

1 **A CELLULAR AUTOMATA MODEL BASED ON IRREGULAR CELLS:**
2 **APPLICATION TO SMALL URBAN AREAS**

3 **Nuno Norte Pinto**

4 Lecturer, Department of Civil Engineering, University of Coimbra, Portugal

5 E-mail: npinto@dec.uc.pt

6
7 **António Pais Antunes**

8 Professor, Department of Civil Engineering, University of Coimbra, Portugal

9 E-mail: antunes@dec.uc.pt

10
11 **Abstract:** We present in this paper the development and application of a cellular automata
12 (CA) model that uses irregular cells for simulating land use change instead of the common
13 regular cells based on the pixels of remote sense images. The use of irregular cells combines
14 urban form with reliable demographic, socio-economic, and building data that are usually
15 available for irregular census blocks. The calibration of the model is made through a
16 procedure based on a particle swarm algorithm that optimizes a measure of agreement derived
17 from the *kappa* index for contingency matrixes. The model was applied for simulating urban
18 change in a small municipality in Portugal that registered high growth rates during the 1990s.
19 Simulation results indicate the model's ability for capturing land use change in small urban
20 areas.

21 **Keywords:** Urban Change, Cellular Automata, Irregular Cells, Small Urban Areas, Particle
22 Swarm.

1 1 INTRODUCTION

2 The problems generated by the rapid growth of urban areas raise important issues to the
3 planning process. These issues are related not only with the motives behind past evolution,
4 but also with the definition of new land use policies capable of responding to the needs of the
5 present and to control the evolution of urban areas towards a sustainable future. The
6 complexity of the problems is such that there are no simple ways to achieve solutions, nor
7 these solutions can be based on a single approach. The comprehensive nature of the planning
8 process is simultaneously a strength, because it derives from the consideration of all
9 ingredients of the problems (physical, demographic, economic, social, etc.), and a potential
10 weakness, because the problems become more and more complex as the natural evolution of
11 societies takes place, especially in the current globalized context, demanding from planners
12 new levels of commitment and accuracy in their research and practice.

13 The complex problems involved in comprehensive urban planning can be advantageously
14 addressed through models capable of capturing the behavior of urban areas and forecast their
15 evolution. Cellular automata (CA) have received a great deal of attention from researchers
16 among the wide range of modeling techniques that are usually applied to urban simulation. In
17 a concise definition “an automaton is a processing mechanism with characteristics that change
18 over time based on its internal characteristics, rules and external input” (Benenson and
19 Torrens, 2004). CA are based on a discrete set of spatial units called cells that together form a
20 cell space. Each cell takes a given (cell) state from a finite set of states. Time is considered in
21 a discrete manner. Each cell, which works as an automaton, then operates state changes over
22 time according to a finite set of transition rules which can be of various types (deterministic or
23 stochastic, unconstrained or constrained). State transition is the result of the application of
24 these rules to each cell considering the neighboring cells.

1 In this paper we present the development and application of a CA model based on irregular
2 cells, a type of model that was used only in a few studies reported in the literature (the first
3 ones are due to Semboloni, 2000 and O'Sullivan, 2001a). CA models commonly use regular
4 cells obtained from satellite images (pixels), which may not represent well the spatial
5 structure of the territory because their form is only constrained by resolution and they only
6 hold information on land use (obtained from remote sensing using automatic classification
7 methods). Our model uses irregular cells based on census blocks – the smallest spatial units
8 used by national statistics systems – to combine form with reliable data on land use,
9 demographics, and buildings. The model was applied in the study of land use change in small
10 urban areas. The CA models dealt with in the literature (at least in leading journals) refer to
11 metropolitan or large urban areas (illustrative examples include Clarke et al, 1997, Silva and
12 Clarke, 2002, and White and Engelen, 1993). However, small and mid-size cities face
13 noteworthy growth problems. Indeed, a very significant share of urban growth is taking place
14 in areas with population below 500,000 (Cohen, 2004). This is in part due to the
15 counter-urbanization trends that are being observed in many regions of the world and,
16 especially, in the developed countries (Cheshire and Hay, 1985; 1989; Ferras, 2007).

17 The paper is divided into three main sections, plus a final section for conclusions. In Section 2
18 a brief literature overview is presented, where the key research on CA and their applications is
19 introduced. Section 3 is dedicated to the presentation of the CA model we have developed in
20 our study. In Section 4 we describe the application of the model to the small municipality of
21 Condeixa-a-Nova, in central Portugal, and provide a forecast analysis of urban growth in this
22 municipality for two ten-year periods. Finally, Section 5 is dedicated to the discussion of
23 results and to the presentation of some conclusions about the study reported in this paper.

24

1 **2 LITERATURE OVERVIEW**

2 The concept of cellular automaton was introduced in the late 1940s by John von Neumann
3 and Stanislaw Ulam, who were facing (independently) the problem of devising sets of
4 mathematical rules for simulating the self-reproduction and evolution of biological systems.
5 The spatial nature of CA led Waldo Tobler to pioneer their application to urban studies
6 (Tobler, 1979). This occurred shortly after Douglass Lee published his paper about the fallacy
7 of large-scale modeling (Lee, 1973), which damaged considerably the confidence on the
8 application of models to urban studies, reducing the intensity of research in the field. Until the
9 1990s, the object of planning was re-centered on short term objectives, on immediate
10 solutions to problems, rather than on the accomplishment of ambitious long-term goals,
11 characteristic of large-scale planning. Despite this, simulation models gained a new breath in
12 this period, thanks to the advent of faster and less expensive computer hardware, along with
13 the development of more powerful database management tools. Couclelis followed Tobler on
14 the exploration of the concept of CA and produced their first applications to urban studies
15 (Couclelis, 1985; 1988; 1997). Other authors, like Batty and Xie (1994; 1997) and White and
16 Engelen (1993; 1997) further explored the concept, extending it with new theoretical features,
17 whereas Semboloni (1991) and Webster and Wu (1999a; 1999b) tested CA as a tool for the
18 evaluation of land use policies. Some of the most important evolutions of CA involve their
19 combination with various other techniques. The capacity of GIS for dealing with large
20 datasets and for providing quality visualization led Li and Yeh (2000), Wu and Webster
21 (2000), and Stephens et al (2007) to couple CA with GIS. Fractal theory was used to assess
22 the quality of CA simulation by comparison of simulated and observed fractal measures of
23 real-world cities (Barredo et al, 2003; Jenerette and Wu, 2001; White and Engelen, 1993). CA
24 was combined with multicriteria evaluation to determine transition rules (Wu and Webster,
25 1998). Neural networks were coupled with CA to calibrate modeling parameters (Almeida et

1 al, 2008; Li and Yeh, 2001; Wu, 2002). The integration of CA with multi-agent simulation
2 (MAS) is at the origin of the concept of geographic automata systems proposed in Benenson
3 and Torrens (2004) and is currently a major subject of research (Diappi and Bolchi, 2006;
4 2008; Vancheri et al, 2008).

5 CA have been applied to a series of real world studies around the world. SLEUTH is a CA
6 model used in the simulation of historical land use evolution in the San Francisco Bay Area
7 (Clarke et al, 1997) which has been largely applied worldwide (Jantz et al, 2004; Silva and
8 Clarke, 2002; Xian and Crane, 2005). Another CA model under intensive use is MOLAND,
9 which has been widely applied to land use simulations across the European Union and also in
10 developing countries (Barredo et al, 2003; Barredo and Demicheli, 2003; Uhel et al, 2006).
11 Ward et al (2000; 2003) developed a CA model coupled with a macro-scale spatial interaction
12 model for the Gold Coast in Australia. Jenerette and Wu (2001) developed a coupled
13 CA/genetic algorithm model to evaluate long term historical urbanization in Phoenix,
14 Arizona.

15 The use of irregular cells in CA models is very limited, but is gaining importance in recent
16 years. Indeed, remote sensing images are commonly used as input data to CA models after
17 pre-processing with automatic land use classification algorithms, making them easy to use
18 and providing different degrees of spatial resolution. There are only few exceptions to the use
19 of regular cells. Semboloni (2000) and Shi and Pang (2000) used Voronoi polygons to
20 develop CA models for simulating urban growth. O'Sullivan (2001a; 2001b) combined CA
21 and graph theory to generate sets of neighborhood-scale irregular cells. Vandergue et al
22 (2000) defined irregular "sectors" to study intra-city migration patterns. Stevens and
23 Dragicevic (2007) and Stevens et al (2007) coupled a high resolution irregular CA with a GIS
24 for modeling neighborhood expansion at the parcel level with GIS, thus having a vector space
25 with geometric features and making use of database capabilities provided by GIS. Moreno et

1 al (2008, 2009) used vector irregular objects for cells that can change shape during time,
2 depending on the interactions with an extended set of neighbors chosen according to their
3 influence on state transition. This vector based CA deals with geographic features of any kind,
4 in particular with irregular polygons that simulate real-world features, being census blocks
5 one possibility. Vliet et al (2009) implemented a variable grid neighborhood to incorporate
6 long distance land use interaction by considering all the land use available for computing the
7 potential for state transition.

8 In general, CA applications are focused on large scale models (regional or metropolitan). To
9 the best of our knowledge, there is no previous work reported in leading scientific journals on
10 the use of CA for small urban areas.

