APPLICATION OF A CELLULAR AUTOMATA MODEL TO THE METROPOLITAN AREA OF BARCELONA [(1)]

Nuno Norte Pinto
Lecturer
npinto@dec.uc.pt

António Pais Antunes
Professor
antunes@dec.uc.pt

R. Dr. Luís Reis Santos, Pólo II da Universidade, 3030-788 Coimbra, Portugal
Telephone +351239797106
Department of Civil Engineering
University of Coimbra
R. Dr. Luís Reis Santos, Pólo II da Universidade, 3030-788 Coimbra, Portugal

Josep Roca Cladera
Professor
josep.roca@upc.edu

Av. Diagonal 649, 4ª Planta, Barcelona, Spain
Telephone +34934016396
Center for land Policy and Valuation
Technical University of Catalonia
Av. Diagonal 649, 4ª Planta, Barcelona, Spain

Key words: cellular automata, urban growth, land use, Barcelona

(1) Esta ponencia se desarrolla como parte de las investigaciones realizadas el marco del Proyecto *EL PROCESO DE URBANIZACION EN LA COSTA MEDITERRANEA: ¿HACIA UN MODELO INSOSTENIBLE DE OCUPACION DEL SUELO? Un análisis retrospectivo (1956-2006) y prospectivo (2006-2026),* financiado por el Ministerio de Ciencia e Innovación, MICINN, en la Convocatoria de Concesión de Ayudas para la realización de proyectos de investigación, en el marco de algunos Programas Nacionales del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2004-2007. Convocatoria 2006. (SEJ2006-09630/GEOG).

Abstract

Cellular Automata (CA) models are among the most popular models for simulating land use change/growth in urban areas around the world and have experienced a significant development over the last twenty years. These models have their origins on the efforts of devising mathematical rules for biological systems and for evolution developed by mathematicians von Neumann and Stanislaw Ulam in the 1940s. Two main features made CA interesting for urban studies, to which they were introduced by Waldo Tobler in the early 1970s. First, their intrinsic spatiality, which is suitable for the simulation of a variety of geographic phenomena. Second, the possibility of simulating complex patterns of, for example, land use starting from a simple conceptual framework that includes the definition of a cell space, a neighborhood, and a finite set of transition rules applied to a finite set of cell states. These models have been developed for different urban contexts and are mainly based on the use of regular cells derived from remote sensing imagery. This is a simplification in the representation of urban areas: on the one hand, regular cells do not represent common urban form and, on the other hand, they do not held information of any type other than land use, obtained from automatic classification.

These issues suggested the consideration of irregular cells linked to reliable information, that is, census blocks. Census blocks are drawn considering the form of urban areas and they are the most reliable source of data on a wide variety of subjects, being a natural choice for the design of CA cells. We present in this paper an application of a CA model to simulate urban change in the Metropolitan Area of Barcelona, in a prospective analysis of 20 years. The model uses irregular cells designed considering census blocks.

We describe the main features of the model and the calibration process, as well as the simulation results. We also discuss some new features that are the core of our current research on CA.

1. Introduction

The rapid growth of urban areas raises important issues to modern societies in general and to the planning processes in particular. Sustainable development is now the key driver of urban growth imposing detailed scrutiny of trends, strategies, and public policies that are aimed to shape desirable futures. The complexity of these problems is such that there are no direct ways to achieve a solution nor these solutions are based on a single approach. On the one hand the consideration of different inputs – physical, sociological, economic, and historical, among several others – gives the comprehensive planning process the necessary tools to tackle complexity. On the other hand, this high level of complexity can be a potential weakness, because the problems become increasingly more complex as the natural evolution of societies takes place, demanding from planners new levels of commitment and accuracy in their research and practice. This knowledge can be used to develop models that aim to explain urban phenomena, retain knowledge from urban systems, and forecast planning scenarios. The study presented in this paper focus on the application of a cellular automata (CA) model to simulate urban change in the Metropolitan Area of Barcelona (MAB). CA models are a matter of intensive research since the late 1980s when information systems were making their entrance on the field of geography. After the introduction of CA to geography by Waldo Tobler (1979), a period of theoretical development took place and the first applications of CA to both theoretical instances and to real world case studies were made (Couclelis, 1985, 1987, White and Engelen, 1993, Batty and Xie, 1994). Since then, numerous variations and improvements were made and CA is widely used for simulating increasingly more complex problems (Wu and Webster, 1998, O'Sullivan, 2001a, Silva and Clarke, 2002, Barredo et al., 2003, Ward et al., 2003, Liu et al., 2008). However, the larger majority of these applications were made considering regular cells, often derived from remote sense imagery. There is only a small group of studies that developed CA models considering irregular cellular fabrics (Semboloni, 2000, O'Sullivan, 2001b, Pinto, 2006, Stevens et al., 2007). This particular characteristic is of great importance for the simulation of local scale problems, where the traditional regular cell, obtained from satellite images, may not represent well the spatial structure of the territory.

