

Go East: A Residential Land Use Model for the Periphery of Rome

G. L. Ciampaglia

MACS Lab., University of Lugano,
V.le Canavée, CH-6850 Mendrisio, Switzerland
(ciampagg@lu.unisi.ch)

B. Tirozzi

Dept. of Physics, University of Rome “La Sapienza”,
P.le Aldo Moro 8, 00100 Rome, Italy
(b.tirozzi@libero.it)

January 27, 2009

Abstract

In this paper we present a model of urban growth and its preliminary application to a case study of the phenomena of residential development in the setting of the eastern periphery of Rome, Italy’s capital city. The modeling approach we use synthesises the two typical paradigms widespread in the community of quantitative urban planning: the traditional one, based on cellular automata (CA), and the (relatively) new one, which is agent-based. In particular, our multi-agent system (MAS) is in-between a reactive MAS, with agents carrying out a two-staged decision process in a complex environment, and a model of statistical physics, since we use populations of agents in order to reduce the number of degrees of freedom of the system. While we explicitly model the consumption of agricultural and undeveloped land due to urban growth, our model may be easily integrated as a socio-economic part into a wider decision support system for environmental planning, e.g. our simulations can produce indicators of environmental impact of the growth of the city: electricity consumption, waste production, etc.

1 INTRODUCTION

It is now a matter of fact that the complex systems perspective has been fully accepted in urban and regional planning studies. Since their introduction, the two major bottom-up approaches, the one of multi-agent systems (see [2, 3]) and cellular automata (see [11, 4]) have contended for the lion’s share of the literature, with the current trend to synthesize the best from both approaches. The work we present in this paper is based on one recent attempt to provide such an interpolation (see [10] for a primer to this methodology, and the references therein) and deals with the development of a major urban area characterized by sprawling phenomena and unregulated residential growth for many decades. From a methodological point of view, connecting MAS with CA is attractive because it turns out to be very easy to model a urban or geographical system with such a combination of approaches. The baseline is to identify “mobile” entities

in the cognitive agents, such as the inhabitants of a city and to subdivide, on the other hand, the urban area in a regular lattice. One then models more or less faithfully the behaviour of the agents, usually at the cognitive level, while assigning to each cell a label describing the urban typology of the part of territory assigned to it. This procedure usually takes the form of the analysis, design and implementation of a software simulator of the complex interactions between agents and the cells, and the simulations produced with such a tool may provide valuable informations of the phenomena one wants to investigate. It is widely accepted that the introduction of multi-agent systems solved the problem of providing a meaningful description of the processes of land use change undergoing in the system, without the need to describe these in terms of abstract interactions between the cells of a CA: it is much more convenient to describe in terms of the agents those socio-economical interactions that – at least one empirically assumes – drive the urban transformations of a city, rather than account for the existence of an ubiquitous process of “update”, only indirectly based on the same interactions, and only explainable in terms of the local state of the neighbourhoods on the lattice of cells. However, this combined approach has disadvantages, no matter the degree of fidelity one can achieve with a well engineered software simulator. It is in fact difficult to calibrate and validate the simulations produced with it, and many times a proper sensitivity analysis is only able to identify critical parameters of the model, but not to truly assess what the uncertainty in the simulations’ measurements is due, especially in presence of bifurcations or phase transitions of the original system. It seems thus important to provide a modelling framework that gives a proper mathematical definition of the entities one is going to model, and that at the same time contemplates the possibility to build a software simulator that cheaply outputs interesting scenarios about the future evolution of the city.

The work we present in this paper is an attempt to go in this direction. The structure of the paper is the following: in section 2 we introduce the main mathematical features of our model, and in particular in 2.1 we detail respectively how the collective decisions of the agents can produce the stochastic dynamics for the evolution of the cellular automaton, and in section 2.2 we look at how it is possible to synthesize a description of the city that enables each agent to take realistic decisions about urban events. In section 3 we discuss the preliminary results we got from running simulations under the assumptions of a stationary dynamics for the configuration of the system. Finally, in section 4 we discuss how the integrated approach to the production of computer tools for environmental assessment and decision making could benefit from our simple methodology.¹

2 A spatial model of urban growth and intelligent agents

To a first approximation, with the CA we model the land uses of the urban system. As is usual for this kind of models, each cell of the CA is a 2-dimensional representation of a given piece of land belonging to the urban area. Our choice was to consider each cell corresponding to an administrative zone, as used in the master plan of the city; thus, we have an irregular lattice Γ , with adjacency relations between cells given by the actual geographical boundaries between zones. Figure 1 shows the cellular decomposition of the CA used in the case study of Rome.