11 **3 MODEL PRESENTATION**

12 The CA model presented in this paper has a simple structure that derives from the classical
13 formulation of CA with the consideration of constrained land use demand, following the
14 concept introduced by White and Engelen (1993). The model operates over an irregular
15 cellular fabric obtained from census blocks. Cell states are classified into a finite set of
16 aggregated classes of land use. Land use interactions take place within a variable
17 neighborhood that is determined through model calibration. Transition rules incorporate
18 planning regulations and simulate land use change based on a composite transition potential
19 that takes into account cell accessibility, land use suitability, and neighborhood interactions
20 within the cell neighborhood. The time step can be defined by the user. Land use demand is
21 determined through the evolution of population and employment densities over time. These
22 topics are explored below in separate subsections.

23 The model was implemented from scratch using the Visual Basic 6 programming language to
24 run on Microsoft Windows.

1 **3.1 Cell Structure**

2 The first component of a CA model is the cell structure. Our model is based on the use of
3 irregular cells. The purpose is to allow individual cells to simulate irregular census blocks.
4 Census blocks hold structured and reliable information on demographics and buildings, and
5 are drawn considering urban form (streets, urban parks, property walls/limits, etc). By
6 combining data and form, they are particularly suited for being used as the cells of a CA
7 model.

8 **3.2 Cell States**

9 Another important component of a CA model is the set of cell states. The number of states
10 can be as small as “housing” or “farm” as a binary state (Caruso et al, 2007) or extremely
11 large as in the MOLAND model (Barredo et al, 2003). In our model, we classify land use into
12 six aggregate states (or land use classes), which are intended to represent the whole set of
13 urban land uses: urban low density (UL), urban high density (UH); industry (I), urban
14 expansion (XU), industrial expansion (XI), and highly restricted uses (R) .States UL and UH
15 are comprehensive classes of urban land uses, differing only on building density. They
16 include residential areas, shopping and office areas, mixed-uses areas, streets, and open
17 spaces. States XU and XI represent non-occupied areas where urban and industrial land uses
18 can be located, respectively. Cells classified in any state but R are identified as active cells:
19 these cells can change state under a given set of conditions. Cells classified as R, which
20 correspond to land classified as agricultural reserve, ecological preservation area, etc., are
21 identified as inactive cells: they can influence land use dynamics but their state cannot
22 change.

23 **3.3 Neighborhood**

1 The third component of a CA model is the neighborhood, that is, the extent to which land use
2 interactions occur. The concept of neighborhood is a key element for CA as it determines the
3 spatial extent of interactions, embodying Tobler’s first law of geography: “everything is
4 related to everything else, but near things are more related than distant things” (Tobler, 1970,
5 p. 236). Instead of using traditional von Neumann or Moore neighborhoods, or combined
6 settings of these two (Benenson and Torrens, 2004), we chose to employ circular
7 neighborhoods. This choice has also been made in several other CA models: it is common to
8 consider the cells located within a given radius in every direction (Barredo et al, 2003; White
9 and Engelen, 1993). In these cases, the number of cells is predetermined considering, in an
10 arbitrary way, a distance perceived as appropriate for defining neighborhood size. In our
11 model, we implemented a variable circular neighborhood where the radius is a calibrated
12 parameter, δ . Its value influences, and is influenced by, all the other parameters of the model.
13 Neighborhood perception is different for different urban areas. Therefore, any model that
14 aims to capture land use interactions in a given territory gains by being able to identify how
15 far these land use interactions really take place.

16 **3.4 Transition Rules**

17 The fourth component of a CA model is the set of transition rules that are behind the
18 dynamics of urban areas. In our model, they are expressed through a potential which reflects
19 the propensity/capacity of each cell to change state in each time step. This potential is a
20 weighted value of land use suitability, accessibility, and neighborhood effects given by the
21 following expression:

$$22 \quad P_{i,s} = (\nu_P \times S_{i,s} + \chi_P \times A_i + \theta_P \times N_{i,s}) \times \xi, \forall i \in \mathbf{C}, s \in \mathbf{S} \quad (1)$$

23 where, for each cell i from the set of cells \mathbf{C} , and for each state s from the set of states \mathbf{S} , $P_{i,s}$ is
24 the transition potential for state s of cell i , $S_{i,s}$ is the land use suitability value for state s of cell

1 i , A_i is the accessibility value of cell i , $N_{i,s}$ is the neighborhood effect for state s of cell i
 2 considering its neighborhood \mathbf{V}_i , ν_P is the calibration parameter for land use suitability, χ_P is
 3 the calibration parameter for accessibility, and θ_P is the calibration parameter for the
 4 neighborhood effect.

5 The first element of the potential is cell suitability, which is considered as a binary value that
 6 is equal to 1 if the cell is suitable for a given land use and 0 otherwise, that is: $S_{i,s} = 1$ if cell i
 7 is suitable for state s , $S_{i,s} = 0$ otherwise. The choice for a binary variable is related to the fact
 8 that suitability and zoning were considered as one single input. It is assumed that planning
 9 regulations define suitable and unsuitable land uses for each cell, according with a
 10 predetermined zoning. Therefore, if one cell is classified as suitable for a given land use then
 11 it is assumed as being capable of attracting that land use.

12 The next element is accessibility. It is assessed considering travel times between cell centroids
 13 over the road network. The network is classified in hierarchical levels – e.g., main roads,
 14 secondary roads, and local roads – according with road capacity and legal speed limits. The
 15 value of accessibility for cell i , A_i , is assessed through the following expression:

$$16 \quad A_i = 1 - \frac{f(T_i^*)}{\left\| \sum_{j \in \mathbf{C}} f(T_j^*) \right\|}, \forall i \in \mathbf{C} \quad (2)$$

17 where $f(T_i^*)$ is an impedance function (typically an exponential function or a power function)
 18 of an aggregate measure of travel time given by

$$19 \quad T_i^* = \alpha_A \times T_{i,C} + \beta_A \times T_{i,V} + \gamma_A \times T_{i,I}, \forall i \in \mathbf{C} \quad (3)$$

20 and $T_{i,C}$ is the travel time from cell i to the municipality's main town, $T_{i,V}$ is the travel time
 21 from cell i to its civil parish (or district) main village, $T_{i,I}$ is the travel time from cell i to the
 22 closest industrial site located in the municipality, and α_A , β_A , and γ_A are calibration parameters.

1 Accessibility is based on the assessment of the proximity of a cell to the urban functions
 2 available at the municipality's main town and the civil parish main village, and to the
 3 employments offered not only in urban areas but also in industrial sites.

4 The last element of the potential is the neighborhood effect, $N_{i,s}$. It is considered as an
 5 aggregate value of the interactions $N_{i,s|j,r}$ between the states (or land uses) s and r , located in
 6 two neighboring cells i and j . It is calculated through the following expression:

$$7 \quad N_{i,s} = \sum_{j \in \mathbf{V}_i} N_{i,s|j,r}, \forall i \in \mathbf{C}, \mathbf{V}_i = \{j \in \mathbf{C} : d_{ij} \leq \delta\}, s, r \in \mathbf{S} \quad (4)$$

8 where the neighborhood effect $N_{i,s}$ is the sum of interactions $N_{i,s|j,r}$ between state s in cell i and
 9 all the states of neighboring cells j that belong to neighborhood \mathbf{V}_i considering the
 10 neighborhood distance parameter δ (which means that attraction or repulsion will only be
 11 taken into account if the cell is located within the neighborhood \mathbf{V}_i) and $d_{i,j}$ is the distance
 12 between cells i and j .

13 For each pair of states, these interactions $N_{i,s|j,r}$ were considered to be a linear function of the
 14 distance between the cells, as depicted in Figure 1. The value of the function is 1 if the
 15 interaction between the states of two cells is characterized by maximum attraction, 0 if they
 16 do not interact, and -1 if the interaction is characterized by maximum repulsion. This value is
 17 obtained as follows:

$$18 \quad N_{i,s|j,r} = 1 - \frac{d_{i,j}}{d_{s,r}^{\max}} \times N_{s,r}^{\max}, \forall i, j \in \mathbf{C}, s, r \in \mathbf{S}; \text{ if } d_{ij} \leq d_{s,r}^{\max} \quad (5)$$

0; otherwise

19 where $N_{s,r}^{\max}$ is the maximum value for the interaction (the ordinate at the origin) between cells
 20 in state s and in state r , $d_{s,r}^{\max}$ is the distance for which the interaction is zero.

21 **Insert Figure 1 approximately here**

1 The transition potential of a cell is increased of a given value that depends on the distance
2 between this cell and each neighboring cell, if their states attract themselves. In contrast, the
3 potential is decreased of a given value if their states repulse themselves. For example,
4 according to the graphic depicted in Figure 1(a), a pair of cells classified in states that show
5 mutual attraction, say UL and UH, would have their potentials increased of about 0.4 (prior to
6 normalization) for a distance between the cells of 0.3 kilometers. Considering the opposite
7 relationship, depicted in Figure 1(b), a pair of two cells classified in states that induce mutual
8 repulsion, say XU and I, would have their potential reduced of about 0.3 (again, prior to
9 normalization) for a distance between the cells of 1.0 kilometer. It is important to note that all
10 the six states influence the potential of cells whatever their states are. However, only the five
11 active states are influenced by the other states (all but R, where construction is highly
12 restricted). This neighborhood effect is a relationship that must be determined for each pair of
13 states considered in the analysis of urban change. The fact that these relationships are very
14 difficult to assess, as they depend on several interdependent factors such as land value,
15 housing demand, public facility location, among several others, suggested the consideration of
16 the two points that define them – $N_{s,r}^{max}$ and $d_{s,r}^{max}$ – as calibrations parameters.

