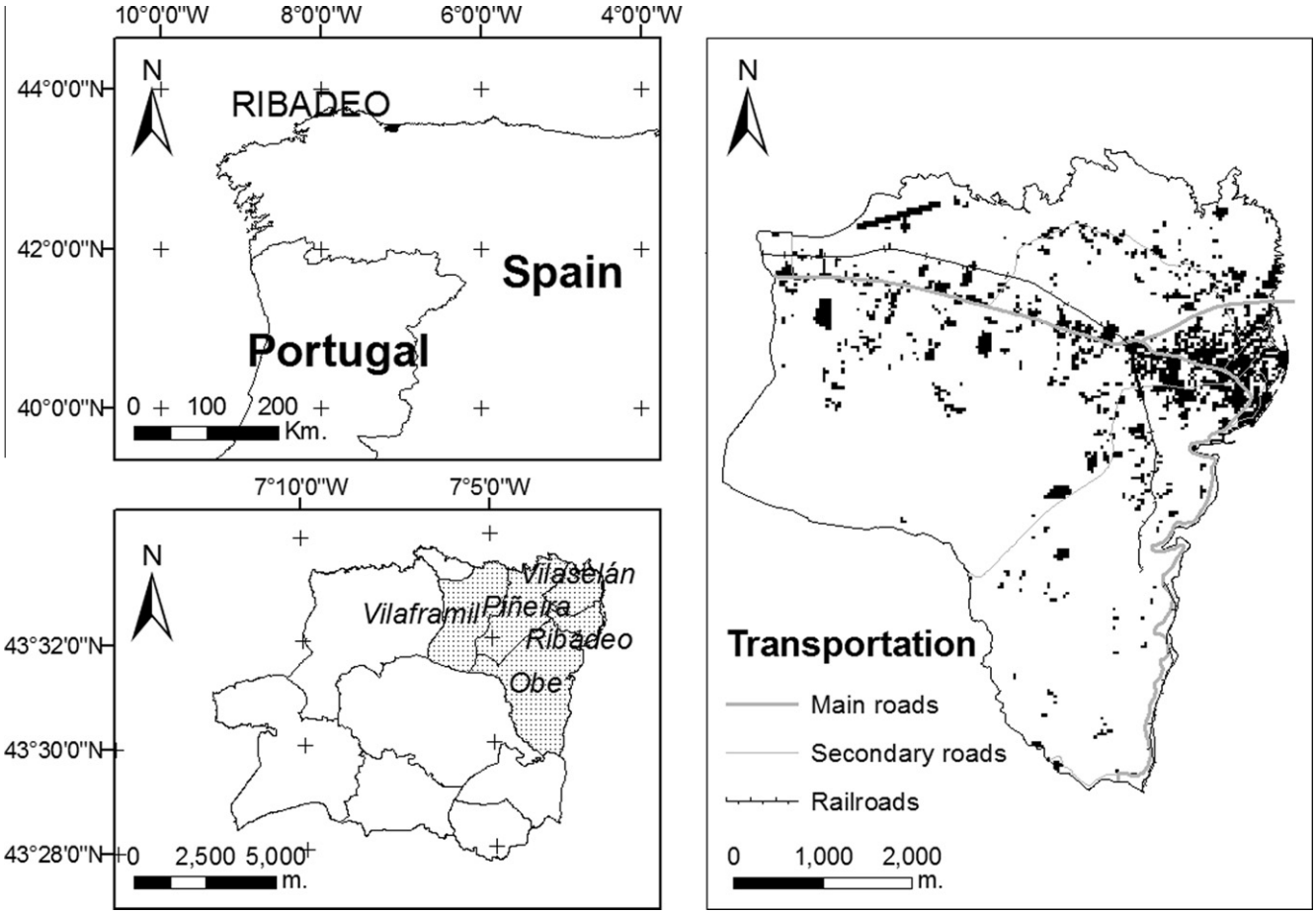

Fig. 1. Location of the study area and urban growth map of 2007.

The results of the two models were compared in order to analyze the differences between the growth patterns obtained based on spatial metrics. Spatial metrics were first developed in the field of landscape ecology to study landscape diversity and complexity. Later, spatial metrics were used by a number of authors to analyze and characterize urban growth patterns in order to find links between the patterns and the dynamics that generated such patterns (Dietzel, Herold, Hemphill, & Clarke, 2005; Herold, Goldstein, & Clarke, 2003; Lagarias, 2007; Seto & Fragkias, 2005; Thomas, Frankhauser, & Biemacki, 2008).