We take a state space which is real-valued and multidimensional, with each cell $c \in \Gamma$ described by a vector of dynamical variables $v(c, t) \in \mathbb{R}^p$, subject to the update rule of the automaton, and by some control variables that follow an exogenous

¹For additional material on this model, please see the following: <http://www.inf.unisi.ch/phd/ciampaglia>

Variable	Description	Update rule
N	Total number of free available flats in the cell	$N + \Delta n$
S_u	Total undeveloped surface in the cell	$S_u - (\Delta s_1 + \Delta s_2)$
S_b	Total built surface due to residential use	$S_b + \Delta s_1$
S_p	Total built surface due to parkings and paved use	$S_p + \Delta s_2$

Table 1: Update rule of the state of a cell due to the construction of an apartment building

dynamics, $\mathbf{w}(c, t) \in \mathbb{R}^q$. Dynamical variables are chosen to have a clear urban and geographical meaning. Examples include: number of households, total surface area of residential buildings, total surface area of empty land. In \mathbf{w} there are instead variables that are not directly subject to the update rule of the automaton: distances from hospitals and services, access points to the transportation network, and other features related to infrastructure. The variables in $\mathbf{w}(c, t)$ can be used to define exogenously the settings for a simulative scenario. Taking \mathbf{v} and \mathbf{w} together for each cell c , the system is described by a n by d matrix $\mathbf{X}(t)$ where $n = |\Gamma|$ and $d = p + q$. The rule that updates the state of the cells of the automaton is not specified as a deterministic function, with fixed adjacency neighbourhood, as in the classical specification of CA. Instead, the updates of the cells' state are performed by decisor agents: the basic idea is that social entities interact locally with the city – that is, agents interact with cells – with possible interactions stereotyped by the elements of a set \mathcal{A} of kinds of interactions.² in modeling the elements of \mathcal{A} , we have in mind classes of events with a clear urban meaning and scope, e.g.: construction of houses, malls and offices, changes in zoning regulations of free lots. It is thus possible to model very easily the effects of these simple interactions; see table 1 for an example of the update rule used to model the construction of an apartment building.

More formally, each $\alpha \in \mathcal{A}$ defines an incremental stochastic dynamic of the form:

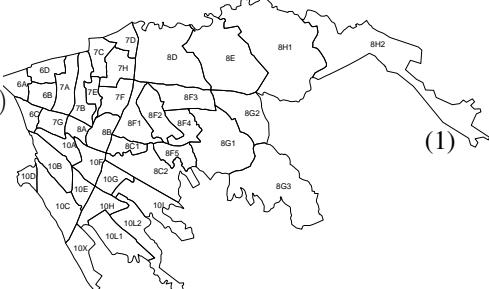
$$v_k(c, t + \Delta t) = v_k(c, t) + \pi_k(\omega, c, t)$$

for $k \in V_\alpha \subseteq \{1, \dots, p\}$, e.g. for some of the variables that make up the components of $\mathbf{v}(c, t)$. We take the standard assumptions so that the counting variables $N_{\alpha, B}(c, t)$ of the number of events of kind α , with increments $\pi \in B \subseteq \mathbb{R}^{n(\alpha)}$, and occurring in c during the time interval $[t, t + \Delta t]$, have law de-

fined by the Poisson distribution with parameter $\lambda_{\alpha, B}(c, t)$. Thus we need to compute the intensity of a non-homogeneous Poisson process (see [7]); the idea is to define it in terms of a density λ^α :

$$\lambda_{\alpha, B}(c, t) = \int_B \lambda^\alpha(c, \pi, t) d\pi \quad (2)$$

²we say “social entity” since an agent may also be e.g. an enterprise looking for a venue where to open a new office.



where $\lambda^\alpha(c, \pi, t)$ is the density of probability that in $[t, t + \Delta t]$, with Δt small enough, one interaction of kind α with increments in $[\pi, \pi + \Delta\pi]$ occurs in c . We then introduce the conditional probability $\beta^\alpha(c, \pi, t)$, that one event has increments in $[\pi, \pi + \Delta\pi]$, given it actually occurred. Then obviously $\lambda^\alpha(c, \pi, t) = \lambda^\alpha(c, t) \cdot \beta^\alpha(c, \pi, t)$, where

$$\lambda^\alpha(c, t) = \int_{\mathbb{R}^{n(\alpha)}} \lambda^\alpha(c, \pi, t) d\pi \quad (3)$$