17 Finally, the stochastic perturbation, ξ , is calculated through the following expression:

$$18 \quad \xi = 1 + [-\log(\rho)]^\sigma \quad (6)$$

19 where ρ is a random variable uniformly distributed in the interval]0,1[and σ is a control
20 parameter for the adjustment of the size of the perturbation (White and Engelen, 1993). This
21 term has a highly distorted distribution so that most values are near one and much larger
22 values occur only infrequently. The main purpose is to introduce the stochastic behavior of
23 agents in the transition potential.

24 The transition potential is calculated for every state and cell, being the final cell potential the

1 highest value for the set of states. For comparison purposes the potential must be normalized,
2 being its value given by:

$$3 \quad P_i = \max \left\{ \frac{P_{i,s}}{\sum_{s \in \mathbf{S}} P_{i,s}} \right\}, i \in \mathbf{C}; s \in \mathbf{S} \quad (7)$$

4 **3.5 Time**

5 The last key component of a CA model is time. Usually, the time step used in CA models is
6 one year long. Our model can use time steps of one, five, or ten years. The definition of the
7 time step is a user decision based on the case that is being modeled. It can be connected with
8 data availability, as the use of a time step similar to the censuses time lag frames CA
9 dynamics with consistent census data. Also, it can be related with the rate at which a given
10 cell changes state. This rate depends in particular of the planning environment – in highly
11 regulated systems these changes are relatively uncommon.

12 **3.6 Land use demand**

13 Our model deals with land use demand differently from classic CA models. Two irregular
14 cells most probably have different areas, with different values for population or employment
15 as well as for building density, supplying the land use market with different amounts of land.
16 If demand was assessed through the number of cells that change state (as is usually the case
17 with CA models), the sum of the newly occupied areas would not match the increase in
18 population or employment that generated the demand. Therefore, our model allocates
19 population or employment considering given thresholds for their densities (that are model
20 inputs) in order to guarantee that it is able to simulate the observed value of land consumption
21 for every state.

1 The demand for urban land uses (UL and UH) depends on three main variables: the variation
2 of population, the variation of household size, and the variation of building density. The
3 increase in the number of households (single persons, single parent families) and the
4 subsequent reduction of the average size of households implies the use of more space for
5 housing, thus increasing land use demand. At the same time, building density tends to
6 decrease over the years because of increasingly exigent requirements of public facilities and
7 public space. The best way to relate these variables is through population density. The model
8 calculates the increase of population during the reference period and distributes that
9 population over the territory, considering thresholds of population densities. Land use demand
10 is the surface of land that is necessary to accommodate this increase in population. The
11 thresholds are calculated for different levels of building density, for different moments in
12 time.

13 Similarly, the demand for the industrial land use (I) is modeled based on employment
14 densities, considering observed values of employment per area for the existing industrial sites.

15 **4 CASE STUDY**

16 The CA model presented in the previous section was tested on a case study involving
17 Condeixa-a-Nova, a small municipality located near Coimbra, in central Portugal (Figure 2).
18 This municipality had very high population growth rates in the 1991-2001 (the 11th highest in
19 mainland Portugal outside the metropolitan areas of Lisbon and Porto). In this section, we
20 present the study area and the dataset created for the application of the CA model. We also
21 present the procedure implemented for calibrating the model, which is based on an
22 optimization technique called particle swarm and on the use of contingency matrices for
23 performance assessment. Finally, we present the results for a forecast analysis of urban
24 change in Condeixa-a-Nova for two ten-year periods.

1 **Insert Figure 2 approximately here**

2 **4.1 Characterization of the Study Area**

3 The municipality of Condeixa-a-Nova has around 150 square kilometers and 15000
4 inhabitants (INE, 2001). The population growth rate for the municipality in the 1991-2001
5 period was 17.8 percent. This value is particularly remarkable because the natural growth rate
6 (birth rate minus mortality rate) was -3.7 percent over the same period. Built areas had an
7 even more impressive 48 percent increase. In 2001 these areas amounted to about 10 percent
8 of the municipality.

9 The municipality is headed by the town of Condeixa-a-Nova and has several smaller villages.
10 The majority of the urban areas are located in the north and western sectors of the
11 municipality. This part of the municipality is relatively flat and mostly covered with
12 agricultural land uses (vegetables, orchards, vineyards, and olive groves). These two sectors
13 also concentrate most of the existing urban expansion areas. The south and southeast sectors
14 of the municipality are hilly to mountainous, with a scattered urban occupation surrounded by
15 scrubland and forest, a large portion of which classified as ecological preservation area.

16 **4.2 Dataset**

17 Case study data were pre-processed to create a dataset in the format that is supported by the
18 model. There were two main data sources: demographic and building data by census block
19 obtained from the national censuses of 1991 and 2001; and land use classification obtained for
20 the same years from municipal master plans. Irregular cells were drawn from the intersection
21 of census blocks (obtained from the geographical databases that support both censuses) with
22 the areas inside the urban limits defined by municipal planning in 1991 and in 2001 (as seen
23 in detail in Figure 3). Land use maps for these two years are depicted in Figure 4.

1 **Insert Figure 3 approximately here**

2 **Insert Figure 4 approximately here**

3 Land use was classified into the six classes (states) required by the model. Urban land use
4 demand was defined by population densities thresholds, determined considering the average
5 household size and dwellings per hectare by the time of each census. This variation is a
6 consequence of the reduction of population density during the inter-censuses period as a result
7 of the two factors discussed previously: (1) the decrease of the average size of households;
8 and (2) the inclusion of much more public facilities and public space in urban areas. The
9 values used for calculating the thresholds are presented in Table 1.

10 **Insert Table 1 approximately here**

11 From 1991 to 2001 the variation of total area for state UH was significant, achieving an
12 impressive 351 percent, while state UL registered a small increase of 2 percent. This
13 difference between UH and UL states is related with the change in population density over the
14 period: land classified as UL in 1991 was re-classified as UH in 2001 due to the change in the
15 population density threshold mentioned above.

16 Accessibility was calculated using a power impedance function with exponent equal to 1.0.
17 This type of function explains traffic flows in the Coimbra area (where Condeixa-a-Nova is
18 located) extremely well.

19 The time step chosen was ten years, the interval between census and between revisions of
20 municipal master plans). The fact that land use is highly regulated in Portugal (particularly
21 since 1990), making state changes for a given cell infrequent, led us to discard the use of
22 shorter time steps.

23 **4.3 Model Calibration**

1 4.3.1 Calibration Approach

2 CA models need to be calibrated in order to ensure the best possible adjustment between the
3 simulated outcome and the reality that is being modeled. Calibration can be achieved through
4 two main approaches: (1) performing a sensitivity analysis for the value of each calibration
5 parameter considering the other parameters fixed; or (2) running an optimization procedure
6 for searching the set of calibration parameters that optimize some fitness function. The first
7 approach is based on a visual comparison of a reference map with a simulation map obtained
8 with some initial set of parameters. The sensitivity analysis is performed to evaluate how
9 much each parameter varies when the other parameters are fixed and to iteratively adjust their
10 values until the simulation map and the reference map meet some similarity criteria. This
11 calibration approach becomes increasingly difficult to apply as the number of parameters
12 increases.

13 In our case, the number of calibration parameters is considerably high: one neighborhood
14 distance parameter (δ), three transition potential parameters (χ_P , v_P , and θ_P), the stochastic
15 parameter (σ), and three accessibility parameters (α_A , β_A , and γ_A), and 40 neighborhood
16 parameters ($N_{s,r}^{max}$ and $d_{s,r}^{max}$ for each pair of cell states, except cell state R and itself). This
17 means a total of 48 calibration parameters. This strongly suggests the use of an optimization
18 approach to calibrate the model – that is, the application of automatic rules to determine the
19 best possible value for a fitness function expressing the goodness of fit of the model.

20 The fitness function we used in this case study is based on the contingency matrix for the
21 reference and simulation maps (Couto, 2003). The contingency matrix for two class maps is a
22 matrix where each element, m_{sr} , expresses the number of cells classified in state s in the
23 simulation map which are in state r in the reference map.

24 Several comparison measures regarding the degree of agreement between two class maps can

1 be extracted from a contingency matrix. The most used one is the *kappa* index, which value
 2 gets closer to one as the similarity between the two maps increases.

3 For our case study, we used a modified version of the *kappa* index to avoid the distortion that
 4 would have been produced if states that cannot take part in the urban dynamics – that is, state
 5 R (highly restricted uses) – were included in the computation. The consideration of cells in
 6 this state would be misleading by producing a large – though meaningless – agreement
 7 between simulation and reference maps.