By introducing randomness in CA models, complex patterns can be obtained and dispersed growth patterns emerge. As suggested by White and Engelen (1993), the stochastic effects are stronger in the early stages of urban growth because growth is more dispersed. As the city grows, the vacant zones among urban patches are filled and nucleate. In this stage, stochastic effects are weaker. Next, the city continues to evolve and expands outwards, generating new dispersed growth, which will be absorbed by the growth of the city core. Thus, the degree of stochasticity is related to the complexity and compactness of the generated urban patterns, and these spatial characteristics can be analyzed by using spatial metrics.

Spatial metrics are criticized as not to be the better tools to validate the results of a model (Brown, Page, Riolo, Zellner, & Rand, 2005). Messina et al. (2008) and Manson (2007) suggest that inferring complex processes from the patterns produced by spatial metrics may have many shortcomings; the same complex process may generate several different patterns or different processes may produce the same pattern due to inherent characteristics of complexity. However, the main aim of this study is to test how the two considered methods introduce randomness in the simulation, not to determine whether the models simulate accurately the urban patterns. Randomness is a characteristic of complex systems. As described above, the level of stochasticity in urban CA determines the compactness and complexity of simulated urban patterns. Therefore spatial metrics that provide a measure of compactness and complexity are suitable to calibrate the level of randomness introduced by each approach.

A number of metrics can be used to measure complexity and compactness. Many of them are commonly used to validate the results of urban CA models. Santé, García, Miranda, and Crecente (2010) reviewed 33 urban CA models, 13 of which used spatial metrics to validate their results. Among these 13 models, 46% used the fractal dimension, 38% used the shape index and an edge density-based index, 31% used the Moran's I index, 23% used the number of patches and the contagion index, the 15% the mean Euclidean nearest neighbor and the largest patch index, whereas the mean patch area, the splitting index, the Simpson's index, and the Spatial Pattern Measure were used in only one of the models.

Taking into account that no single metric can capture complex urban patterns (Seto & Fragkias, 2005), the selection of the metrics was based on the most widely used metrics in previous research for validation of urban CA models, and on their ability to quantify the complexity and the compactness (or dispersion) of urban spatial patterns, since both characteristics are related to the degree of stochasticity. Among the metrics used to measure structural complexity (fractal dimension, shape index, and edge density index), the fractal dimension was selected because the fractal structure of cities has been confirmed by many studies (Lagarias, 2007; Thomas et al., 2008; White, Engelen, & Uljee, 1997). To analyze the compactness or dispersion of the urban pattern, the number of patches, the mean patch area and the Euclidean mean nearest neighbor distance were used. The number of patches was selected because this is the most widely used metric to measure dispersion and it is well complemented by mean patch area and the Euclidean mean nearest neighbor distance. A high number of patches may indicate that the urban pattern is dispersed. Yet, it must be analyzed together with the mean patch area to check if the patches are large or small, and with the mean Euclidean nearest neighbor to check if the patches are close of far from one another.

Further explanations on the calculation and meaning of these metrics are given below:

- Number of patches (NP): number of patches for each land-use class.
- Mean patch area (AREA_MN): average area of all the patches corresponding to a land-use class, expressed in hectares.
- Area weighted mean patch fractal dimension (FRAC_AM). This index suggests the degree of complexity of the patches. The value of the index approaches 1 for patches with simple perimeters such as squares, and 2 for patches with highly complex and space-filling perimeters. Because FRAC_AM is an area-weighted index, the fractal index of largest patches will be given more weight when calculating the mean. Larger patches are weighted because the shape of smaller patches is highly dependent upon the spatial resolution of the images:

$$\frac{\sum_{j=1}^{n}\left(\frac{2\ln(0.25p_{ij})}{\ln a_{ij}}\right)a_{ij}}{n_i} \tag{7}$$

  where p_{ij} is the perimeter of each patch j in land-use class i, a_{ij} is the area of each patch j in land-use class i and n_i is the number of patches in land-use class i.
- Euclidean mean nearest neighbor distance (ENN_MN): this index calculates the Euclidean distance mean value in meters, to the nearest neighboring patch based on the shortest distance from cell center to cell center.

The above spatial metrics were calculated for the patterns generated using the two methods for introducing stochasticity in the models, stochastic perturbation R and the Monte Carlo approach, for different values of δ and α. Table 1 shows the results obtained for the spatial metrics considered.