This means that we are able to decouple the problem of modelling an interaction of kind α into two sub-models: the first accounts for how many interactions occur in a cell during a small time interval, the second allows us to generate the values that characterize the event that results from an interaction between an agent and a cell. If we consider Δt to be the time unit, then $\lambda_{\alpha, B}(c, t)\Delta t$ is the average number of events per time unit. Of course, in the spirit of CA modeling, this quantity has to depend on the local configuration of the system, where the concept of ‘locality’, for any given α , plays the same role of that of neighbourhood in the classic definitions of CA. A reasonable assumption for this class of systems is that, for values of Δt small enough, the information about the events occurring in a cell during Δt does not change the configuration of the system, so that the intensities are constant during the time interval. This assumption let us to recover a time homogeneous process for each step of update of the system, which is of great aid in the simulation of the process.

2.1 Decision dynamics of the multi-agent system

We now briefly detail the first of the two sub-models, which enables us to compute $\lambda^\alpha(c, t)$. Each α has a population of decisor agents. Unlike those MAS where each agent is equipped with a complex agenda and models faithfully a given sociological class of individuals, we take the opposite approach and reduce the degree of freedoms of the system by describing populations of temporary and anonymous agents. An agent may be inactive, or may be active and locatable in a specific zone, e.g. resident in a cell. The decision process of an agent is composed of four actions or steps: activation (A), diffusion (D), update (U) and leaving (L). After activation, an agent ‘enters’ into the automaton and is placed in a cell. From there it may either diffuse – that is, jump – to another cell, update the state of the cell by realizing an event of kind α (and subsequently become inactive and ‘exit’ the CA), or leave the decision process and become inactive, with the same consequences of the update step. We define four processes, at the global level of the whole automaton, for the above actions. Let these processes have each intensity Λ_i^α , for $i \in \{A, L, D, U\}$. This means that, as an example, Λ_A^α is the average number of agents belonging to the population of agents α that become active in any cell of the CA per time unit. The idea is to consider the global intensities Λ_i^α as parameters of the model and then, thanks to the property of composition of Poisson processes, distribute (or, generally speaking, assign) the overall rate of activations, diffusion, etc. among the cells. The following four formulas explain how this idea is put in practice:

$$\lambda_A^\alpha(c, t) = \Lambda_A^\alpha(t) \cdot \frac{F^\alpha(c, t)}{\sum_{c' \in \Gamma} F^\alpha(c', t)} \quad \lambda_D^\alpha(c, t) = \Lambda_D^\alpha \cdot \frac{F^\alpha(c, t)}{\sum_{c' \in \Gamma} F^\alpha(c', t)} \quad (4)$$

$$\lambda_L^\alpha(c, t) = \Lambda_L^\alpha \quad \lambda_U^\alpha(c, t) = \Lambda_U^\alpha \cdot G^\alpha(c, t) \quad (5)$$

The first formula says that the intensity for activations of agents in a cell c is distributed proportionally to a global ‘attractivity’ force $F^\alpha(c, t)$. Intuitively speaking, $F^\alpha(c, t)$

models how much the cell c is favorable to the occurrence of interactions of kind α due to the regional context in which c is: it integrates informations at a macroscopic or regional level, such as residential centrality in the case when α is the interaction of renting a flat; a similar definition has been used in the second formula for the intensity $\lambda_D^\alpha(c, t)$, of jumps that have cell c as a target, that is, regardless of the origin cell. The effect of these two definitions is to bias the exploration of the urban space, that the agents do during their life cycle, towards those areas of the city that exercise more force of attraction than others. The dependency of $\Lambda_A^\alpha(t)$ on time, as we shall see later, lead us to manipulate the number of events of kind α that actually occur during a simulation. The third formula defines the average time an agent spends in the active state to be $1/\Lambda_L^\alpha$. The intensity $\lambda_U^\alpha(c, t)$ is proportional to another attractivity force called $G^\alpha(c, t)$, which integrates information on the state of c , as usual with respect to α , at a more local and detailed level.