8 The modified *kappa* index, k_{mod} , that we used as fitness function was calculated as follows:

$$9 \quad k_{mod} = \frac{n \sum_{i \in \mathbf{S}^*} m_{ii} - \sum_{i \in \mathbf{S}^*} \left(\sum_{j \in \mathbf{S}^*} m_{ij} \times \sum_{j \in \mathbf{S}^*} m_{ji} \right)}{n^2 - \sum_{i \in \mathbf{S}^*} \left(\sum_{j \in \mathbf{S}^*} m_{ij} \times \sum_{j \in \mathbf{S}^*} m_{ji} \right)}, \quad \mathbf{S}^* = \mathbf{S} / \{R\} \quad (8)$$

10 where n is the total number of elements in the contingency matrix and m_{ij} is a generic element
 11 of the matrix.

12 To optimize the fitness function we used a recent optimization algorithm called particle
 13 swarm (from now on referred to as PS) which is suitable for dealing with a large number of
 14 parameters. PS is inspired on the flight of bird flocks, which move to some goal starting from
 15 an entirely random distribution in space and then organize themselves behind a leader in a
 16 stable flight formation where the relative position and velocity of birds is maintained through
 17 time. Analogously, the search of the solution space (the possible flight formations) for the
 18 optimum set of calibration parameters is made within the PS algorithm by adjusting the
 19 trajectories of sets of calibration parameters called particles (the birds). Details on the PS
 20 algorithm can be found on Kennedy (1997) and Parsopoulos and Vrahatis (2002).

21 In our case, the parameters are calibrated by running the CA model embedded on a PS

1 algorithm, as depicted in Figure 5. We start with a given number of particles – each particle is
2 a set of 48 parameters – with random positions and velocities, and calculate through a CA run
3 the value of the fitness function (k_{mod}) for each particle. Then, in successive iterations, we
4 update the position and velocity of each particle according to a pair of rules that operate over
5 their previous position and velocity, and the previous position and velocity of the particle with
6 the highest fitness function (swarm leader). The updating process is carried out until the value
7 of the fitness function no longer improves.

8 **Insert Figure 5 approximately here**

9 In Figure 6 we show how the values of three parameters of the model (α_A , β_A , and γ_A) evolve
10 through the application of the PS algorithm. Initially, they vary widely and then progressively
11 converge, more or less quickly, to some final, stable value.

12 **Insert Figure 6 approximately here**

13 4.3.2 Calibration results

14 The CA model was calibrated using the datasets for 1991 and 2001. A group of 60 particles of
15 48 parameters was randomly generated considering plausible ranges for every parameter, and
16 used in the CA runs of each PS iteration. The algorithm was set to stop when the fitness
17 function did not improve more than 0.1 percent in five consecutive iterations. The number of
18 particles was chosen after Trelea (2003).

19 The fitness function reached a final k_{mod} of 0.621, with an improvement of 184 percent when
20 compared to the best run of the initial group of 60 CA runs. The value for $kappa$ was 0.774,
21 which shows that the use of this commonly-used index instead of k_{mod} would overrate
22 simulation quality by 24.7%. The simulation land use map is depicted in Figure 4c. Rather
23 than capturing the trend towards urban concentration observed in Condeixa-a-Nova (Figure

1 4b), the model favored the emergence of surrounding urban centers. It was however capable
2 of identifying areas of change, such as those indicated by the circles displayed on Figure 7.

3 **Insert Figure 7 approximately here**

4 The final values obtained through the PS algorithm for the set of calibration parameters
5 matched what would be expected. The neighborhood distance parameter (δ) was set to 4
6 kilometers. The values obtained for the accessibility parameters were 0.445 for α_A , 0.238 for
7 β_A , and 0.568 for γ_A , which means that the model favored accessibility to the main industrial
8 site. Regarding the transition potential, the values obtained for the parameters were 0.341 for
9 ν_P , 0.130 for χ_P , and 0.933 for θ_P , which means that the model identified neighborhood
10 interactions as the key component of the potential. The relationships corresponding to the
11 neighborhood effect parameters for four pairs of cell states are depicted in Figure 8. Finally,
12 the value obtained for the stochastic parameter σ was 1.229, which matches the results
13 obtained by White et al (1997) for a good degree of predictability.

14 **Insert Figure 8 approximately here**

15 The model was calibrated using the k_{mod} for a contingency matrix expressed in terms of the
16 number of cells (Table 2a). In general, the values of the conditional k_{mod} , which measure the
17 agreement between simulation and reference maps for a specific state, are in line with the
18 global agreement obtained for the case study (0.771). However, the conditional k_{mod} for the
19 UH state is considerably lower. The difference between the values for UL population density
20 in 1991 and UH population density in 2001 is very small, which may have had some
21 influence on the final result, by reducing the capacity of cells in state UL to receive
22 population when slightly larger amounts of population were being allocated to cells with the
23 same size in state UH.

24 It is interesting to compare the contingency matrix expressed in terms of the area of cells

1 (Table 2b) with the previous matrix. The value of k_{mod} is slightly higher (0.643) when the area
2 of cells is considered, and the same happens with the global agreement (0.778). This may
3 indicate that the model could be better calibrated with cell areas instead of cell counts. The
4 values of the conditional k_{mod} are similar or higher for all the states (with the exception of
5 state UH). This reinforces the idea of using cell areas for calculating k_{mod} .

6 **Insert Table 2 approximately here**

7 The procedure used to define cells may have reduced the quality of calibration due to an
8 inappropriate problem definition. It was strictly based on the intersection of census blocks
9 with the areas inside the urban limits. In larger urban areas census blocks are small and
10 originated small cells whereas some smaller villages are located almost completely within a
11 single large census block, which originated large cells. This has led to a wide range of
12 variation in size for cells in urban land uses and to a significant supply of large cells with UL
13 and XU states. The model found good opportunities for allocating population to these larger
14 cells, because of two main reasons. The first reason is the favorable location of some of these
15 large cells. This is, for example, the case with Ega, Belide, and Zambujal. All these villages
16 are mainly contained in one large cell particularly well located in terms of accessibility.
17 Moreover, this cell usually has neighboring cells in state R, a fact that enhances their ability to
18 attract urban land uses. On the contrary, the cell space of the center of Condeixa-a-Nova is
19 strongly disaggregated and cell sizes present some homogeneity. The second reason relates to
20 the possibility of allocating a large amount of population to a large cell without infringing the
21 UH limit for population density.

22 The existence of large cells inside urban centers produced a significant difference for the ratio
23 between simulated and reference areas for each state. The large value of +30 percent
24 registered for UH state along with the negative value of -6 percent for UL state and +3

1 percent for XU state is a direct consequence of the model's preference for larger cells. This
2 preference affected mainly UL cells: in the simulation about 127 hectares (34 cells) changed
3 from UL to UH state, while in reality only 68 hectares (48 cells) changed this way. On the
4 opposite direction was the change from XU to UH state: the simulation assigned only about
5 32 hectares (30 cells) to this change while in reality about 46 hectares (48 cells) changed from
6 the former state to the latter. The same behavior is observed in the change from XU to UL
7 cells: the simulation assigned 44 hectares (27 cells) to this change while in reality there were
8 91 hectares that changed between these two states for a total of 85 cells. These values
9 corroborate the idea that the model favors the occupation of well located large cells in UL
10 state: it assigns a large amount of population to large UL cells and only a small amount to
11 small UL and XU cells.

12 The number of cells whose transition was exactly determined by the simulation was lower
13 than expected considering the high value of k_{mod} . From a total of 94 cells (205 hectares)
14 classified in urban land uses which experienced state transition only for 19 (23 hectares) the
15 simulation matched reality. Nonetheless, the model was able to identify approximately the
16 areas where change has occurred – which is what really matters for planning purposes,
17 particularly at the municipal planning scale (Figure 7).

18 **4.4 Forecast analysis**

19 A forecast of land use change was made using the parameters obtained from calibration for
20 the years 2011 and 2021 (which are coincident with future censuses), starting from the
21 reference situation in 2001 (Figure 9a). A population growth rate of 10 percent was
22 considered for the 2001-2011 and 2011-2021 periods. Figure 9(b) and Figure 9(c) depicts
23 simulation maps for 2011 and 2021 respectively. Simulation results point to the densification
24 of Condeixa-a-Nova's central areas, especially in the areas closer to the industrial site – and

1 this is indeed what is currently happening.

2 **Insert Figure 9 approximately here**

3 **5 CONCLUSION**

4 In this paper we presented the development and application of a CA model based on irregular
5 cells that aim to improve the simulation of urban change by enhancing the representativeness
6 of the underlying spatial structure. The model fits into a body of research that has been
7 growing fast in recent years. This can be attributed to the fact that the consideration of
8 irregular cells in a model makes urban form issues easier to cope with. Furthermore, if the
9 irregular cells correspond to census blocks, vast amounts of reliable demographic and
10 socio-economic data can be used in the application of the model. In addition to being based
11 on irregular cells, the CA model has a number of other interesting features, e.g., cell
12 neighborhoods are variable, cell transition potentials depend on an aggregate accessibility
13 measure, and land use demand is defined considering population and employment growth.
14 Taken together, these features make the model quite innovative.