As shown in Table 1, the values of the spatial indices show opposite trends for the two methods used. For the Monte Carlo approach, the increase in the value of δ causes a decrease in the number of patches, which suggests a more compact growth pattern. The reason for this is that as the value of δ increases, the exponential curve that scales the transition probability will be more skewed and similar values will be more distant to one another, therefore the Monte Carlo method will be able to differentiate better between cells with similar values, such that those with higher

Table 1
Results obtained for different values of δ using the two methods for introducing randomness in the model.

δ	NP	AREA_MN	FRAC_AM	ENN_MN
Monte Carlo				
1	460	0.55	1.17	110.47
3	339	0.71	1.18	110.36
5	285	0.86	1.18	115.08
7	248	0.99	1.18	116.90
9	220	1.09	1.18	114.23
α	NP	AREA_MN	FRAC_AM	ENN_MN
Stochastic perturbation				
1	204	1.19	1.18	114.15
3	343	0.71	1.18	105.50
5	417	0.58	1.16	107.02
7	496	0.49	1.14	105.73
9	518	0.47	1.14	106.75
	NP	AREA_MN	FRAC_AM	ENN_MN
Real data for 2007				
	232	1.08	1.18	106.91

values will always be selected for transition and more compact patterns are generated. Conversely, in the stochastic perturbation approach, the value of α acts directly on the stochastic perturbation, such that if the intermediate values of the probability of development for a cell are given less weight and extreme values are given more weight, more dispersed patterns are generated.

The comparison of both methods using different values of δ and α reveals that the spatial metrics of the generated growth patterns do not show remarkable differences between the two methods. However, it can be observed that the stochastic perturbation produces a wider range of values for all metrics and increases in the value of α produce higher variations of the results, specially for high values of α. This is particularly evident in metrics FRAC_AM and ENN_MN. This fact involves a higher complexity in the calibration of the stochastic perturbation method to get urban patterns with a complexity and dispersion similar to that of the real patterns.

After having compared the results obtained for both methods, we analyzed which of the methods for introducing randomness in the model yielded the values that were nearest to the real values. To this end, we compared the results of the models with the spatial metrics obtained from real maps produced in 2007 (Table 1). As shown in Table 1, the values that best fit the real values are those obtained using $\delta = 9$ for the Monte Carlo approach and $\alpha = 1$ for the stochastic perturbation approach.

Further tests were conducted by varying the values of δ and α, to check whether better results were obtained. The stochastic perturbation was very difficult to control; for a value of $\alpha = 2$ the obtained patterns were too dispersed and for values between 1 and 2 the results were not better that for $\alpha = 1$. With the Monte Carlo approach the best results were obtained with a value of $\delta = 7.5$, for which the NP and the AREA_MN were almost equal to the real values. Table 2 illustrates the superiority of the stochastic perturbation in simulating the degree of fractality and the superiority of the Monte Carlo approach in simulating patch dispersion.

To verify which method produced the most stable results, the coefficient of variation was used to analyze the variability in the spatial metrics obtained for the patterns from 100 simulations of each model. Low stochastic perturbations were used in order to avoid that variability in results could be attributed to a high degree

Table 2
Spatial metrics obtained for the real map and the simulated map using stochastic perturbation ($\alpha = 1$) and the Monte Carlo method ($\delta = 7.5$).

	NP	AREA_MN	FRAC_AM	ENN_MN
Real 2007	232	1.08	1.18	106.91
Simulated 2007 $\delta = 7.5$	231	1.04	1.16	116.21
Simulated 2007 $\alpha = 1$	204	1.19	1.18	114.15

Table 3
Coefficients of variance for the values of spatial metrics obtained from 100 simulations of the model using the Monte Carlo approach and the stochastic perturbation method.

	NP	AREA_MN	FRAC_AM	ENN_MN
CV (Monte Carlo $\delta = 9.9$)	0.024	0.026	0.005	0.018
CV (Stochastic Pert. $\alpha = 0.1$)	0	3.61E−16	1.16E−15	1.23E−16

