Calibration of a cellular automata model with the particle swarm algorithm

Nuno Norte Pinto[1], António Pais Antunes[2], Josep Roca Cladera[3]

[1,2] Department of Civil Engineering, University of Coimbra
R. Luis Reis Santos, Pólo II da Universidade, 3030-788 Coimbra
Telephone: (+351)239797106
Fax: (+351)239797147
Email[1]: npinto@dec.uc.pt
Email[2]: antunes@dec.uc.pt

[3]Center for Land Policy and Valuation, Technical University of Catalonia
Av. Diagonal, 649, 4ª planta, 08028 Barcelona, Spain
Telephone: (+34)934016396
Fax: (+34)933330960
Email: josep.roca@upd.edu

1. Introduction

Cellular automata (CA) models have long been applied to simulate the evolution of urban areas. The large majority of CA models reported in the literature make use of regular cells derived from remote sensed images to represent land use and the use of irregular cells is scarce (Moreno et al., 2008, Stevens and Dragicevic, 2007). However, regular cells are not directly connected to the information that underlies the drivers of land use change – population, employment, or built up area indicators. We proposed a CA model that operates over a cell structure derived from irregular cells obtained from census blocks, which hold reliable data and can be easily classified for their land use (Norte Pinto and Pais Antunes, 2010).

Calibration plays a critical role in modelling because it connects reality to model representation. CA model calibration has been a subject of different approaches using different types of procedures, from sensitivity analysis to optimization-based methods. SLEUTH (Silva and Clarke, 2002) is uses both visual calibration and a brute force computational procedure to compare model and reference data. Li and Yeh (2001) coupled a CA model with an artificial neural network to calibrate it. Barredo et al. (2003) used basic sensitivity analysis to calibrate the weighting parameters for the spatial interactions between land uses.

2. Cellular automata model

The CA model has a simple structure that derives from the classical formulation of CA with the consideration of constrained land use demand, following the concept introduced by White and Engelen (1993). The model operates over an irregular cellular fabric obtained from census blocks. Cell states are classified into a finite set of aggregated classes of land use. Land use interactions take place within a variable neighborhood which distance value is determined through model calibration. Transition rules intend to incorporate planning regulations and simulate land use change based on a composite transition potential that takes into account cell accessibility, land use suitability, and neighborhood interactions within the cell neighborhood, calculated by the following expression

$$P_{i,s} = \left(\nu_P \times S_{i,s} + \chi_P \times A_i + \theta_P \times N_{i,s}\right) \times \xi, \forall i \in \mathrm{C}, s \in \mathrm{S}$$

where, for each cell i from the set of cells C, and for each state s from the set of states S, $P_{i,s}$ is the transition potential for state s of cell i, $S_{i,s}$ is the land use suitability value for state s of cell i, A_i is the accessibility value of cell i, $N_{i,s}$ is the neighborhood effect for state s of cell i considering its neighborhood V_i, ν_P is the calibration parameter for land use suitability, χ_P is the calibration parameter for accessibility, θ_P is the calibration parameter for the neighborhood effect, and ξ is the stochastic parameter. The model has 30 more calibration parameters which define the linear relationships of neighborhood effect interactions between each pair of land uses, generically depicted in Figure 1(a) for attraction and Figure 1(b) for repulsion. The time step can be defined by the user. Land use demand is determined through the evolution of population and employment densities over time. The flowchart for the CA model is depicted in Figure 2. Further details on the structure of the model can be found in Norte Pinto and Pais Antunes (2010).