15 The CA model was tested in Condeixa-a-Nova, a small, fast-growing municipality in central
16 Portugal, with promising results that we expect to be validated as soon as the census of 2011
17 becomes available. The measure used to assess the agreement between simulated and
18 reference maps – a variant of the *kappa* index for contingency matrixes – was clearly above
19 commonly accepted thresholds (for example, the ones utilized for automatic land use
20 classification from remote sensing images). This is remarkable, because the evolution of land
21 use in small, fast-growing municipalities takes place according to patterns that are typically
22 much more irregular than the ones generally encountered in metropolitan and large urban
23 areas. For the calibration of the parameters of the model, we used a particle swarm algorithm.
24 To the best of our knowledge, this type of optimization algorithm, which is particularly suited

1 for the calibration of a large set of interdependent parameters, was never used before in the
2 calibration of CA models.

3 Despite the promising results given by the model, there is significant ground for improvement.
4 Land use demand is one of the issues that need careful attention in order to avoid biases
5 induced by large cells that became too attractive for development. There are two alternatives
6 for tackling this shortcoming: to divide the large cells in smaller ones at the outset (which
7 raises a modifiable area unit problem); to allocate population to cells probabilistically as a
8 function of their transition potential. Other issues to improve are the computation of cell
9 transition potential (e.g., accessibility could be measured taking into account different
10 transport modes) and the extension of the variable neighborhood concept (e.g., to include
11 discontinuous spatial structures and time-evolving configurations). A last issue to address
12 relates to the consideration of a multi-scale approach. Urban phenomena take place at
13 different spatial scales, which suggests that they should be simulated taking scale into account.
14 Within this approach our local-scale CA model would be embedded into a macro-scale CA
15 model to simulate the socio-economic forces that drive urban change at regional level. In the
16 upcoming future, our research on CA models based on irregular cells will be focusing on
17 these issues.

18 **ACKNOWLEDGEMENTS**

19 The authors wish to thank the valuable comments from the reviewers that significantly
20 contributed to improve our paper. Nuno Pinto also wishes to acknowledge the support
21 received from Fundação para a Ciência e a Tecnologia under grant SFRH/BD/37465/2007.

22 **REFERENCES**

23 Almeida, C. M., J. M. Gleriani, E. F. Castejon, and B. S. Soares-Filho, 2008, "Using neural

1 networks and cellular automata for modelling intra-urban land-use dynamics" *International*
2 *Journal of Geographical Information Science* **22**(9) 943 - 963

3 Barredo, J., M. Kasanko, N. McCormick, and C. Lavallo, 2003, "Modelling dynamic spatial
4 processes: simulation of urban future scenarios through cellular automata" *Landscape and*
5 *Urban Planning* **64**(3) 145-160

6 Barredo, J. I. and L. Demicheli, 2003, "Urban sustainability in developing countries'
7 megacities: modelling and predicting future urban growth in Lagos" *Cities* **20**(5) 297-310

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

10 Batty, M. and Y. Xie, 1997, "Possible urban automata" *Environment and Planning B:*
11 *Planning and Design* **24**(2) 175-192

12 Benenson, I. and P. M. Torrens, 2004, *Geosimulation - Automata-based modeling of urban*
13 *phenomena* (John Wiley & Sons Ltd, Chichester)

14 Caruso, G., D. Peeters, J. Cavailhès, and M. Rounsevell, 2007, "Spatial configurations in a
15 periurban city. A cellular automata-based microeconomic model" *Regional Science and*
16 *Urban Economics*, **37**(5) 542-567.

17 Chesire, P. and D. Hay, 1985, "Problemas de declive y crecimiento en las ciudades de
18 Europa" *Estudios Territoriales* **19** 31-56

19 Chesire, P. and D. Hay, 1989, *Urban Problems in Western Europe: An Economic Analysis*
20 (HarperCollins Publishers Ltd, London)

21 Clarke, K., S. Hoppen, and L. Gaydos, 1997, "A self-modifying cellular automaton model of
22 historical urbanization in the San Francisco Bay area" *Environment and Planning B: Planning*
23 *and Design* **24**(2) 247-261

- 1 Cohen, B., 2004, "Urban growth in developing countries: A review of current trends and a
2 caution regarding existing forecasts" *World Development*, **32**(1) 23-51.
- 3 Couclelis, H., 1985, "Cellular worlds - a framework for modeling micro-macro dynamics"
4 *Environment and Planning A* **17**(5) 585-596
- 5 Couclelis, H., 1988, "Of mice and men - what rodent populations can teach us about complex
6 spatial dynamics" *Environment and Planning A* **20**(1) 99-109
- 7 Couclelis, H., 1997, "From cellular automata to urban models: New principles for model
8 development and implementation" *Environment and Planning B: Planning and Design* **24**(2)
9 165-174
- 10 Couto, P., 2003, "Assessing the accuracy of spatial simulation models" *Ecological Modelling*
11 **167**(1-2) 181-198
- 12 Diappi, L. and P. Bolchi, 2006, "Gentrification waves in the inner-city of Milan", in
13 *Innovations in Design & Decision Support Systems in Architecture and Urban Planning* Eds J.
14 van Leeuwen and H. Timmermans (Springer, Dordrecht, The Netherlands) pp 187-201
- 15 Diappi, L. and P. Bolchi, 2008, "Smith's rent gap theory and local real estate dynamics: A
16 multi-agent model" *Computers, Environment and Urban Systems* **32**(1) 6-18
- 17 Ferras, C., 2007, "El enigma de la contraurbanización: Fenómeno empírico y concepto
18 caótico" *EURE Journal of Latin American Urban and Regional Studies* **33**(98) 5-25
- 19 INE, 2001, *Censos 2001*(Instituto Nacional de Estatística, Lisbon)
- 20 Jantz, C. A., S. J. Goetz, and M. K. Shelley, 2004, "Using the SLEUTH urban growth model
21 to simulate the impacts of future policy scenarios on urban land use in the
22 Baltimore-Washington metropolitan area" *Environment and Planning B: Planning and*
23 *Design* **31**(2) 251-271

- 1 Jenerette, G. D. and J. Wu, 2001, "Analysis and simulation of land-use change in the central
2 Arizona - Phoenix region, USA" *Landscape Ecology* **16**(7) 611-626
- 3 Kennedy, J., 1997. "The particle swarm: Social adaptation of knowledge" in Proceedings of
4 1997 *IEEE International Conference on Evolutionary Computation (ICEC '97)*, Indianapolis,
5 IN, 303-308
- 6 Lee, D., 1973, "Requiem for large-scale models" *Journal of the American Planning*
7 *Association* **39**(3) 163-178
- 8 Li, X. and A. Yeh, 2001, "Calibration of cellular automata by using neural networks for the
9 simulation of complex urban systems" *Environment and Planning A* **33**(8) 1445-1462
- 10 Li, X. and A. G. O. Yeh, 2000, "Modelling sustainable urban development by the integration
11 of constrained cellular automata and GIS" *International Journal of Geographical Information*
12 *Science* **14**(2) 131-152
- 13 Moreno, N., A. Ménard, and D. J. Marceau, 2008, "VecGCA: a vector-based geographic
14 cellular automata model allowing geometric transformations of objects" *Environment and*
15 *Planning B: Planning and Design* **35**(8) 647-665
- 16 Moreno, N., F. Wang, and D. J. Marceau, 2009, "Implementation of a dynamic neighborhood
17 in a land-use vector-based cellular automata model" *Computers, Environment and Urban*
18 *Systems* **33**(1) 44-54.
- 19 O'Sullivan, D., 2001a, "Exploring spatial process dynamics using irregular cellular automaton
20 models" *Geographical Analysis* **33**(1) 1-18
- 21 O'Sullivan, D., 2001b, *Graph-based cellular automaton models of urban spatial processes*,
22 PhD dissertation, Bartlett School of Architecture and Planning, University College, London
- 23 Parsopoulos, K. E., and M. N. Vrahatis, 2002, "Recent approaches to global optimization

- 1 problems through Particle Swarm Optimization" *Natural Computing* **1** 235–306
- 2 Semboloni, F., 1997, "An urban and regional model based on cellular automata" *Environment*
3 *and Planning B: Planning and Design* **24**(4) 589-612
- 4 Semboloni, F., 2000, "The growth of an urban cluster into a dynamic self-modifying spatial
5 pattern" *Environment and Planning B: Planning and Design* **27**(4) 549-564
- 6 Shi, W. and M. Y. C. Pang, 2000, "Development of Voronoi-based cellular automata -an
7 integrated dynamic model for Geographical Information Systems" *International Journal of*
8 *Geographical Information Science* **14**(5) 455 - 474
- 9 Silva, E. and K. C. Clarke, 2002, "Calibration of the SLEUTH urban growth model for Lisbon
10 and Porto, Portugal" *Computers, Environment and Urban Systems* **26**(6) 525-552
- 11 Stevens, D. and S. Dragicevic, 2007 "A GIS-based irregular cellular automata model of
12 land-use change" *Environment and Planning B: Planning and Design* **34**(4) 708-724
- 13 Stevens, D., S. Dragicevic, and K. Rothley, 2007, "iCity: A GIS-CA modelling tool for urban
14 planning and decision making" *Environmental Modelling & Software* **22**(6) 761-773
- 15 Straatman, B., R. White, and G. Engelen, 2004 "Towards an automatic calibration procedure
16 for constrained cellular automata" *Computers, Environment and Urban Systems* **28**(1-2),
17 149-170
- 18 Tobler, W., 1970, "A Computer Movie Simulating Urban Growth in the Detroit Region",
19 *Economic Geography*, **4**(2) 234-240.
- 20 Tobler, W., 1979, "Cellular geography", in *Philosophy in Geography* Eds S. Gale and G.
21 Olsson (D. Reidel, Boston) pp 379-386
- 22 Trelea, I. C., 2003, "The particle swarm optimization algorithm: convergence analysis and

1 parameter selection" *Information Processing Letters*, **85** 317–325.