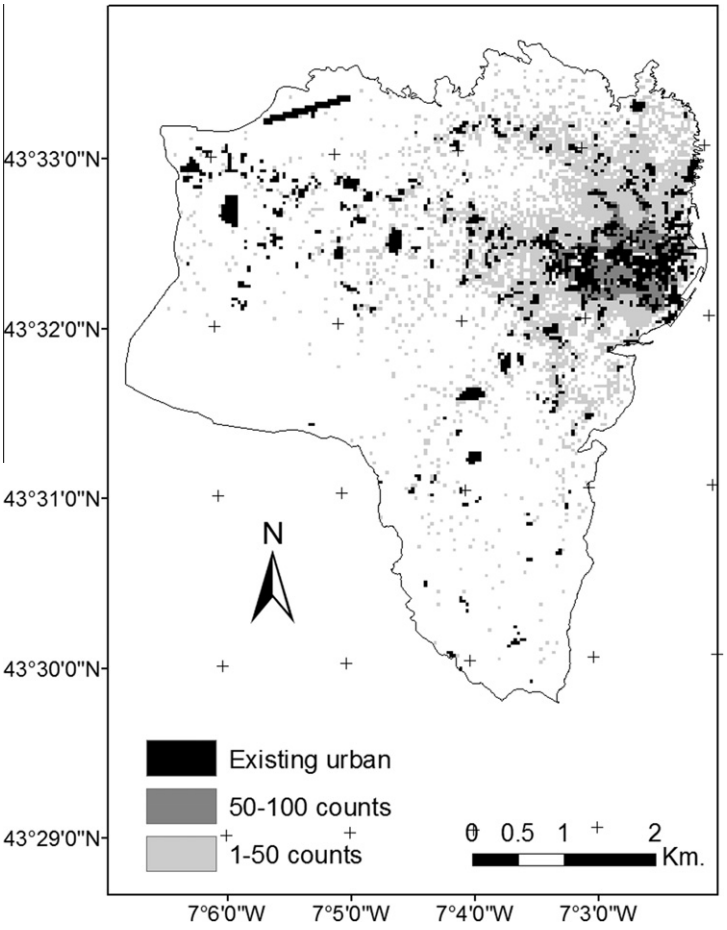

Fig. 2. Number of times that each cell was developed throughout the 100 simulations performed with the Monte Carlo approach ($\delta = 9.9$).

of randomness. As mentioned above, the two analyzed approaches introduce randomness in different ways. In the Monte Carlo approach, higher values of δ produce larger differences between the resulting probabilities. As a result, it will be more difficult that cells with lower probabilities may be converted. Thus, higher values of δ yield lower stochastic perturbations. Conversely, with the stochastic variable, higher values of α produce higher differences between high and low values of the stochastic variable. Because the stochastic variable scales directly the probability of change, the stochastic variable produces greater perturbations and therefore more randomness.

Bearing this in mind, a high value of δ ($\delta = 9.9$) was used for the Monte Carlo approach, since δ usually ranges from 1 to 10 (Wu, 2002), whereas a low value of α ($\alpha = 0.1$, since α cannot get values equal or lower than 0) was used for stochastic perturbation. Thus the observed variability is due to the method and not to the degree of randomness, since greater values of α and lower values of δ will only introduce higher levels of randomness and consequently more variability. One hundred of simulations using each approach were carried out and coefficients of variation were calculated for each landscape metric.

As shown in Table 3, the coefficients of variation for all the indices of the patterns modeled with the Monte Carlo approach are higher. These results show that, for the most common range of values of both indices, the Monte Carlo approach produces more variable patterns than stochastic perturbation.

The variability in the results can also be confirmed by defining a map of development probability based on the data obtained from the 100 simulations performed (Yeh & Li, 2006). Thus, the number of times that a cell is converted to urban in the 100 simulations is graphically represented. If the results are uniform, there will be a lot of cells that are converted to urban in most of the simulations, i.e. there will be a large number of cells with a high probability to be converted throughout the 100 simulations. If the results are very variable, there will be a lot of cells with a low probability to be converted to urban in a given simulation. As shown in Fig. 2, the results of the Monte Carlo approach are highly variable because many cells were converted to urban throughout the 100 simulations. The opposite is observed for stochastic perturbation (Fig. 3). According to these maps, the Monte Carlo approach introduces a lot of stochastic uncertainty (Brown et al., 2005), such that the results are more difficult to predict.

In addition, we analyzed the types of growth generated in each simulation and the area that corresponded to each type of growth. Xu et al. (2007) classified urban growth into three types: (a) spontaneous growth, in which the new urban patches have no direct connection with the existing urban patches, (b) edge-expansion growth, which refers to the newly developed urban area spreading out from the fringe of existing urban patches, and (c) infilling growth, which refers to converting the vacant lots existing between urban patches or within an urban patch to urban land. These authors have proposed an index to identify the three growth types, which is determined by dividing the length of the common boundary of a newly grown urban area and the pre-growth urban patches by the total perimeter of the newly grown area. If the value of the index is equivalent to zero, the patch is identified as spontaneous growth because it does not have any common boundaries with existing patches; if the index takes a value between 0 and 0.5, the patch corresponds to edge-expansion growth; and if it takes a value above 0.5, the patch is identified as infilling growth.