Before getting into the discussion on our approach to the modeling of the configuration of the system, that is, on the definition of F^α and G^α , we have to explain how, starting from the intensities of the local processes, the above definitions are brought together and give the overall dynamics of the MAS along a simulation's step $[t, t + \Delta t]$. In turn, this let us to compute the intensities $\lambda^\alpha(c, t)$ of the non-homogeneous processes of update events of the CA's state. If we analyse the number of agents entering and leaving a cell due to the four actions just introduced, we can derive a first order differential equation for the probability $P^\alpha(c, t)$ on the space of the states of an agent:

$$\frac{dP^\alpha}{dt} = \frac{F^\alpha(c, t)}{T^\alpha(t)} \cdot \left(P^\alpha(\bar{\Gamma}, t) \cdot \Lambda_A^\alpha(t) + (1 - P^\alpha(\bar{\Gamma}, t)) \cdot \Lambda_D^\alpha \right) + \\ - P^\alpha(c, t) \cdot \left(\Lambda_D^\alpha + \Lambda_L^\alpha + \Lambda_U^\alpha \cdot G^\alpha(c, t) \right) \quad (6)$$

Where $T^\alpha(t) = \sum F^\alpha(c, t)$ is the normalizing factor, and $P^\alpha(\bar{\Gamma}, t)$ is the probability that an agent is passive and thus not in any cell of Γ . It is then possible to find a solution, for the stationary case $dP^\alpha/dt = 0$:

$$P^\alpha(c, t) = \frac{F^\alpha(c, t) \cdot \left(P^\alpha(\bar{\Gamma}, t) \cdot \Lambda_A^\alpha(t) + (1 - P^\alpha(\bar{\Gamma}, t)) \cdot \Lambda_D^\alpha \right)}{T^\alpha(t) \cdot (\Lambda_D^\alpha + \Lambda_L^\alpha + \Lambda_U^\alpha \cdot G^\alpha(c, t))} \quad (7)$$

The assumption of stationarity is motivated by the fact that the configuration of the system is determined by variables that, with respect to P^α , have a slow dynamics. Now, since we are dealing with Poisson processes, $\lambda_U^\alpha(c, t)$ approximates the probability that one event occurs in a small time step of duration Δt ; thus $\lambda_U^\alpha(c, t) \cdot P^\alpha(c, t)$ gives the probability that an active agent performs the update step in the cell – that is, that an event occurs – and if we multiply that by the number of agents in the population, we obtain the average number of events of kind α occurring in $[t, t + \Delta t]$. For this construction to work properly, it is important to ensure that the number of events does not fluctuate too much around this expected value. This condition is satisfied, thanks to the law of large numbers, in the limit of the number of agents growing to infinity. In practical simulations, this condition is usually fulfilled with a number of active agents on the order of the thousands.

2.2 The urban configuration

As already stated, $G^\alpha(c, t)$ is the piece of information with which active agents decide whether to make an update in the current cell or not. In [9], from which our work

inherits the basic modeling framework, there is not a single and general model for such a force; instead, the authors use fuzzy decision theory and develop several models of G^α , one for each kind of event $\alpha \in A$, by refining and aggregating multiple indicators – by and large, demographic and geographic data – with the aid of fuzzy t -norms and t -conorms. This approach is very powerful when empirical models of the different α are available: in such cases fuzzy modeling is indeed a suitable tool for the translation in the formal language of mathematics.

The definition of the force of attraction $G^\alpha(c, t)$ we give is inspired by principal component analysis (PCA). PCA is a widely used technique for multivariate analysis (see [6]), and can be also viewed as a simple form of unsupervised learning.³ A common data analysis task that can be done with PCA is the identification, for each principal component, of a subset of the original variables that have highest correlation in absolute terms with that component, beyond a certain threshold (see [5]). Usually human experts are able to synthesize meaningful indices using the principal components. In our case, if we look at how a kind of interaction α is defined, there's already a subset of variables having a special status with respect to α . These variables have their dynamical behaviour influenced by interactions of kind α , that is, those with index $k \in V_\alpha$ in equation (1). One can compute the cell's scores, for these variables, in the space spanned by the principal components, and assess the cell's attractivity towards the kind of interactions α by looking at those scores. This is, conceptually, the opposite of the operation of principal components' identification stated above: we choose to model the force of attraction in terms of a fixed subset of variables, and then PCA automatically synthesizes an index that measures how much any cell is attractive, with respect to those variables. This technique has some disadvantages – which we'll discuss later – but it allows for a simple and general purpose model of the urban forces G^α .