This paper is divided into three main parts. Section 2 is dedicated to presenting the CA model and its application to a series of test instances to assess model performance. Section 3 is dedicated to presenting the application of the model to the Metropolitan Area of Barcelona. Section 4 is dedicated to presenting the ongoing research on CA with some considerations about the development of a multi-scale CA model. Finally, some conclusions are drawn.

2. The Cellular Automata Model

The cellular automata model presented in this paper was designed to maintain the simplicity of the original concept. This section presents a brief overview about the different components of the model. For a detailed reading on CA and all their features see Pinto and Antunes (2007).

The cell structure is drawn from irregular cells. The aim is to use cells designed from real-world irregular census blocks. The model works with a set of six aggregate cell states (or land use classes): (1) urban low density (UL), (2) urban high density (UH), (3) industry (I), (4) non-urbanized urban areas (XU), (5) non-urbanized industrial areas (XI), and (6) areas where construction is highly restricted (R). Besides cell state, a set of other cell parameters are defined: cell accessibility, cell suitability, and neighborhood effect. The neighborhood effect simulates the spatial interaction between each pair of land uses: this interaction has a normalized value and ranges from -1 if two land uses repulse each other (e.g. cell states I and UH) to 0 if they not interact, or from 0 to +1 if they attract each other (e.g. cell states UL and R). This interaction is modeled by a linear function that decays with distance until it is no longer observed.

The neighborhood within which spatial interaction occurs is variable and is calibrated as a model parameter. Transition rules are applied through the consideration of a transition potential. The potential function gives a calibrated value of accessibility, land use suitability, and neighborhood effect. Land use demand is proportional to the increase of population, as well as to the variation of construction density. The assessment of model performance was made using contingency matrices and the corresponding k index (Couto, 2003). The comparison of the simulation map with the reference map through a measure of similarity is appropriate because it is oriented for the analysis of the entire territory as a distributed structure and not only as a centralized urban layout. But there are urban land uses (cell state R in the present classification) that were not considered in the changing dynamics.

