Environmental Modelling & Software 69 (2015) 77–89

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Partial validation of cellular automata based model simulations of urban growth: An approach to assessing factor influence using spatial methods

Pablo Barreira González*, Francisco Aguilera-Benavente, Montserrat Gómez-Delgado

Departamento de Geología, Geografía y Medio Ambiente, Universidad de Alcalá, Facultad de Filosofía y Letras, c/Colegios n°2, 28801, Alcalá de Henares, Spain

A R T I C L E I N F O

Article history:
Received 13 November 2013
Received in revised form
3 March 2015
Accepted 13 March 2015
Available online

Keywords:
Cellular automata models
Urban growth
Partial validation
Map comparison
Frequency maps

A B S T R A C T

Cellular Automata (CA) based models have a high aptitude to reproduce the characteristics of urban processes and are useful to explore future scenarios. However, validation of their results poses a major challenge due to the absence of real future data with which to compare them. A partial validation applied to a CA-based model for the Madrid Region (Spain) is presented as a proposal for determining the influence of given factors on the results and testing their spatial variability. Several simulations of the model were computed by different combinations of factors, and results were compared using flexible map comparison methods in order to study spatial pattern matches and similarities between them. Main and total effects of these factors were calculated for each method, by applying a simplified Global Sensitivity Analysis approach. Frequency maps showing the most frequent cells with changed land use in the results were generated.

1. Introduction

Land use change processes, and particularly the expansion of urban areas, are among the most significant changes currently taking place in Europe (European Environment Agency, 2006; Kasanko et al., 2006), the U.S.A. (Buyantuyev and Wu, 2009; Wu et al., 2011) and other less developed countries all over the world (Barredo et al., 2004). These phenomena have substantial environmental and regional implications that require meticulous assessment (Berling-Wolff and Wu, 2004; Grimm et al., 2008).

In this context, the usefulness of urban simulation tools is evident: besides helping us to better understand the processes and driving forces of urban growth (Cheng and Masser, 2003), urban simulation models can reproduce past land use dynamics and simulate the possible evolution of these processes in the future (Hansen, 2010; Paegelow and Camacho-Olmedo, 2008; Sante et al., 2010). They can therefore be useful for assessing changes in cities and metropolitan growth patterns (Aguilera et al., 2010, 2011; Berling-Wolff and Wu, 2004; Li et al., 2008) and the possible effects of such growth on the landscape and region where they occur (Aguilera and Talavera, 2009; Forman, 1995; Mitsova et al., 2011).

Among the several types of simulation models, Cellular Automata (CA) has shown a high capacity for reproducing the main characteristics of urban expansion processes (Benenson and Torrens, 2004; Batty, 2007), such as complexity or self-similarity (Itami, 1994; Frankhauser, 1998; Torrens, 2000). Previous theoretical work carried out by White and Engelen (1993), Couclelis (1997), White et al. (1997), Benenson and Torrens (2004), Couclelis (2005), Batty (2007) among others, and also existing applications (Barredo et al., 2003; Petrov et al., 2009; Hewitt et al., 2014) have shown the potential of CA-based models to simulate urban growth and land use dynamics in different contexts. The *NASZ* model scheme, a CA-based model framework proposed by White et al. (1997), which includes four parameters (*neighbourhood*, *accessibility*, *suitability* and *zoning status)*, has become widely popular when simulating urban growth, leading to a considerable proliferation of simulation tools and studies based on this approach (see Sante et al., 2010; Triantakonstantis and Mountrakis, 2012 for a detailed review).

Urban growth simulation using CA-based models follows three main steps (Fig. 1). Once a model is implemented, a calibration process is necessary in order to adjust its internal parameters.

* Corresponding author. Tel.: +34 918856384; fax: +34 918854439.
E-mail addresses: pablo.barreira@edu.uah.es (P. Barreira González), f.aguilera@uah.es (F. Aguilera-Benavente), montserrat.gomez@uah.es (M. Gómez-Delgado).

http://dx.doi.org/10.1016/j.envsoft.2015.03.008
1364-8152/

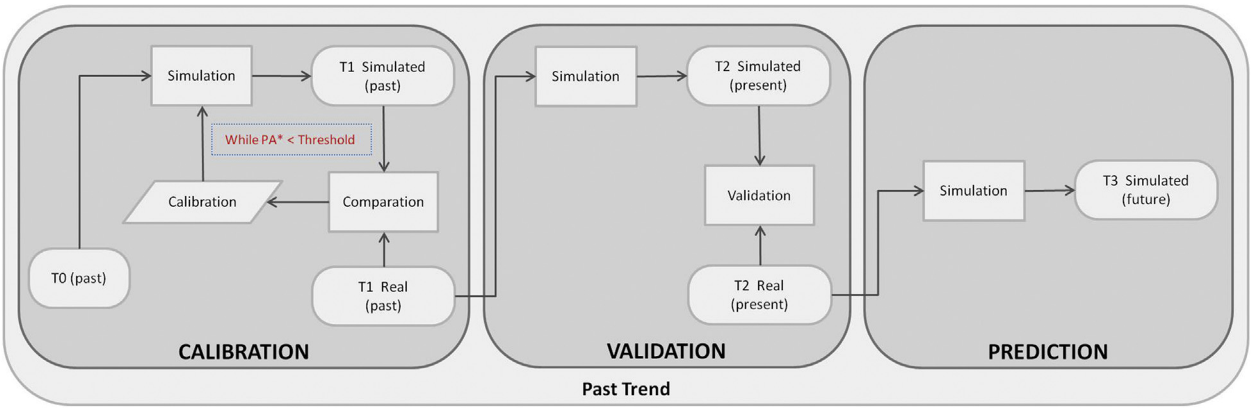

Fig. 1. Simulation process with CA-based models. Validation is performed before simulation. *PA = Percentage of Agreement.