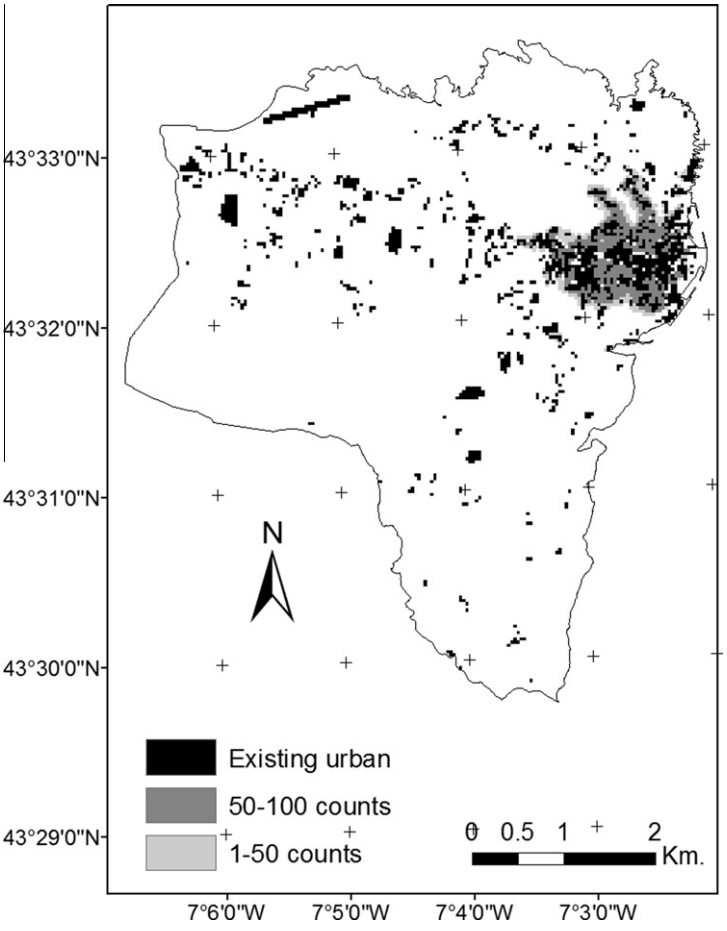

Fig. 3. Number of times that each cell was developed throughout the 100 simulations performed with the stochastic perturbation method ($\delta = 0.1$).