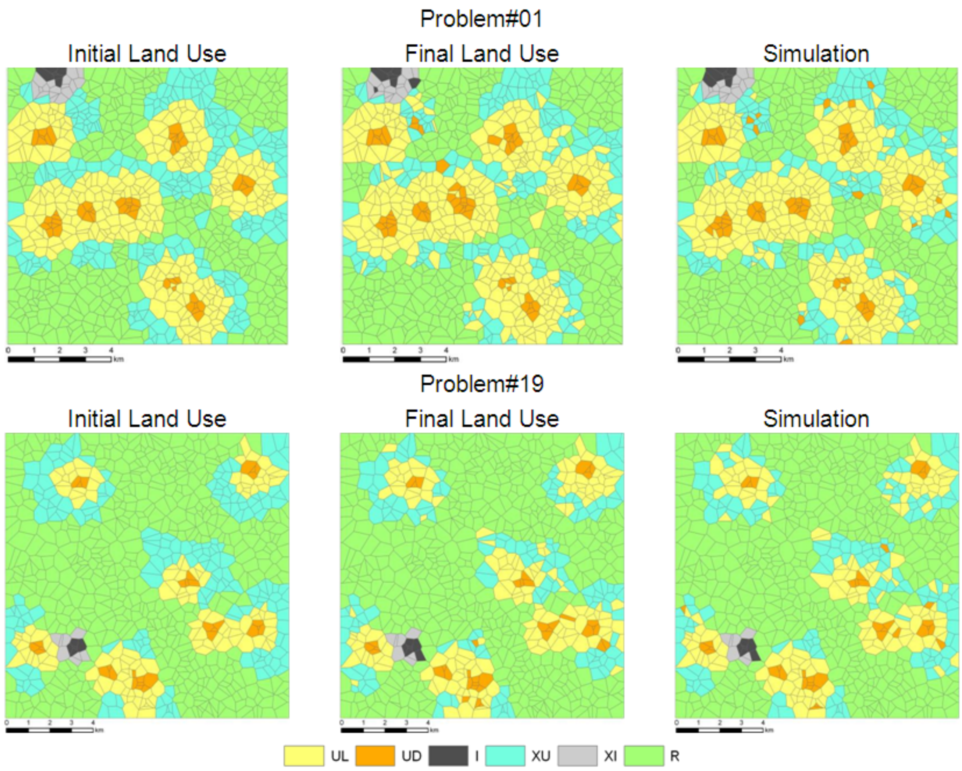
The consideration of the entire set of cell state for the calculation of the k index value (named k) would produce a distortion on its significance. To avoid this distortion, a modification of the k measure was considered, named kMod, accounting only the cell states that take part in the urban change dynamics. The calibration of the model was made through an optimization procedure called Particle Swarm with the goal of producing an extensive search of the set of calibration parameters that optimize the fitness function chosen for the model. This new type of optimization algorithm has given promising results for complex optimization problems. It is an optimization paradigm that simulates the movement of a group of individuals towards some goal, where the success of each individual influence its own searches and those of their peers (Kennedy, 1997, Parsopoulos and Vrahatis, 2002).

3. Model Performance

3.1 Test Instances

A set of 20 theoretical test instances was created with the aim of simulating municipalities not only in size but also in the total number of cells. These prototypes of spatial structures, depicted in Figure 1, were generated following the natural evolution of a territory, considering accessibility and a given probability of location regarding the distance to the centre of the settlements. There are two reference maps generated for each instance representing the initial and final land use configuration for a period of ten years. Then, a simulated land use map is generated by the CA model and compared with the reference map for the final year.

Figure 1. Two examples of test instances



3.2 Model Performance

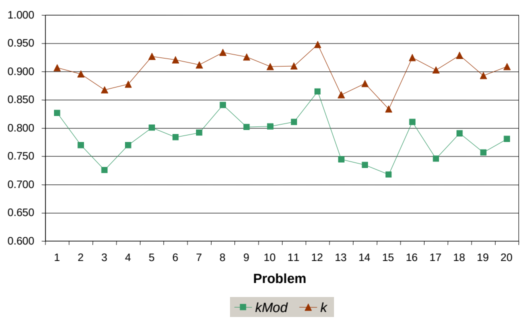
The ability of the CA model for simulating land use evolution was assessed for the 20 test instances. Global k_{Mod} results are depicted in Figure 2 (a). These results can be considered good for a simulation: 50 percent of the problems achieved a k_{Mod} around 0.800 or higher and 75 percent of them exceeded 0.750. k_{Mod} is a measure of agreement between modeled and reference maps which does not take into account inactive cell state. Figure 2 (a) also presents the variation of the absolute k measure for the set of test problems.

For 65 percent of the problems, the agreement exceeded 0.900 and 95 percent exceeded 0.850. These values are commonly accepted as good agreement between modeled and reference situations (Barredo et al., 2003). Note that the difference between the values of k and k_{Mod} is often higher than 0.100, justifying the choice of k_{Mod}. The model was also capable of simulating correctly the total area consumed by each cell state. The average relationship between total modeled area and total reference area for each cell state was of about 1.0% for UL cells, -5% for UH cells, and -1% for XU cells.

