Computers, Environment and Urban Systems 54 (2015) 119–131

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

# From raster to vector cellular automata models: A new approach to simulate urban growth with the help of graph theory

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

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

ARTICLE INFO

*Article history:*
Received 24 April 2015
Received in revised form 22 June 2015
Accepted 21 July 2015
Available online xxxx

*Keywords:*
Cellular automata
Urban growth
Graph theory
Vector
Simulation models

ABSTRACT

The majority of celullar automata (CA) based models developed to simulate land use changes or urban growth employed a raster representation of the space which reduces reality into regular pieces. This approach is perfectly valid when simulating on large study areas. However, the use of vector representation of space becomes more useful in small extensions since they represent land use covers in a more realistic way. In this regard, a new vector CA-based prototype is presented making use of cadastral parcels and cellular space is combined with graph theory to get a better operability. Neighbourhood factors are defined more flexibly since every parcel is different from the rest. Accessibility, suitability and zoning were computed and added to the prototype's transition rules. A little municipality within the Community of Madrid, one of the most dynamic spaces in terms of urban growth over the last decades, was selected to test the prototype.



## 1. Introduction

Land use change, and urban growth dynamics in particular, are two of the phenomena that have contributed to land transformation in recent times (Seto, Fragkias, Güneralp, & Reilly, 2011). These kinds of change have a major environmental impact wherever they occur, with loss of natural land and the alteration of natural systems being two of the most significant consequences (Alberti & Marzluff, 2004; Berling-Wolff & Wu, 2004; Lauf, Haase, Hostert, Lakes, & Kleinschmit, 2012).

Studies of land use change as a dynamic system have enabled researchers to define the components and interrelations (Macgill, 1986) from which land use change emerges (Itami, 1994). One good way to describe these systems is to model their behaviour using tools that enable a spatial simulation of their associated dynamics (Aljoufie, Zuidgeest, Brussel, van Vliet, & van Maarseveen, 2013; Jantz, Goetz, Donato, & Claggett, 2010; White, 1998).

Urban growth represents one of the most significant land use changes, and simulation models constitute an essential tool for understanding the behaviour of expanding urban systems and determining their main characteristics, such as emergence, self-similarity, self-organisation and non-linear behaviour (Barredo, Kasanko, McCormick, & Lavalle, 2003; Frankhauser, 1998; Portugali, 2000). These models can be used to reproduce the dynamics related to these kinds of systems (e.g. Benenson & Torrens, 2004), explore the factors that act as driving forces (e.g. Jokar Arsanjani, Helbich, Kainz, & Darvishi Boloorani, 2013; Leao, Bishop, & Evans, 2004), analyse urban growth patterns (e.g. Aguilera, Valenzuela, & Botequilha-Leitao, 2011; Shafizadeh Moghadam & Helbich, 2013) and how they affect land (e.g. Mitsova, Shuster, & Wang, 2011), and lastly, study future consequences derived from several scenarios (e.g. Sante, Garcia, Miranda, & Crecente, 2010; Verburg, Schot, Dijst, & Veldkamp, 2004; Zhang, Ban, Liu, & Hu, 2011).

A wide diversity of tools, approaches and spatial resolutions can be used to simulate urban dynamics: depending on the simulation method employed, these include cellular automata (CA) based models, artificial neural networks, fractal modelling, linear/logistic regression, multicriteria evaluation techniques, agent based models and decision tree modelling. Of these, the most widely used models over the last two decades have been CA-based models (see Sante et al., 2010; Triantakonstantis & Mountrakis, 2012). The main reasons for their success are their simplicity, flexibility, transparency and intuitiveness. They are also able to accurately reproduce emerging complex dynamics such as cities (Liu, 2012; White & Engelen, 1993).

Nevertheless, several aspects of urban growth modelling have not yet been satisfactorily solved through the use of CA-based models, such as the representation of time built into model ticks, the introduction of randomness, raster cellular space (which does not take into account the real structure of the land) or the use of scales that are inappropriate for urban or regional planning. As a consequence, only a few models have been integrated into decision-making processes in planning (Triantakonstantis & Mountrakis, 2012).

The CA-based models that have proliferated most to date are those which represent space using raster surfaces (Barredo, Demichelli, Lavalle, Kasanko, & McCormick, 2004; Batty, 1998; Dietzel & Clarke,

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

http://dx.doi.org/10.1016/j.compenvurbsys.2015.07.004

2004; Li, Yang, & Liu, 2008; Sante et al., 2010; White & Engelen, 1993; Wu, 2002). This method of representation is strongly influenced by remote sensing (O'Sullivan & Torrens, 2000), in which satellite or aerial images represent the geographical space as a set of pixels (cells). This has been the most widely used format in this kind of model, partly due to its powerful spatial analysis capabilities and its capacity for mathematical calculation, yielding an appreciable reduction in computational time.

However, an important question arises during the modelling process: is it appropriate to use raster surfaces in urban growth simulations? Urban space transformations do not follow the pattern of a regular structure such as that used in a raster data model; instead, they usually fit into the pre-existing land structure (basically plots or land parcels) so that it is irregular parcels which change rather than regular cells. Therefore, the use of regular cells as the minimum space representation units may be excessively simple in some cases and model outputs can be sensitive to cell size (Jantz & Goetz, 2005; Ménard & Marceau, 2005; Moreno, Wang, & Marceau, 2009; Pinto & Antunes, 2010).