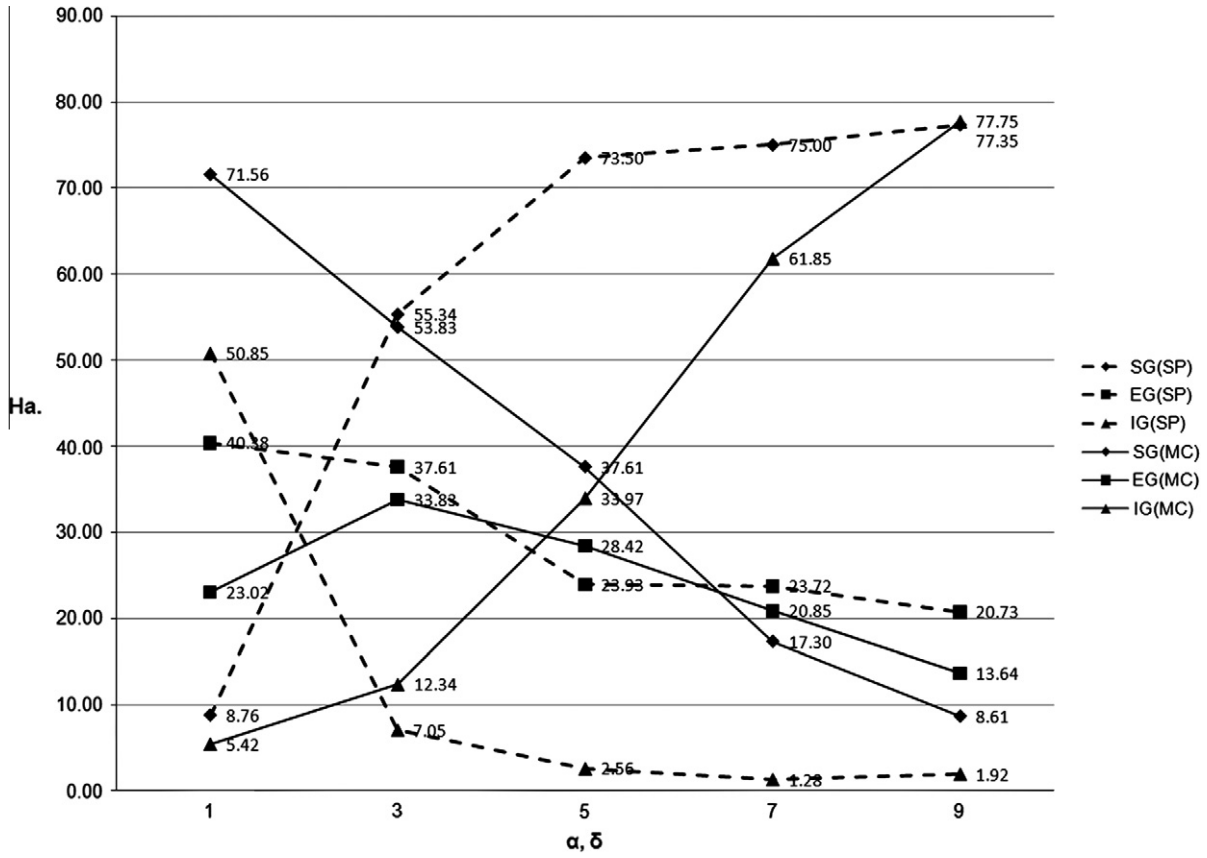

Fig. 4. Area (ha) for spontaneous growth (SG), edge-expansion growth (EG) and infilling growth (IG) obtained from the simulations performed with the Monte Carlo approach Monte Carlo (MC) and stochastic perturbation (SP) for different values of δ and α.

In our analysis, the index proposed by Xu et al. (2007) was applied to every patch, and the patches that corresponded to each type of growth were grouped in order to calculate the areas for the three types of growth obtained in each simulation using the Monte Carlo approach and the stochastic perturbation method (Fig. 4). If the patterns are scattered and complex, the amount of spontaneous growth will be high, whereas if they are compact and regular, infilling growth will be higher.

As shown in Fig. 4, in the simulations performed using the stochastic perturbation method the increase in spontaneous growth and the decrease in infilling growth are very abrupt for low values of α. It can be identified a point around $\alpha = 5$ where the increase of spontaneous growth and decrease of infilling growth becomes almost asymptotic. This is due to the logarithmic curve which is used to scale the stochastic perturbation. In the Monte Carlo approach, the curve used to control the degree of randomness is exponential. Consequently, when randomness increases, the increase in spontaneous growth and decrease in infilling growth becomes more linear.

In regard to the analysis of edge growth, in the stochastic perturbation approach, edge growth decreases with the increase in stochasticity. The Monte Carlo approach shows a similar behavior, but there comes a point where the trend is reversed ($\delta = 3$). This is related to the fact that a slight increase in stochasticity causes a decrease in infilling growth and a increase in edge growth and spontaneous growth. Conversely, when the increase in stochasticity is higher, both infilling and edge growth decrease, which leads to a further increase in spontaneous growth. In the stochastic perturbation approach, the logarithmic shape of the function causes a more abrupt increase of the stochasticity, which prevents this intermediate step, because it causes a direct decrease in edge and infilling growth that leads to an increase in spontaneous growth. This fact confirmed the difficulty of calibration of the stochastic perturbation.

5. Discussion and conclusions

Most urban CA models incorporate a stochastic component with a view to generating the complexity inherent in urban systems. Assuming that complexity is a state between order and chaos, and that achieving complexity requires that the right degree of randomness be introduced in the models, the analysis of the most suitable method to introduce the appropriate degree of randomness becomes essential. For this reason, this paper has explored the influence of different stochastic components on the spatial patterns of urban growth generated by a CA model. To this end, we compared the results obtained using two methods for introducing randomness in the model, based on stochastic perturbation and the Monte Carlo approach, respectively.