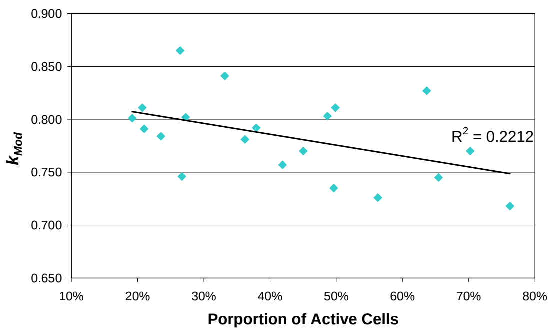
For industrial land uses this value increased due to the simplicity of the problem: only one cell out of two or three cells that have changed state produced a significant difference between modeled and reference areas. A tenuous relationship with a Pearson factor of 0.470 can be established between the proportion of active cells and the performance of the model: the smaller this proportion is, the better the model performs (Figure 2 b). The quality of the simulation is also assessed by the capacity of the model to identify areas of change.

The comparison of simulated and reference maps depicted on Figure 1 show that although the model was not able to correctly match the change for a large number of cells, it succeeded on choosing cells close to those ones due to a good calculation of the transition potential, identifying areas of change.

Figure 2. (a) Global k_{Mod} and k results for the set of test problems and (b) Relationship between the proportion of active cells and model performance



(a)



(b)

4. Case Study Application – Metropolitan Area of Barcelona

The model was applied to the Metropolitan Area of Barcelona using available censuses data for 1996 and 2001. This application consisted of making a first run to calibrate the model with reference data from censuses and a second run for a prospective analysis of land use growth considering the best set of calibration parameters obtained in the previous run.

This prospective analysis was made for two periods of ten years. The fact that calibration was made with a five years data set relates only to data availability; there is no implication of the use of ten years periods in the prospective analysis. Land use maps for reference years of 1996 and 2001 and for the calibration of the model are depicted in Figure 3. Land use maps for the prospective analysis are depicted in Figure 4.