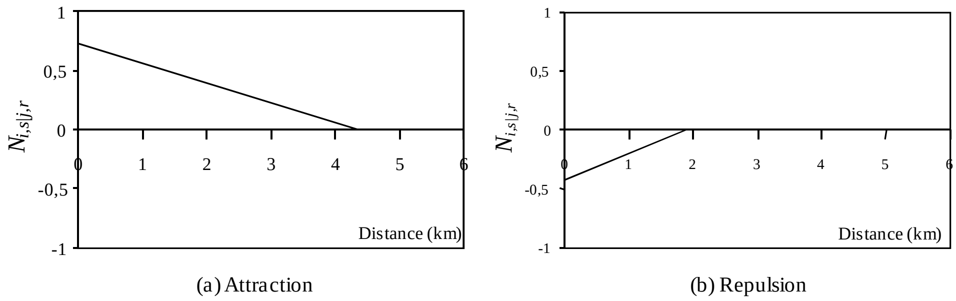

Figure 1. Generic neighborhood effect relationships

3. Calibration with particle swarm

The high number of calibration parameters indicates the use of an optimization procedure to ensure a good search of the solution space. The calibration of the CA model is processed though an optimization procedure that uses a fitness measure based on *kappa* index from contingency matrixes (Couto, 2003). We used a modified version of the traditional *kappa* (named k_{Mod}) to avoid the distortion that would have been produced if states that cannot take part in the urban dynamics – for example, agricultural or ecological reserve land – were considered. The inclusion of cells in this state would be misleading by producing a larger – though meaningless – agreement between simulation and reference maps.

The optimization algorithm chosen was the particle swarm (PS), which roots are in the simulation of social behaviors, in the study of the synchronized movement of bird flocks and fish schools (for further details please see Kennedy, 1997, and Parsopoulos and Vrahatis, 2002). This algorithm is suitable for dealing with a high number of dimensions (our calibration parameters) because it has a simple formulation which ensures that the complex interdependences between the parameters are taken into account in the calibration process. The algorithm makes use of a swarm of p particles (from a few to traditionally up to 120, but with no upper limit) will fly through the search space during *n* iterations. The larger the swarm is, the better the search space is searched. Each particle has *D* dimensions: in our CA model each calibration parameter is represented by a PS dimension. Hence, there will be 48 dimensions for each particle. The algorithm retains the position and the velocity of each particle in every iteration, calculating their new values considering the group leader and their individual best positions. The flowchart for the PS algorithm is depicted in Figure 2. Note that CA are an embedded process that is called as many times as the number of PS iterations multiplied by the number of particles.