The authors who have used the Monte Carlo approach in their models argue that this approach provides a more realistic way of considering randomness and deciding which cells must change state, because the probability for a cell to change state is proportional to the transition potential of the cell. However, in the light of the results of this research, it can be affirmed that the Monte Carlo approach produces a greater degree of stochastic uncertainty (Brown et al., 2005) in the resulting patterns, yielding more scattered urban patterns and therefore a higher amount of spontaneous growth. This greater degree of stochastic uncertainty produces more different patterns in each run of the model. This makes more difficult to reproduce real growth patterns, so the model will yield worse results.

On the other hand, Monte Carlo method permits a better control of the degree of randomness introduced in the model due to the usage of an exponential curve to scale this factor. Thus more compact urban patterns are obtained. The logarithmic curve used in the stochastic perturbation produces more abrupt variations in the degree of randomness, i.e. produces higher changes in the degree of randomness for lower variations in the coefficient that controls the shape of the curve. Therefore, the fine-tuning of the degree of randomness is more difficult when using the stochastic perturbation, so the model calibration needed for the matching of simulated and real patterns will be more hardly achieved.

These two shortcomings could be overcome by taking advantage of the lower degree of stochastic uncertainty introduced by the stochastic perturbation and the better capability of the Monte Carlo approach to scale randomness. This may be achieved by scaling the stochastic perturbation with an exponential curve.

Finally, this paper points to the need to perform analyses of this kind, which allow for the selection and calibration of the stochastic component that best suits the dynamics of the analyzed area. Future research should be focused on methods to better control the degree of randomness, for example, conducting sensitivity analysis of how different kinds of curves can scale the randomness introduced in the models.

References

Almeida, C. M., Batty, M., Monteiro, A. M. V., Camara, G., Soares-Filho, B. S., Cerqueira, G. C., et al. (2003). Stochastic cellular automata modeling of urban land use dynamics: Empirical development and estimation. *Computers, Environment and Urban Systems, 27*, 481–509.

Brown, D. G., Page, S., Riolo, R., Zellner, M., & Rand, W. (2005). Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science, 19*(2), 153–174.

Cecchini, A. (1996). Urban modelling by means of cellular automata: Generalised urban automata with the help on-line (AUGH) model. *Environment and Planning B – Planning & Design, 23*, 721–732.

Cheng, J. Q., & Masser, I. (2004). Understanding spatial and temporal processes of urban growth: Cellular automata modelling. *Environment and Planning B – Planning & Design, 31*, 167–194.

Clarke, K. C., Hoppen, S., & Gaydos, L. (1997). A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. *Environment and Planning B – Planning & Design, 24*, 247–261.

Dietzel, C., & Clarke, K. C. (2004). Spatial differences in multi-resolution urban automata modeling. *Transactions in GIS, 8*, 479–492.

Dietzel, C., & Clarke, K. C. (2006). The effect of disaggregating land use categories in cellular automata during model calibration and forecasting. *Computers, Environment and Urban Systems, 30*, 78–101.

Dietzel, C., Herold, M., Hemphill, J. J., & Clarke, K. C. (2005). Spatio-temporal dynamics in California's central valley: Empirical links to urban theory. *International Journal of Geographical Information Science, 19*, 175–195.

He, C. Y., Okada, N., Zhang, Q. F., Shi, P. J., & Zhang, J. S. (2006). Modeling urban expansion scenarios by coupling cellular automata model and system dynamic model in Beijing, China. *Applied Geography, 26*, 323–345.

Herold, M., Goldstein, N. C., & Clarke, K. C. (2003). The spatiotemporal form of urban growth: Measurement, analysis and modeling. *Remote Sensing of Environment, 86*, 286–302.

Itami, R. M. (1988). Cellular worlds, models for dynamic conceptions of landscape. *Landscape Architecture, 78*(5), 52–57.

Jantz, C. A., & Goetz, S. J. (2005). Analysis of scale dependencies in an urban land-use-change model. *International Journal of Geographical Information Science, 19*, 217–241.

Jenerette, G. D., & Wu, J. G. (2001). Analysis and simulation of land-use change in the central Arizona–Phoenix region, USA. *Landscape Ecology, 16*, 611–626.

Kocabas, V., & Dragicevic, S. (2006). Assessing cellular automata model behaviour using a sensitivity analysis approach. *Computers, Environment and Urban Systems, 30*, 921–953.

Lagarias, A. (2007). Fractal analysis of the urbanization at the outskirts of the city: Models, measurement and explanation. *Cybergeo: European Journal of Geography* (on-line). Systèmes, Modélisation, Géostatistiques, document 391. <http://www.cybergeo.eu/index8902.html> (accessed 22.10.08).