Figure 3 Land use maps for the Metropolitan Area of Barcelona

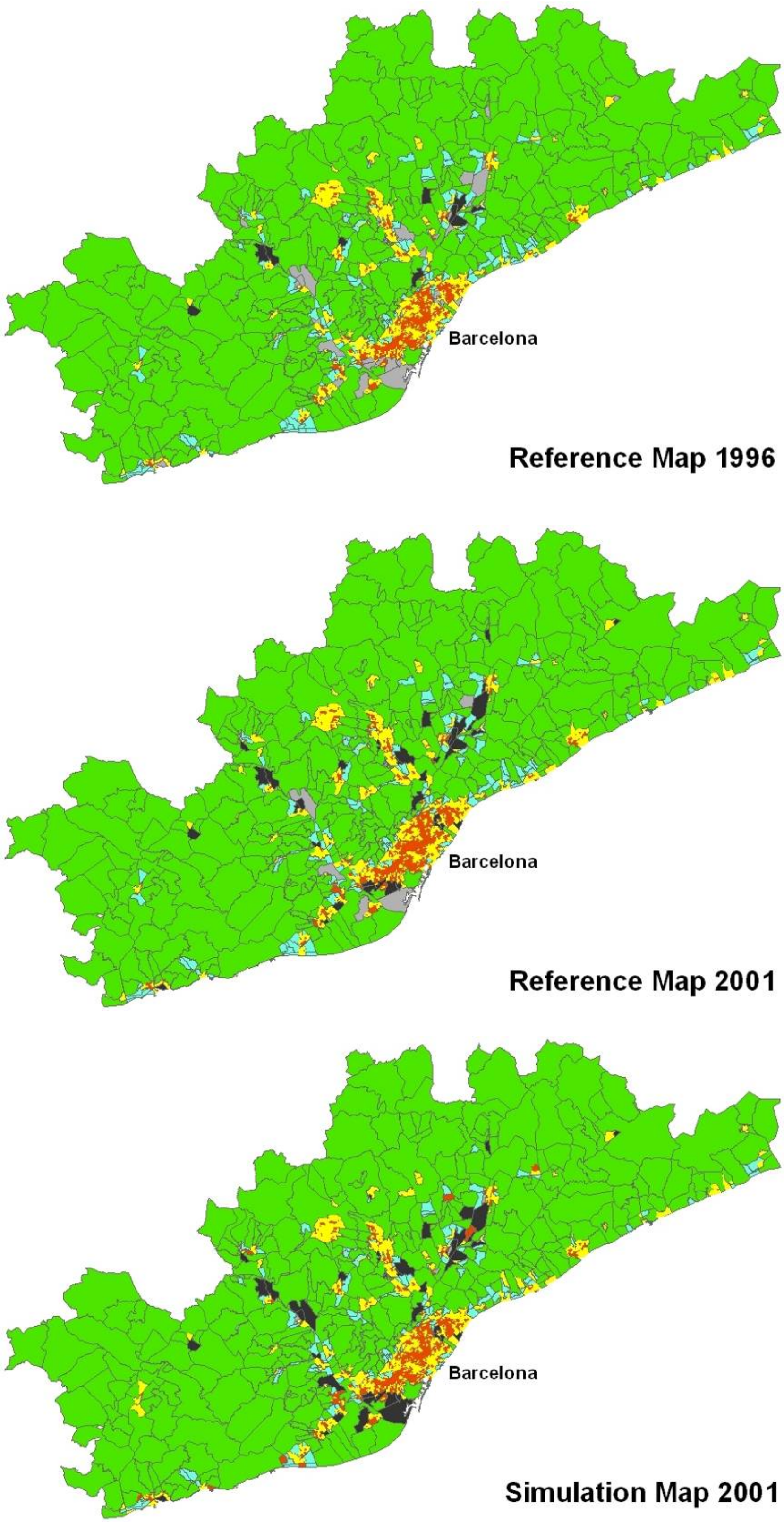
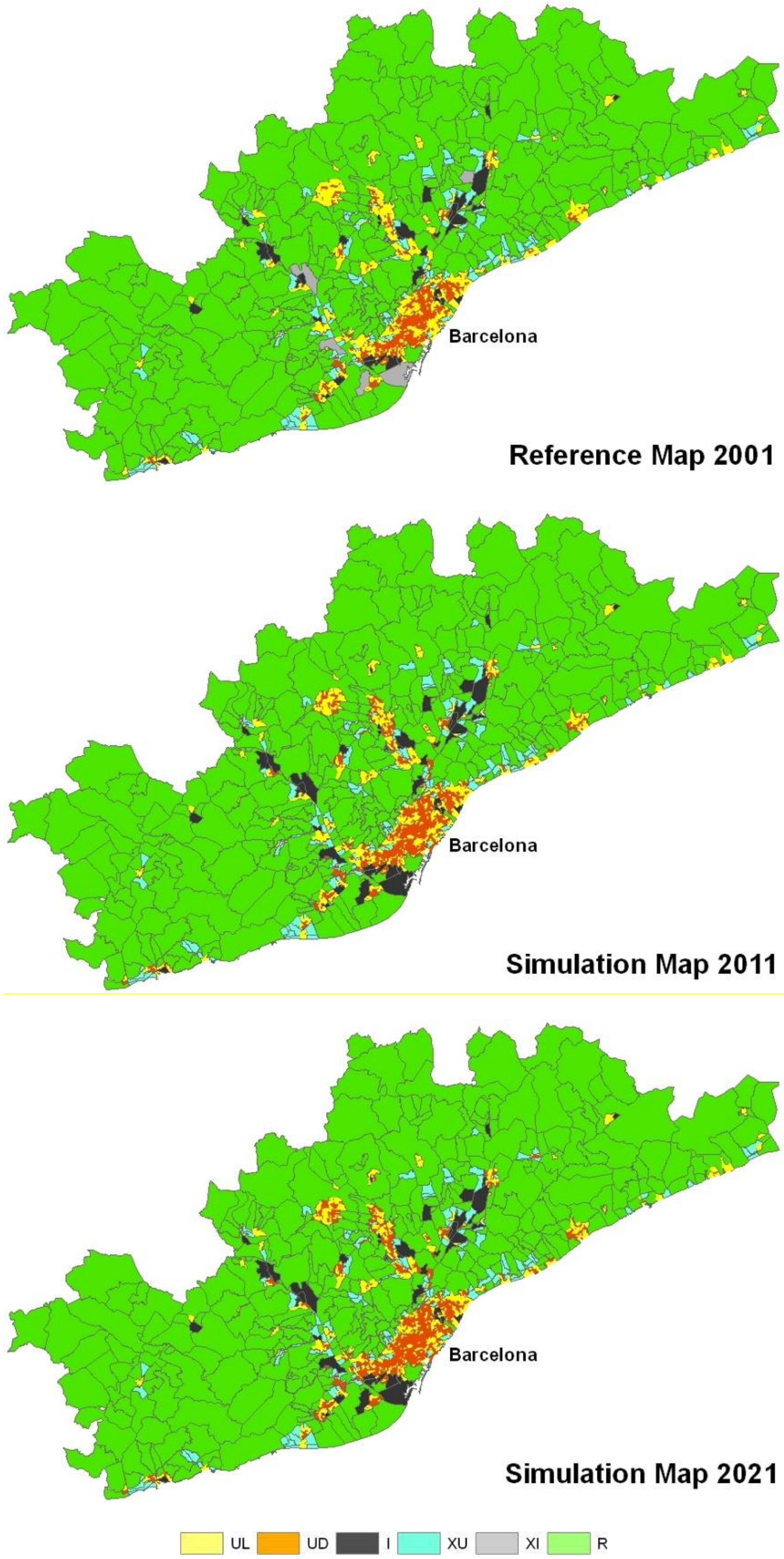