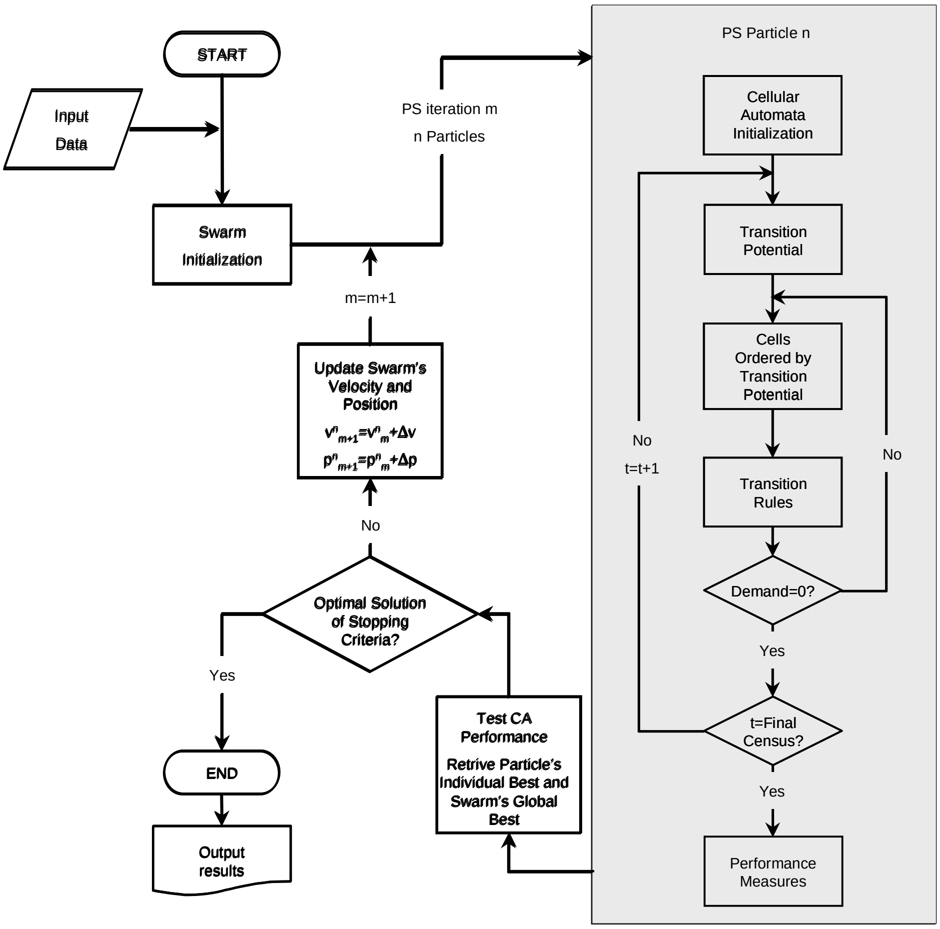

Figure 2. CA model (grey) and PS algorithm flowchart

4. Model results

The model was tested using a set of twenty test instances generated to simulate plausible spatial structures. These test instances have two reference land use maps (initial and final) for two moments in time, comprising information about population, employment and accessibility considering a road network. Three examples are depicted in Figure 3. Land use was classified with a set of aggregate cell states: urban low density (UL) and urban high density (UH), non-urbanized urban areas (XU); industry (I), non-urbanized industrial areas (XI); and areas where construction is highly restricted (R).

Global k_{Mod} results for the entire set of problems are depicted in Figure 4. These results can be considered good for a simulation process: 50 percent of the problems achieved a

k_{Mod} around 0.800 or higher and 75 percent of them exceeded 0.750. Figure 4 also presents the variation of the absolute *kappa* measure for the set of test problems. For 65 percent of the problems, the agreement exceeded 0.900 and 95 percent exceeded 0.850. Overall accuracy for the k_{Mod} measure also exceeded 0.850 for 75 percent of the cases. These values are commonly accepted as very good agreement between modeled and reference situations (Barredo et al., 2003).