Li, X., & Liu, X. (2006). An extended cellular automaton using case-based reasoning for simulating urban development in a large complex region. *International Journal of Geographical Information Science, 20*, 1109–1136.

Li, X., & Yeh, A. G.-O. (2004). Data mining of cellular automata's transition rules. *International Journal of Geographical Information Science, 18*, 723–744.

Liu, X. H., & Andersson, C. (2004). Assessing the impact of temporal dynamics on land-use change modeling. *Computers, Environment and Urban Systems, 28*, 107–124.

Liu, X. P., Li, X., Liu, L., He, J., & Ai, B. (2008). A bottom-up approach to discover transition rules of cellular automata using ant intelligence. *International Journal of Geographical Information Science, 22*, 1247–1269.

Manson, S. M. (2007). Challenges in evaluating models of geographic complexity. *Environment and Planning B – Planning & Design, 34*(2), 245–260.

Menard, A., & Marceau, D. J. (2005). Exploration of spatial scale sensitivity in geographic cellular automata. *Environment and Planning B – Planning & Design, 32*, 693–714.

Messina, J. P., Evans, T. P., Manson, S. M., Shortridge, A. M., Deadman, P. J., & Verburg, P. H. (2008). Complex systems models and the management of error and uncertainty. *Journal of Land Use Science, 3*(1), 11–25.

Phipps, M., & Langlois, A. (1997). Spatial dynamics, cellular automata, and parallel processing computers. *Environment and Planning B – Planning & Design, 24*, 193–204.

Portugali, J., & Benenson, I. (1995). Artificial planning experience by means of a heuristic cell-space model: Simulating international migration in the urban process. *Environment and Planning A, 27*, 1647–1665.

Samat, N. (2006). Characterizing the scale sensitivity of the cellular automata simulated urban growth: A case study of the Seberang Perai Region, Penang State, Malaysia. *Computers, Environment and Urban Systems, 30*, 906–920.

Santé, I., García, A. M., Miranda, D., & Crecente, R. (2010). Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning, 96*, 108–122.

Seto, K. C., & Fragkias, M. (2005). Quantifying spatiotemporal patterns of urban land-use change in four cities of China with time series landscape metrics. *Landscape Ecology, 20*, 871–888.

Thomas, I., Frankhauser, P., & Biemacki, C. (2008). The morphology of built-up landscapes in Wallonia (Belgium): A classification using fractal indices. *Landscape and Urban Planning, 84*(2), 99–115.

White, R. (1998). Cities and cellular automata. *Discrete dynamics in nature and society* (Vol. 2, pp. 111–125). India: Overseas Publishers Association.

White, R., & Engelen, G. (1993). Cellular-automata and fractal urban form: A cellular modeling approach to the evolution of urban land-use patterns. *Environment and Planning A, 25*, 1175–1199.

White, R., & Engelen, G. (1997). Cellular automata as the basis of integrated dynamic regional modelling. *Environment and Planning B – Planning & Design, 24*, 235–246.

White, R., & Engelen, G. (2000). High resolution modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems, 24*, 383–400.

White, R., Engelen, G., & Uljee, I. (1997). The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics. *Environment and Planning B – Planning & Design, 24*, 323–343.

Wu, F. (2002). Calibration of stochastic cellular automata: The application to rural–urban land conversions. *International Journal of Geographical Information Science, 16*, 795–818.

Wu, F., & Martin, D. (2002). Urban expansion simulation of Southeast England using population surface modelling and cellular automata. *Environment and Planning A, 34*, 1855–1876.

Wu, F., & Webster, C. J. (1998). Simulation of natural land use zoning under free-market and incremental development control regimes. *Computers, Environment and Urban Systems, 22*, 241–256.

Xie, Y. (1996). A generalized model for cellular urban dynamics. *Geographical Analysis, 28*, 350–373.

Xu, C., Liu, M. S., Zhang, C., An, S. Q., Yu, W., & Chen, J. M. (2007). The spatiotemporal dynamics of rapid urban growth in the Nanjing metropolitan region of China. *Landscape Ecology, 22*, 925–937.

Yang, Q., Li, X., & Shi, X. (2008). Cellular automata for simulating land use changes based on support vector machines. *Computers and Geosciences, 34*, 592–602.

Yeh, A. G.-O., & Li, X. (2006). Errors and uncertainties in urban cellular automata. *Computers, Environment and Urban Systems, 30*(1), 10–28.