Figure 4 Prospective simulation of land use growth



Census blocks were used for cells in their original configuration. Land use was classified in the six cell states required by the model. The model achieved a high k_{Mod} value, 0.905. This high value is due to a relative low number of cells that changed state in simulation mainly because the model favored the densification of cells with larger areas in which larger amounts of population could be allocated. There are some problems regarding cell definition that may have influenced the simulation. On the one hand, they can represent both a traditional *Eixample* block and an entire municipality which imposes some problems on cell comparison. On the other hand, smaller municipalities may be represented by just one cell which means an important loss of representativity. This issue of cell representativity is one of the subjects of the ongoing research.

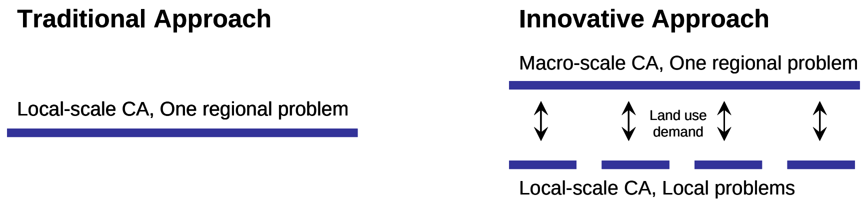
The model was used to prospectively simulate land use evolution on the Metropolitan Area of Barcelona for two consecutive periods of ten years, with a population growth rate of 5 percent. This simulation used the best set of calibration parameters determined during the calibration phase, starting from the reference situation in 2001. The model reinforced the trend of densification of the main urban areas.

4. Ongoing Research

Cellular automata has been under intensive research not only regarding its application but also considering their multiple components. The strength of the mathematical concept has been progressively relaxed in order to bring it closer to the complexity of urban phenomena. These relaxations need to be carefully addressed and they are part of the ongoing research on CA. Some of these issues are briefly introduced in this section.