Consequently, although a raster surface may be appropriate for urban growth simulation on large study areas, this representation is to some extent less attractive when the aim is to describe an urban system and its associated dynamics at a more reduced extent, close to urban planning scales (Bardají, 2011). At this level, the use of the real (and irregular) spatial units employed in urban planning could help build a more realistic model. In the Spanish planning practice the urban planning units are the existing cadastral parcel. Therefore, their use is not only a scale issue; they constitute a more realistic representation of the spatial reference unit.

To address these two difficulties, the present article describes the implementation of a CA-based urban growth model prototype to explore the viability of vector land representation at a municipal scale. Irregular space structure was based on cadastral parcels (also known as plots), which constitute the geographical space division (according to land property) used in urban planning. Thus, the use of this structure to simulate urban growth will yield models capable of working at a scale closer to that employed in urban planning.

Attaining this goal necessarily involves a significant relaxation of the formal structure of CA-based models. Irregular space representation entails implementing CA-based models in vector format; thus, it was necessary to adopt a computational solution since the use of this kind of format involves high computational costs. This was addressed through the use of graph theory, by abstracting vector representations to graphs. The prototype was entirely developed in Python (Van Rossum & Drake, 2008), currently the most widely used open-source programming language for GIS. It has been tested in Los Santos de la Humosa, a municipality within the Community of Madrid, one of the most urbanised areas in Spain in recent decades. We do not expect to reproduce past tendencies or dynamics in that region in this stage of the modelling process. The purpose of the test is to verify the computational and operational viability of the prototype.

The present article is organised as follows: the conceptual basis of CA-based modelling and its relaxation is described in Section 2. In Section 3, the prototype is presented; Subsection 3.1 describes the area where the prototype was tested and the input data, Subsection 3.2 defines the conceptual basis of the prototype, Subsection 3.3 details how the vector structure is abstracted to a graph and Subsection 3.4 outlines implementation of the prototype. Section 4 presents the results and the discussion and finally, the conclusions are given in Section 5.

## 2. CA-based model structure and its relaxation as the basis of a new prototype

### 2.1. Conceptual basis of CA

Over the last few decades, several definitions of 'cellular automata' have appeared in the literature (Benenson & Torrens, 2004; Ménard & Marceau, 2005; Torrens & O'Sullivan, 2001; White, 1998; White & Engelen, 2000; Wolfram, 1984; Wu et al., 2009). In general, CA can be considered as a mechanism that reproduces processes whose characteristics change over time based on their status, rules and inputs (Benenson & Torrens, 2004). Wolfram (1984) proposed a more detailed description of the 5 main characteristics of CA: (i) discrete space representation, (ii) discrete time representation, (iii) each cell has one state from a set of possible states, (iv) transition rules depend only on the neighbourhood of a cell and its state, and finally, (v), the state of the cells changes according to the same transition rules (Fig. 1).

CA trace their origins to the 1940s when Stanislaw Ulam and Konrad Zuse established the theoretical basis of what they called 'computing spaces' for representing discrete physical systems, partially based on the work developed by Alan Turing (1936). Subsequently, John Von Neumann gave shape to these theories (Itami, 1994) by defining the CA (originally 'cellular space') as we currently know them, implementing the neighbourhood factor. As a result of these seminal works, the development and use of CA continued, with the Game of Life by John Conway (Gardner, 1970) representing a milestone in mathematical research and in the application of CA. Later on, Waldo Tobler (1970) demonstrated the great utility of cellular spaces for implementing models that simulated or represented dynamic systems in the field of geography (what would later be called 'cellular geography', Tobler, 1979).

The Game of Life, one of the most classic and fundamental examples of a two-dimensional CA, consisted of a set of cells with two possible states, dead or alive, whose evolution depended on the state of their neighbouring cells (transition rules). When the Game started running, spatial patterns were generated whose characteristics have been widely studied in the literature, such as emergence, self-similarity, self-organisation and non-linear behaviour.

The concept of emergence is based on the implementation of local transition rules in a system. These local rules have the capacity to generate complex patterns at global scales, so that the whole is more than the sum of its parts (Torrens, 2000). Self-organisation refers to the capacity of these systems to generate ordered patterns at global scales that emerge from local rules. On some occasions, these systems have the capacity to generate patterns whose parts are similar to the complete pattern, regardless of the scale. This characteristic is known as self-similarity. Finally, the behaviour of these systems is non-linear (non-linear behaviour), leading to patterns not previously seen.

These characteristics are typical of complex systems, and they have been and continue to be applied in several research fields. More specifically and in relation to the present study, these characteristics are typical of systems such as cities, and urban areas in general. It was these similarities among patterns and Game of Life characteristics that aroused interest in applying this methodology to simulate urban processes.

Based on the definition of formal CA proposed by Wolfram, many urban growth simulation models have been developed to date, paving the way for several applications (Barredo et al., 2003; Batty, 2007; Itami, 1994; Sante et al., 2010; Triantakonstantis & Mountrakis, 2012; White & Engelen, 1993). The proliferation of these models and the inherent complexity of urban dynamics have led to the introduction of a series of relaxations in their initial structure to render the models more adaptable and thus more capable of simulating these dynamics much more realistically.

### 2.2. From formal and strict CA structure to flexible CA

A strict CA structure as defined by Wolfram (1984) presents several limitations when simulating a complex phenomenon such as urban growth, as Couclelis (1985) has pointed out. Thus, it has become necessary to use models based on a CA structure, but more complex (White, 1998) and with some slight modifications to enhance them, using relevant aspects and theories related to urban shape, structure and dynamics. Several authors have discussed about the need to relax the formal