2 Uhel, R., D. Ludlow, J. Fons, N. Blanes, O. Gómez, H. Savolainen, M. Kasanko, J. Barredo,
3 C. Lavallo, L. Petrov, and V. Sagris 2006, "*Urban sprawl in Europe: The ignored challenge*",
4 Joint Research Centre of the European Commission and European Environment Agency,
5 Copenhagen, Denmark

6 Vancheri, A., P. Giordano, D. Andrey, and S. Albeverio, 2008, "Urban growth processes
7 joining cellular automata and multiagent systems. Part 2: computer simulations" *Environment*
8 *and Planning B: Planning and Design* **35**(5) 863-880

9 Vandergue, D., J.-P. Treuil, and D. Drogoul, 2000, "Modelling urban phenomena with
10 cellular automata", in *Applications of Simulation to Social Science* Eds G. Ballot and G.
11 Weisbuch (Hermes Science Publishing, Stanmore, Middlesex) pp 127-140

12 Vliet, J. V., R. White, & S. Dragicevic, 2009, "Modeling urban growth using a variable grid
13 cellular automaton" *Computers, Environment and Urban Systems* **33**(1) 35-43.

14 Ward, D., A. Murray, and S. Phinn, 2000, "A stochastically constrained cellular model of
15 urban growth" *Computers, Environment and Urban Systems* **24**(6) 539-558

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

18 Webster, C. J. and F. Wu, 1999a, "Regulation, land-use mix, and urban performance. Part 1:
19 theory" *Environment and Planning A* **31**(8) 1433-1442

20 Webster, C. J. and F. Wu, 1999b, "Regulation, land-use mix, and urban performance. Part 2:
21 simulation" *Environment and Planning A* **31**(9)1529-1545

22 White, R. and G. Engelen, 1993, "Cellular automata and fractal urban form: a cellular
23 modelling approach to the evolution of urban land-use patterns" *Environment and Planning A*

1 **25**(8) 1175-1199

2 White, R. and G. Engelen, 1997, "Cellular automata as the basis of integrated dynamic
3 regional modelling" *Environment and Planning B: Planning and Design* **24**(2) 235-246

4 White, R., G. Engelen, and I. Uljee, 1997, "The use of constrained cellular automata for
5 high-resolution modelling of urban land-use dynamics" *Environment and Planning B:
6 Planning and Design* **24**(3) 323-343

7 Wu, F., 2002, "Calibration of stochastic cellular automata: the application to rural-urban land
8 conversions" *International Journal of Geographical Information Science* **16**(8) 795–818

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

12 Wu, F. L. and C. J. Webster, 2000, "Simulating artificial cities in a GIS environment: urban
13 growth under alternative regulation regimes" *International Journal of Geographical
14 Information Science* **14**(7) 625-648

15 Xian, G. and M. Crane, 2005, "Assessments of urban growth in the Tampa Bay watershed
16 using remote sensing data" *Remote Sensing of Environment* **97**(2) 203-215

17

18

1 **Figure Captions**

2 Figure 1 - Generic neighborhood effect relationships

3 Figure 2 - Location and land use map (2001) for the municipality of Condeixa-a-Nova

4 Figure 3 - Irregular cells in the center of Condeixa-a-Nova

5 Figure 4 - Condeixa-a-Nova calibration maps

6 Figure 5 - CA model (grey) and PS algorithm flowchart

7 Figure 6 – Evolution of three CA model parameters during the application of PS

8 Figure 7 - Details of urban change in Condeixa-a-Nova main town

9 Figure 8 - Neighborhood effect relationships in Condeixa-a-Nova

10 Figure 9 - Condeixa-a-Nova forecast maps

11

12 **Table Headings**

13 Table 1 - Density thresholds for urban cell states

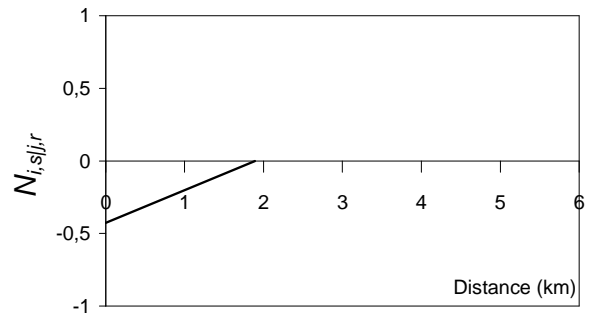
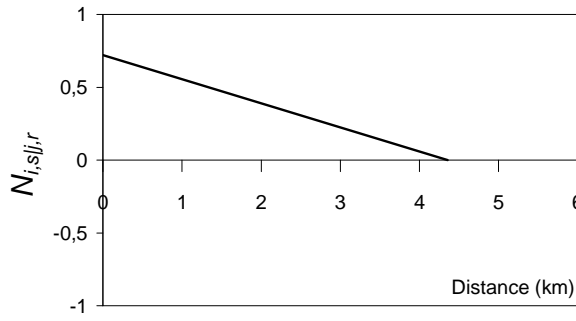
14 Table 2 - Contingency matrices for Condeixa-a-Nova