Currently, research on CA is addressing the issue of scale. Different urban phenomena can be observed and modeled considering, on the one hand, cities as parts of larger regional systems and, on the other hand, cities themselves, where the driven forces of urban change depend on local variables. Therefore, the use of both regional and local scales is considered useful to correctly simulate a wide set of complex urban phenomena that occur on different spatial scales (or layers). The consideration of a regional scale of approach will improve the simulation of large scale urban phenomena. The evolution of population and employment observed during large periods and the flows between cities can be used to estimate degrees of relationship between them. Considering that these relationships depend largely on neighborhood conditions and that large scale cells usually have considerable and reliable data (municipalities, cities) it is clear that a CA-based model may be used to simulate urban evolution at a regional, therefore macro-scale. At this scale, the assessment of aggregate land use demand – for housing, industrial, or tertiary land uses – through population and employment growth can be more representative of urban growth than the disaggregate amount determined for each land use, which is traditionally used by other CA models. In opposition, at a city/neighborhood scale – the local scale – land use change results from the distribution of each land use (considered in a disaggregate level). In fact, land use demand can be estimated as the amount of land necessary for each land use considering population and employment growth. Then, a set of disaggregate land uses can be assigned to different cells in order to meet land use demand. This multi-scale approach is the consideration of two levels of simulation: the regional level, in which the goals are to simulate regional evolution and to assess land use demand from population and employment evolution through time; the local level, in which the goal is to assign different land uses to different locations (cells or parts of cells) in order to meet land uses demand considering local interaction. Figure 5 depicts the difference between a traditional single scale approach and the proposed multi-scale one. The multi-scale approach promotes the separation between land use allocation and regional trends, in opposition to the consideration different scale phenomena in the land use allocation process at the local scale.

Figure 5 Multi-scale Approach



The use of irregular cells implies the definition of a robust method for cell design, mainly related to data availability. In a multi-scale approach local scale cells should be as close to the urban structure as possible; at a regional scale, municipalities or other intermediate levels of administration can be considered as a natural option. The use of irregular cells is proposed as the basis for the simulation at regional and local scales.

Neighborhood is also a critical issue. Neighborhood is normally (if not entirely) considered by the formal concept that grounds the mathematical formulation of CA. This concept is based only on the consideration of a set of physical neighbors for each cell: these neighboring cells could be those which are directly connected to the cell considered or they could be the group of cells that are within a given range from that cell (which can be set by a radius or a number of cells in every direction). But this concept is far from being representative of real world dynamics. The concept of neighborhood must be able to reproduce how agents interact, considering both spatial and functional levels of interaction. Neighborhood must shift from the concept of a limited area to a larger and possibly disconnected part of the territory both at regional and local scales.

A multi-scale approach requires some attention on the definition of transition rules for both scales. At a regional level of analysis, it is important to notice that the goal is to simulate the macro interactions that are observed within a region. These interactions can be assessed using macro-scale indicators of population, employment, and commuting flows, for example. Transition rules and cell states will relate to this indicators and to measures of interaction. At a local scale, the goal is to simulate the distribution and growth of land use in a disaggregate manner. These phenomena depend on a series of factors related with land suitabilities, land use demand, accessibility, among many others. At this scale, transition rules will relate to measures of urban potential for each land use considering all the other land uses. The CA concept of cell state implies the consideration of a finite set of acceptable states for each cell. This concept is far from reality in urban phenomena. Homogeneity is very difficult to observe in any parcel of land of any dimension.

Land suitability is also an important issue to attend. At a regional level, the analysis is made with an aggregate perspective and land suitability must be taken into account as a measure of comparison for particular issues as general environmental quality or landscape and wildlife protection policies. At a local level, the comparison between land suitabilities for every land uses is determinant for the assessment of the demand for different and/or competitive land uses. Therefore, at this level of analysis it is imperative to develop a robust set of land suitability indicators.
Another important issue regards the calibration process. The simulation of urban phenomena is generally difficult to calibrate and validate due to its complexity. There are too many variables at stake and their behaviors are extremely complex and strongly interdependent. The calibration based on the Particle Swarm technique proved to be feasible. This technique has been successfully applied to problems where there is a large set of calibration parameters. However, in order to obtain even more reliable results, it is prudent to consider other calibration procedures based on other techniques, such as artificial neural networks or supported vector machines, combined with sensitivity analysis to their own calibration parameters.