Let $\mathbf{X} = \mathbf{X}(t)$ be the dynamic matrix of multivariate data we apply PCA to.⁴ \mathbf{X} is constituted of $n = |\Gamma|$ samples or data points, one for each cell c , and of $d = p + q$ observations of demographic and economic variables, that is, the concatenation of vectors $\mathbf{v}(c, t)$ and $\mathbf{w}(c, t)$. The principal components are standardized linear combinations of the urban variables, $\mathbf{y}_k = \mathbf{a}'_k \mathbf{X}$. Since the variables have different units of measure, we compute the coefficients \mathbf{a}_k by diagonalization of the sample correlation matrix \mathbf{R} of the data, $\mathbf{R} = \mathbf{A}\mathbf{D}\mathbf{A}'$: the \mathbf{a}_k are then the orthogonal columns of \mathbf{A} . Since the sample points are centered around their mean μ , for the i -th cell it is possible to compute the projection on the k -th component component as $y_{ik} = \mathbf{a}'_k(\mathbf{x}_i - \mu)$. In the terminology of PCA, y_{i1}, \dots, y_{id} are called the scores of the i -th sample. These 'raw' scores, however, are not suitable for computing an index. First, the original variable we take into account might correlate negatively with the component. Since our data are centered, and the transformation induced by the PCA is just a rotation of the space, values of that variable that are less than the mean have actually negative score on the component, while we want them to give a positive contribution to the total score of the cell. A symmetric argument holds if the correlation is positive. Moreover, it is desirable to take into account how much of the variance of the original variable the principal component is able to explain. Finally, since we are going to use this index to define an

³A very good review is [8], in which PCA is presented as a learning problem in a linear model with latent variables, under the assumption that the hidden state is constant and constituted of independent normal variables and that the linear dependency of the observations on the state is affected by additive Gaussian noise with infinitesimal variation.

⁴Matrices are denoted by upper case bold letters, vectors are always column vectors and are denoted with bold lower case letters; thus $\mathbf{a}'\mathbf{b}$ is the scalar product of two vectors. When referencing matrices, i, j are the row/column indices for the original data, while h, k always refer to the principal components.

intensity, which is positive definite, we want each score to give a positive contribution as well. Since it is possible to compute the correlation between the j -th variable and the k -th component as $l_{jk} = a_{jk}\sqrt{\lambda_k}$, then the above considerations lead us to define the contribution that each score gives as:

$$y_{ijk}^* = \frac{1}{\pi} \tan^{-1} \left(\frac{l_{jk} y_{ik} - \mu_{jk}}{\sigma_{jk}} \right) + \frac{1}{2} \quad (8)$$

so that $0 < y_{ijk}^* < 1$. μ_{jk} and σ_{jk} are the sample mean and standard deviation of the raw scores of the cells of the j -th variable with respect to the k -th principal component. Now let us consider the i -th cell c_i and the kind of interaction α . Let us denote with g^α the weighted sum of the scores (8) on the first $r \leq d$ principal components:

$$g^\alpha(c_i, t) = \sum_{j \in V_\alpha} \sum_{k=1}^r y_{ijk}^* \quad (9)$$

where, as stated, j is restricted to range in V_α , the subset of variables whose dynamics is affected by the update rule of α . r is chosen so that the first r components explain at least 50% of the variance of the original data. Finally, to define the dynamics of the force $G^\alpha(c_i, t)$, we introduce a temporal delay to smooth the changes of the term in (9):

$$G^\alpha(c_i, t + \Delta t) = (1 - \epsilon) G^\alpha(c_i, t) + \epsilon g^\alpha(c_i, t + \Delta t) \quad \text{for } t > 0 \quad (10)$$

$$G^\alpha(c_i, 0) = g^\alpha(c_i, 0) \quad (11)$$

where $0 < \epsilon \leq 1$ acts as a learning rate.

The definition of $F^\alpha(c, t)$ ‘averages’ $G^\alpha(\cdot, t)$ over the regional context of c :

$$F^\alpha(c, t) = \sum_{c' \in \Gamma} i(c') h(d(c, c')) G^\alpha(c', t) \quad (12)$$

$$h(x) = h(x; m, n, h_0) = m \left(1 - \frac{x^n}{x^n + h_0} \right) \quad (13)$$