15

16

17

1
2
3



4

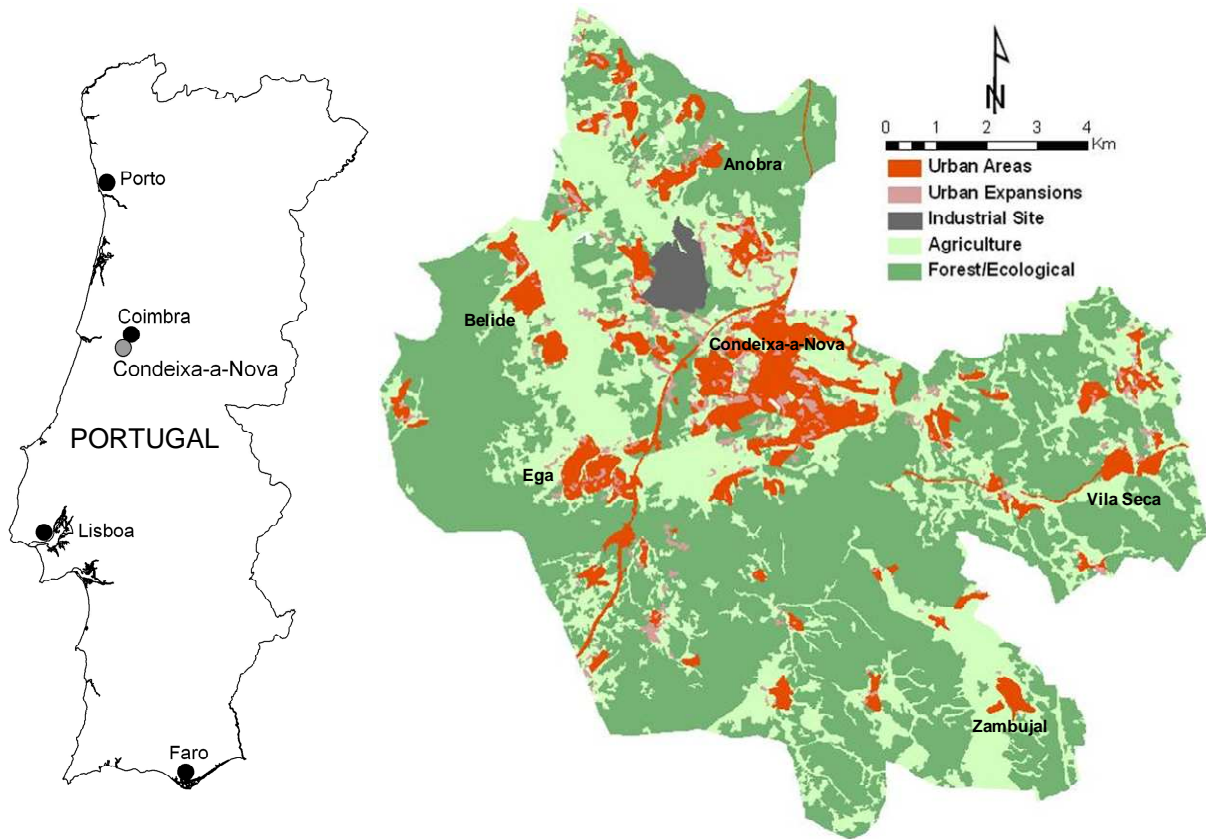
(a) Attraction

(b) Repulsion

5
6
7
8
9

Figure 1 - Generic neighborhood effect relationships

1
2



3

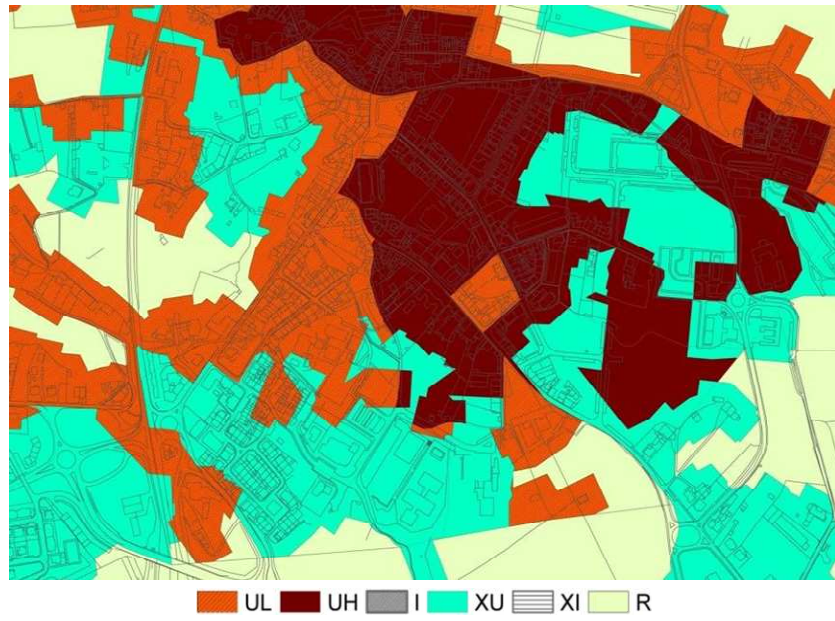
4 Figure 2 - Location and land use map (2001) for the municipality of Condeixa-a-Nova

5

6

7

1



2

Figure 3 - Irregular cells in the center of Condeixa-a-Nova

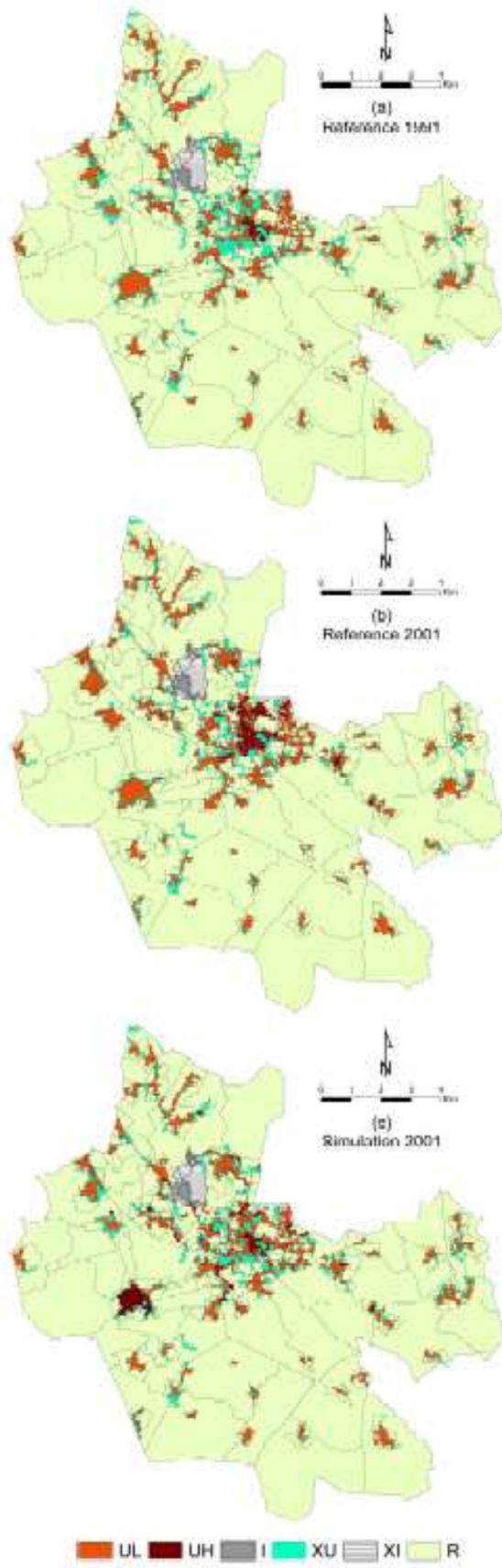
3

4

5

6

7

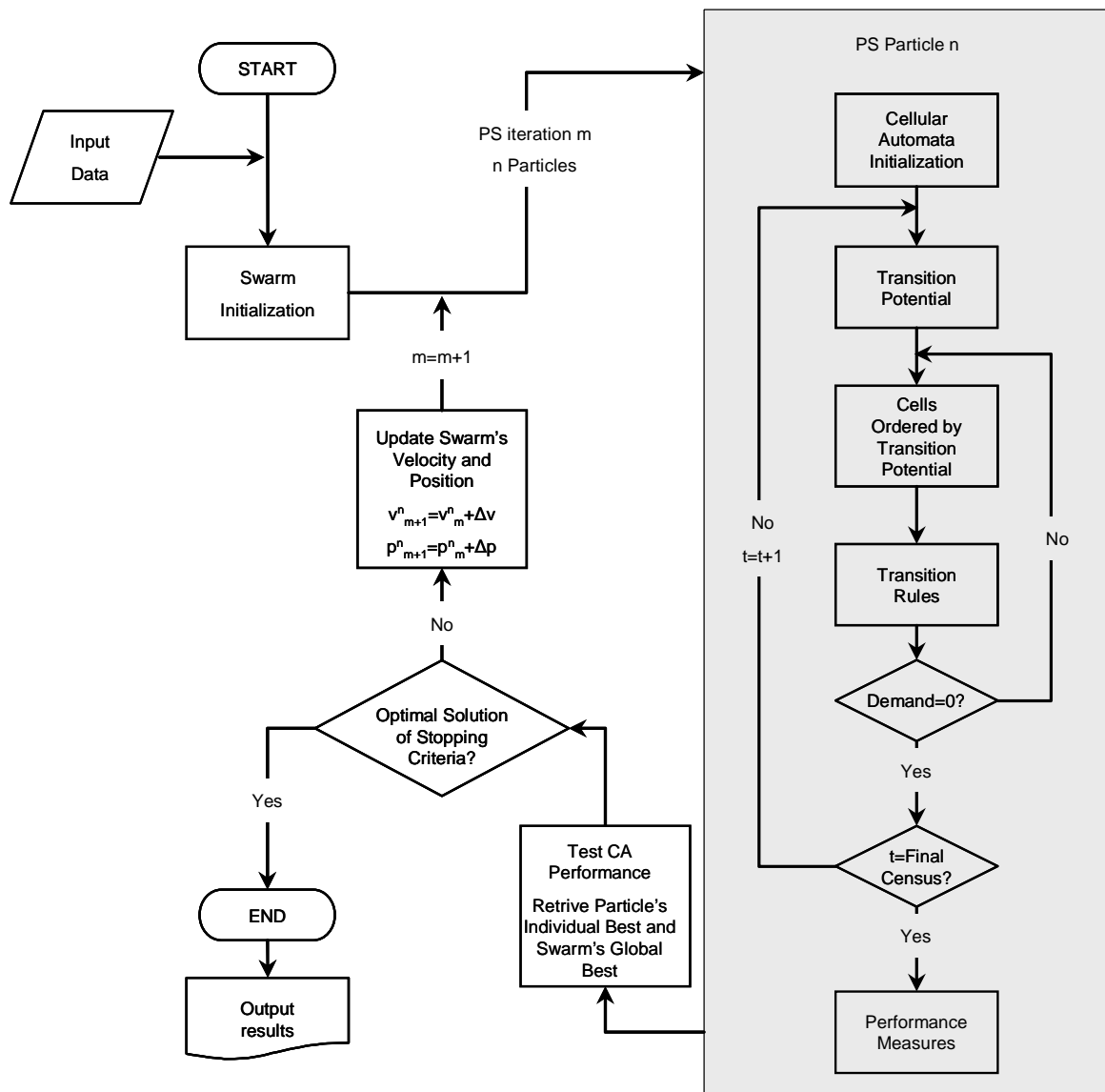


1

2

3

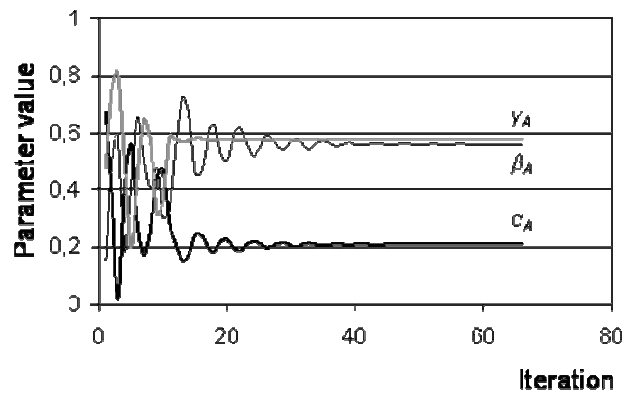
Figure 4 - Condeixa-a-Nova calibration maps