Finally, models should be able to break with historical trends under given conditions. Common urban transformations are usually the result of single decisions localized in time as major urban renovation operations. These transformations are very difficult to capture by any kind model. CA work with a set of transition rules that are calibrated considering the historic evolution of a problem. The model should be able to generate new transition rules after observing a trend that can not be explained by the past evolution.

5. Conclusions

The results obtained for the set of test instances show promising possibilities of using cellular automata (CA) for modeling urban change based on irregular cells. The assessment of the components of the transition potential has to be improved in order to better simulate land use interaction and neighborhood effects. The concept of neighborhood needs to be closer to real urban structures rather than to its mathematical concept. The assessment of neighborhood relationships is another field that needs careful research. The case study application for the Metropolitan Area of Barcelona (MAB) proved that a CA model using irregular cells can be used to simulate land use change in large metropolitan/regional areas. There are, however, some issues regarding cell definition and representativity that should be taken into account.

The consideration of a single CA model for the entire MAB may be insufficient to correctly simulate and capture the driving forces of land use at this scale. This reinforces the idea of a multi-scale CA model to better capture and differentiate metropolitan/regional from local interactions.

Our current research is focused on this approach and there are two integrated models under development that operate at these two spatial levels. The comparison of the results obtained from the multi-scale CA model with the results presented in this paper for a single scale model may help to clarify the quality of the proposed approach.

Bibliography

Barredo, J., Kasanko, M., Mccormick, N. & Lavalle, C. (2003) Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landscape and Urban Planning, 64(3), 145-160.

Batty, M. & Xie, Y. (1994) From cells to cities. Environment & Planning B: Planning & Design, 21(Celebration Issue), 531-548.

Couclelis, H. (1985) Cellular worlds - a framework for modeling micro-macro dynamics. Environment and Planning A, 17(5), 585-596.

Couclelis, H. (1987) Cellular-dynamics - how individual decisions lead to global urban change. European Journal of Operational Research, 30(3), 344-346.

Couto, P. (2003) Assessing the accuracy of spatial simulation models. Ecological Modelling, 167(1-2), 181–198.

Kennedy, J. (1997) The particle swarm: Social adaptation of knowledge. 1997 IEEE International Conference on Evolutionary Computation (ICEC '97). INDIANAPOLIS, IN.

Liu, X., 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(11), 1247 - 1269.

O'sullivan, D. (2001a) Exploring spatial process dynamics using irregular cellular automaton models. Geographical Analysis, 33(1), 1-18.

O'sullivan, D. (2001b) Graph-based cellular automaton models of urban spatial processes. Bartlett School of Architecture and Planning. London, University College.

Parsopoulos, K. E. & Vrahatis, M. N. (2002) Recent approaches to global optimization problems through Particle Swarm Optimization. Natural Computing, 1235–306.

Pinto, N. N. (2006) A microsimulation approach for modelling the growth of small urban areas. Department of Civil Engineering. Porto, Portugal, School of Engineering of the University of Porto.

Pinto, N. N. & Antunes, A. P. (2007) Cellular Automata and Urban Studies: a Literature Survey. Architecture, City and Environment, 1(3), 368-399.

Semboloni, F. (2000) The growth of an urban cluster into a dynamic self-modifying spatial pattern. Environment and Planning B: Planning and Design, 27(4), 549-564.

Silva, E. & Clarke, K. C. (2002) Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal. Computers, Environment and Urban Systems, 26(6), 525-552.

Stevens, D., Dragicevic, S. & Rothley, K. (2007) iCity: A GIS-CA modelling tool for urban planning and decision making. Environmental Modelling & Software, 22(6), 761-773.

Tobler, W. (1979) Cellular geography. IN GALE, S. & OLSSON, G. (Eds.) Philosophy in Geography. Boston, D. Reidel.

Ward, D., Murray, A. & Phinn, S. (2003) Integrating spatial optimization and cellular automata for evaluating urban change. Annals of Regional Science, 37(1), 131-148.

White, R. & Engelen, G. (1993) Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environment and Planning A, 25(8), 1175-1199.

Wu, F. & Webster, C. J. (1998) Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environment and Planning B: Planning and Design, 25(1), 103-126.