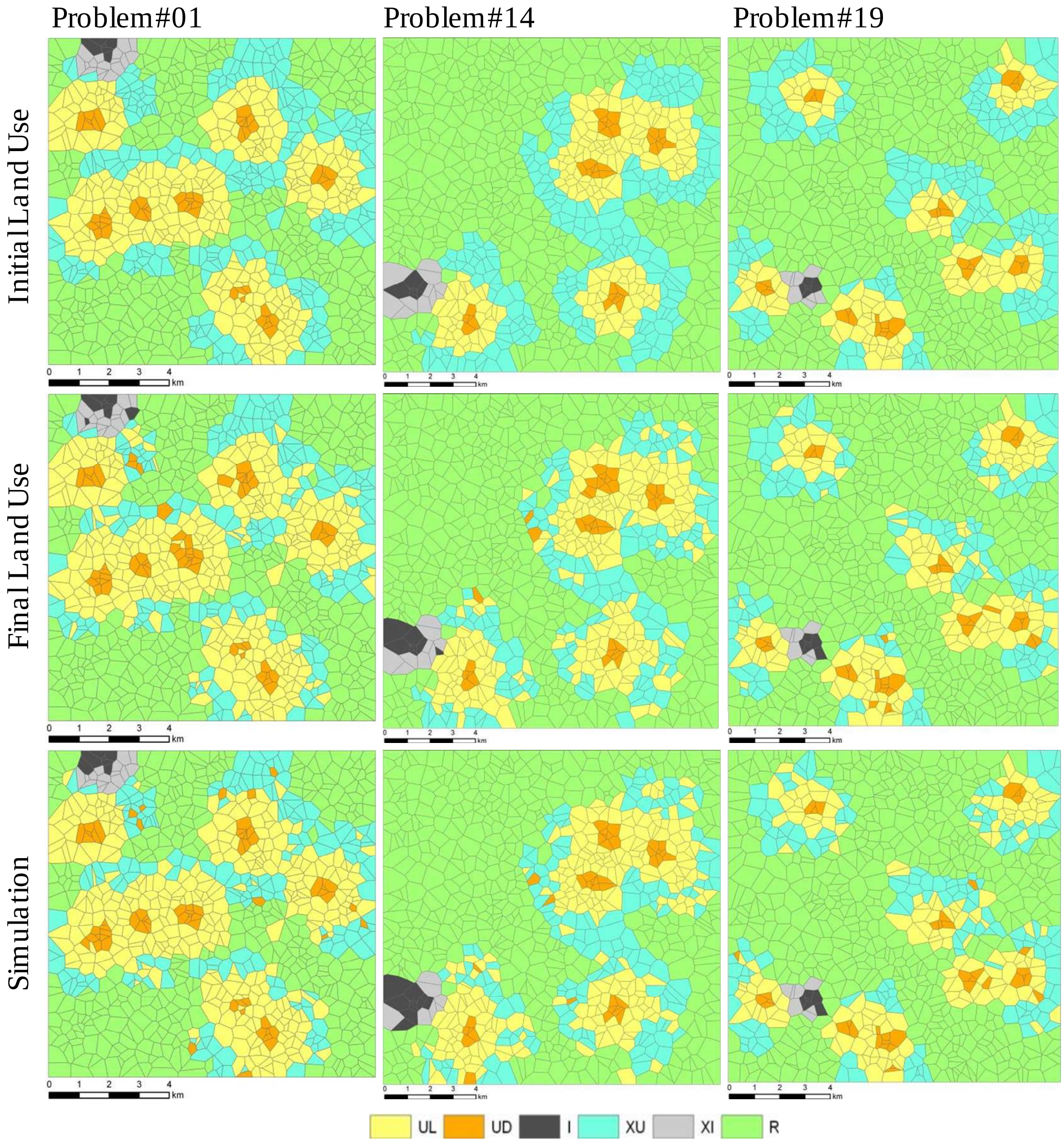

Figure 3. Three examples of test instances.

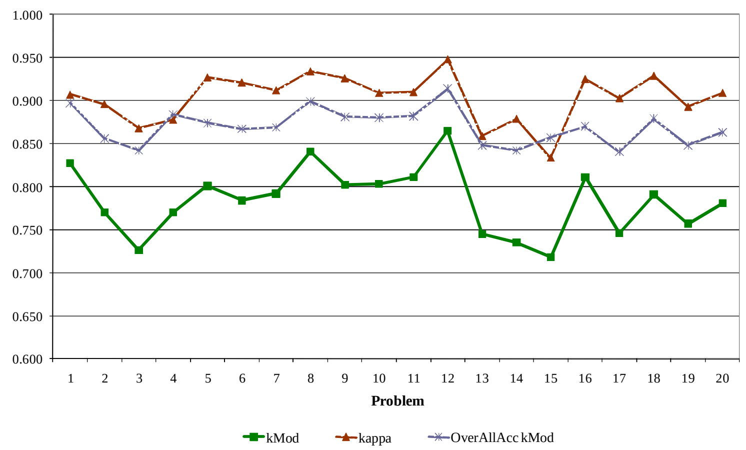

Figure 4. Global k_{Mod} and *kappa* results for the set of twenty test problems.

5. Concluding remarks

The results obtained for the set of test instances show that the use of the PS algorithm ensures an efficient search of good sets of calibration parameters for the CA model. The average value of the fitness measure k_{Mod} is high and is equal or higher than the values founded in the literature for other CA models. Current developments of our CA models – focusing on a multi-scale approach – also use the PS optimization for model calibration.

6. Acknowledgments

Nuno Pinto wishes to acknowledge the support received from Fundação para a Ciência e a Tecnologia under grant SFRH/BD/37465/2007.

7. References

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.

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. Proceedings of the 1997 IEEE International Conference on Evolutionary Computation (ICEC '97), Indianapolis, IN, 303-308.

Li, X. & Yeh, A. G. O. (2001) Calibration of cellular automata by using neural networks for the simulation of complex urban systems. Environment and Planning A, 33(8), 1445-1462.

Moreno, N., Ménard, A. & Marceau, D. J. (2008) VecGCA: a vector-based geographic cellular automata model allowing geometric transformations of objects. Environment and Planning B: Planning and Design, 35(4), 647-665.

Norte Pinto, N. & Pais Antunes, A. (2010) A cellular automata model based on irregular cells: application to small urban areas. Environment and Planning B: Planning and Design, 37(6), 1095-1114.

Parsopoulos, K. E. & Vrahatis, M. N. (2002) Recent approaches to global optimization problems through Particle Swarm Optimization. Natural Computing, 1, 235–306.

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. (2007) A GIS-based irregular cellular automata model of land-use change. Environment and Planning B: Planning and Design, 34(4), 708-724.

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.