1
2
3
4
5

Figure 5 - CA model (grey) and PS algorithm flowchart

1



2

3

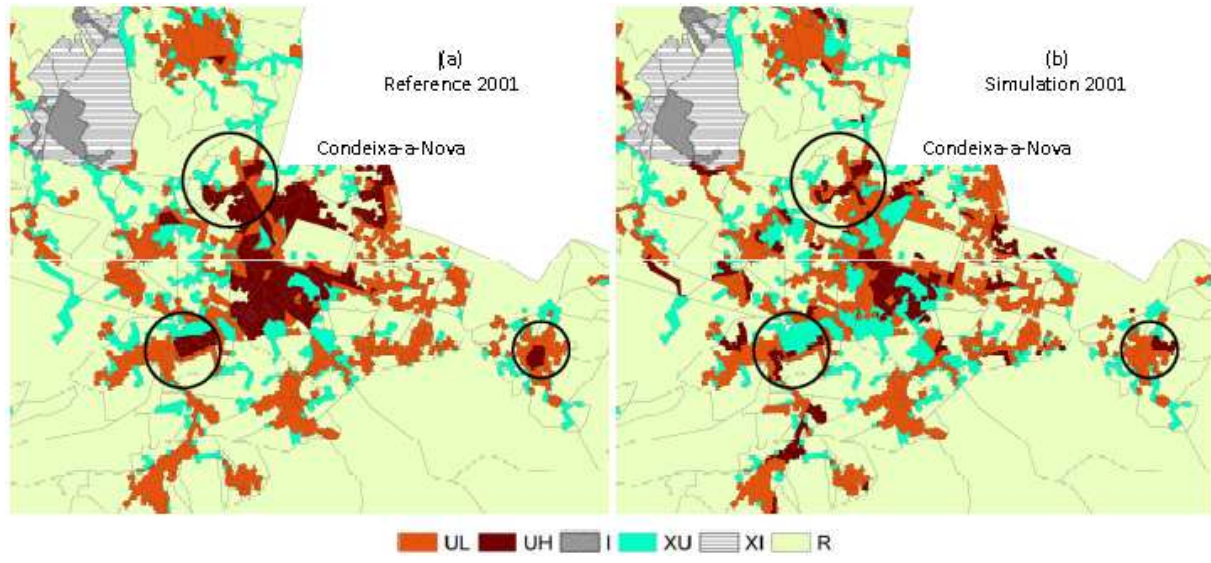
Figure 6 – Evolution of three CA model parameters during the application of PS

4

5

6

1



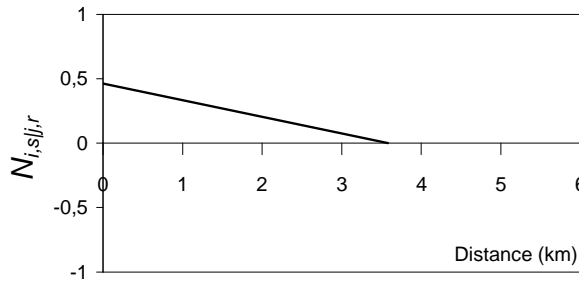
2

3

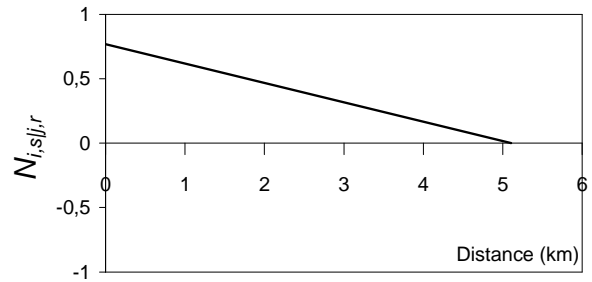
Figure 7 - Details of urban change in Condeixa-a-Nova main town

4

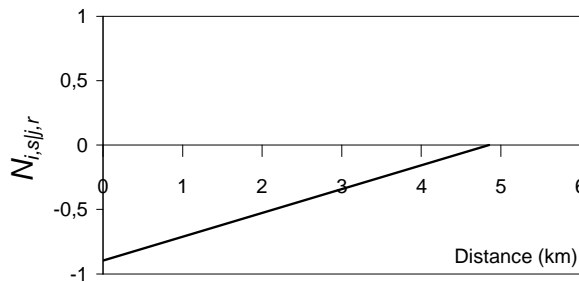
5



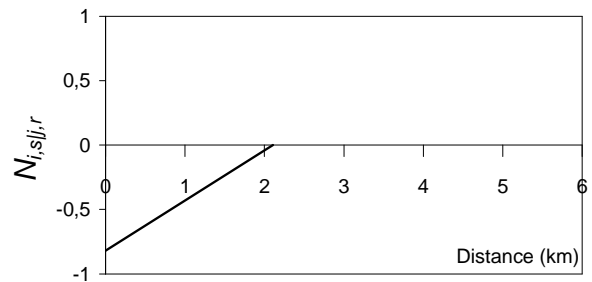
— State UL and State UH



— State XI and State I



— State UH and State I



— State XI and State R

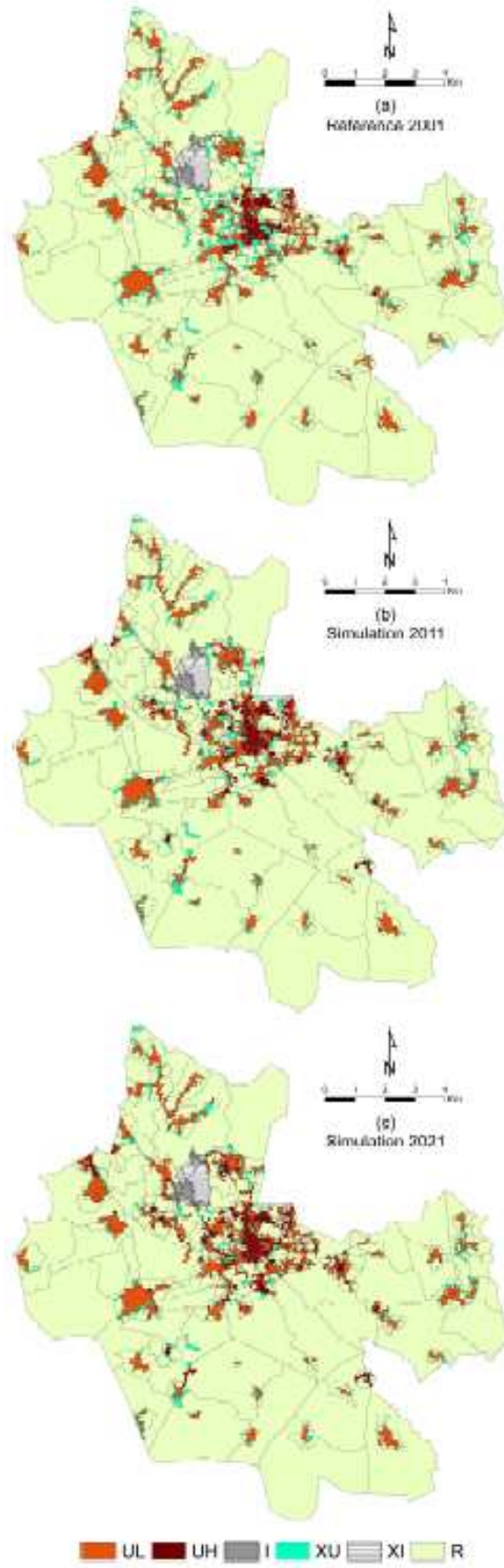
1

2

Figure 8 - Neighborhood effect relationships in Condeixa-a-Nova

3

1



2

3

4

Figure 9 - Condeixa-a-Nova forecast maps

1

Table 1 - Density thresholds for urban cell states

State	Year	Household Size (Inhabitants)	Construction Density (Dwellings per hectare)	Population Density (Inhabitants per hectare)
UL	1991	2,5	10	13,2
	2001	2,2	6	6,6
UH	1991	2,5	5	25,0
	2001	2,2	3	12,5

2

3

4

5

1

Table 2 - Contingency matrices for Condeixa-a-Nova

Simulation Map	Reference Map					Sum	Conditional k_{mod}
	UL	UH	I	XU	XI		
UL	328	43	0	20	0	391	0,718
UH	33	40	0	18	0	91	0,362
I	0	0	3	0	2	5	0,598
XU	74	40	0	402	0	516	0,610
XI	0	0	2	0	10	12	0,831
Sum	435	123	5	440	12		

(a) Number of Cells - Overall Agreement 0,771

Simulation Map	Reference Map					Sum	Conditional k_{mod}
	UL	UH	I	XU	XI		
UL	735,9	66,6	0,0	27,9	0,0	830,4	0,732
UH	125,0	45,7	0,0	22,0	0,0	192,7	0,161
I	0,0	0,0	25,4	0,0	5,7	31,1	0,809
XU	75,4	34,6	0,0	339,5	0,0	449,5	0,678
XI	0,0	0,0	39,9	0,0	119,2	159,1	0,728
Sum	936,3	146,9	65,3	389,4	124,9		

(b) Area in hectares - Overall Agreement 0,778

2

3