Calibration is performed at a past period of time ($t_0 - t_1$) using historical land use data, and in essence the model has to reproduce a real dynamic within that period and compare its prediction with real data at time t_1. When the comparison results are optimal, it is assumed that the model can reproduce past dynamics that continue to the present (demonstrated through a validation process using real data), enabling subsequent trend simulations whereby present dynamics are projected into the future. Hence, calibration and validation process are closely related since they compare simulation results with real data (Engelen and White, 2008). Many of the studies developed, have therefore adopted a predictive approach (Paegelow and Camacho-Olmedo, 2008), embodied in a methodology similar to that shown in Fig. 1.

However, other studies (Berdoulay, 2009) have highlighted the problem, or at least the complexity, of adopting a purely predictive approach to simulations of urban dynamics based on a positivist conception of the urbanization process. Within this approach, urban CA models are developed (among other reasons) to improve urban planning. Therefore, simulating only past trend dynamics into the future may not be enough to inform planners about the possible futures that may happen. In addition to this, urban processes are subject to a significant degree of uncertainty related to political and administrative decision-making processes, unforeseen economic circumstances or emergence of new influential factors (Barredo et al., 2003).

Thus, we believe that the aim of these models should be to simulate several possible future situations under different circumstances. This alternative approach would employ urban growth simulation models in conjunction with future scenarios, not only to simulate past trends into the future (Fig. 1), but also to explore different development possibilities (Fig. 2), such as external exploratory scenarios (Borjeson et al., 2006).

Hence, after the model calibration process and comparison with real data to confirm that the model reproduces past dynamics, the model may be used to simulate different future scenarios. These scenarios may require changes to the model in order to simulate different possible futures. Therefore, the results of future simulations should be subjected to a second validation process (Fig. 3). The main problem at this stage (future time) lies in the fact that there are no real data to compare the model's simulation with. Therefore, the concept of validation in this context should be adapted and made flexible, since only partial validation is possible (Paegelow and Camacho-Olmedo, 2008).

There is no agreement within the scientific community about which set of procedures should comprise a validation process (see Tables 1–6 in Sante et al., 2010), even less in relation to a partial validation. However, stability and robustness determination of the results through a Sensitivity Analysis (SA) is considered one of the most important processes. In the field of spatially explicit models, we can find many examples in the literature that consider variations in the input variables and analyse quantitatively the differences in the outputs (e.g. Crosetto et al., 2001; Crosetto and Tarantola, 2001; Store and Kangas, 2001; Gómez-Delgado and Tarantola, 2006; Lilburne and Tarantola, 2009; Ligmann-Zielinska and Sun, 2010; Chu-Agor et al., 2011; Plata-Rocha et al., 2012). In recent years, there is an increasing effort to develop SA based methods that obtain spatially-explicit representation of the results of the SA (e.g. Brown et al., 2005; Chen et al., 2010, 2013; Ligmann-Zielinska and Jankowski, 2014, Marrel et al., 2011). These can quantify the influence that each parameter implemented in the model has on the results generated. Such SA-based methods have been widely applied to spatial models such as urban growth or land use change models (based on Multicriteria Evaluation, Agent-based models, etc). With regard to the spatial CA-based models, these methods have traditionally been applied only in order to determine the influence of a single parameter, such as neighbourhood or randomness (Garcia et al., 2011; Kocabas and Dragicevic, 2006; Ménard and Marceau, 2005; Pan et al., 2010), or a given characteristic of the data used, such as spatial or temporal resolution (Dietzel and Clarke, 2004; Jantz and Goetz, 2005; Liu and Andersson, 2004; Samat, 2006; Yeh and Li, 2006).

However, bearing in mind the CA-based models' complexity and peculiarity, it may be of interest to conduct a partial validation

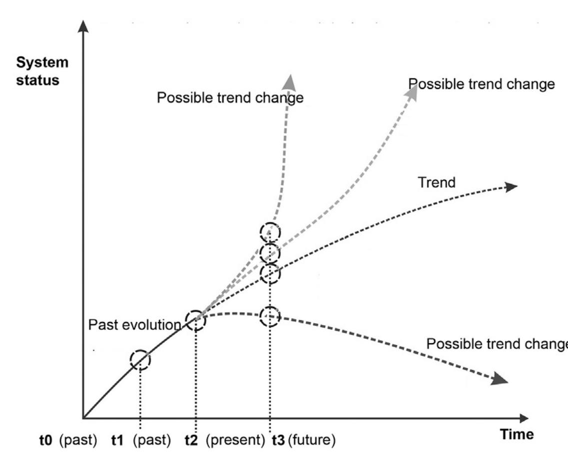

Fig. 2. Possible changes on urban dynamics trends may make it necessary to use scenarios: evolution of time is shown in the axis of abscissa and the state of the different variables that describe the system, i.e. urban growth quantity, are shown in the axis of ordinates.