In (12) the regional context of a cell generalizes the concept of neighborhood of a CA by means of a simple gravitational model: the contribution of each cell c' is proportional to $G^\alpha(c', t)$ (e.g. its mass), and since (13) is a monotonically decreasing step function, decreases with a measurement of the distance between the cells. Such a measurement, which should be taken with respect to the transportation network of the city, is modeled by the term $i(c')h(d(c, c'))$. Usually, the network is explicitly modeled as a labeled digraph, and one takes a suitable graph-theoretic measure of integration of a node in a graph; however, the only data we had for our case study were the distances, from the center of mass of the developed areas of the cell, to the nearest access point to the transportation network (d_{net}), hospital (d_{hos}), university or school (d_{edu}) and major shopping or service center (d_{ser}). If we allow k to range in $\{net, hos, edu, ser\}$, our integration measure is:

$$i(c) = \prod_k h(d_k(c); 2, n, \mu_k) \quad (14)$$

The parameter $n = n_k$ can be set so that the k -th factor $\approx 3/2$ when $\mu_k - d_k(c) \approx \sigma_k$, which equals to reward those cells that are better integrated, and conversely to weight less those that are ‘distant’ from hospitals, schools, etc. Finally, note that, by using in (12) the euclidean distance $d(c, c')$ between the centers of the cells, we are taking the gross assumption of a homogeneous transportation network.

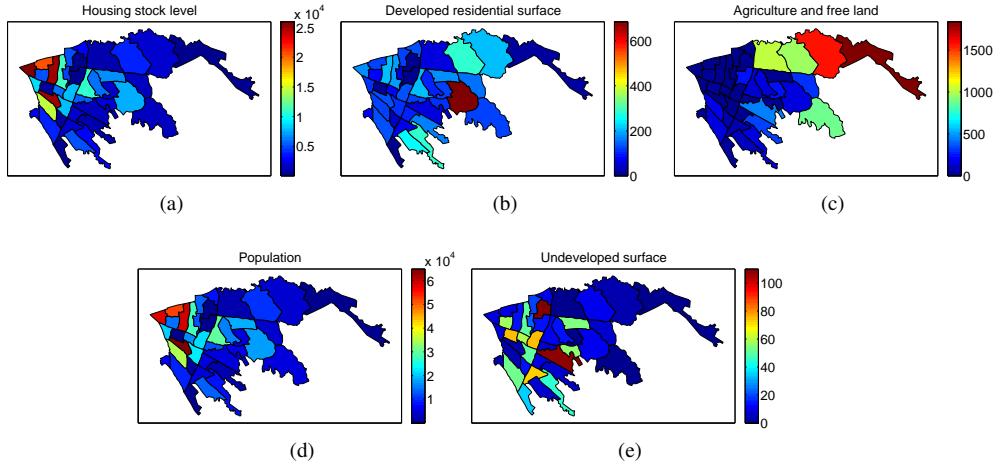


Figure 2: (a)–(e): Initial configuration of the CA. In each figure, the interval between the minimum measurement of the variable and the maximum is mapped on a colormap, such that the relative distribution of values on the cells is portrayed in terms of the difference between “warm” and “cool” colors. (a): Initial level of the housing stock. (b): Initial amount of residential surface. (c): Initial amount of surface due to agriculture or free land. (d): Initial resident population. (e): Initial amount of undeveloped or interstitial surface.

3 Simulation results

The model has been implemented for a case study on the eastern area of the city of Rome, Italy. The CA has $|\Gamma| = 40$ cells, each grossly corresponding to an administrative zone,⁵ and each cell is described by $d = 21 + 6$ demographic and economic variables. Figure 2 shows the initial configuration of the system, taking into account the subset of the dynamical variables in ν related to the phenomena of residential growth. We performed multiple simulations to see if the model was able to show a plausible behaviour with respect to the phenomena of residential growth of the area under study. In these simulations we set $\Lambda_A^\alpha(t) = \Lambda_A^\alpha(0)$ for every t and for every $\alpha \in \mathcal{A}$, so we expect to see a stationary dynamics for the variables of the configuration of the system, a condition easily checkable by inspecting that the trajectory of $G^\alpha(c, t)$, after a transient growth, reaches a steady level. The parameters $\Lambda_L^\alpha, \Lambda_D^\alpha, \Lambda_U^\alpha$ are set to constant values so that agents from any population have, on the average, three jumps to explore the CA before passing to the inactive state. The overall rate of events occurring during a simulation is controlled by setting $\Lambda_A = \sum_\alpha \Lambda_A^\alpha$ to a constant value, and then taking fixed ratios to define the global intensity of activation for each population of agents: e.g. 20% of all activations are from the population of agents looking for a house, etc. We varied Λ_A from 0.1 to 10, and for each value we executed multiple simulations and averaged $\nu(c, t)$ over the simulations. We deem this strategy for the determination of the activation rates to be reasonable, since the city is far from being in a period of ex-

⁵According to the Italian law, the zones we take into account are called “zone urbanistiche” and constitute a refinement of the subdivision of the urban area into municipalities. In our model, we decided to further subdivide some bigger zones into smaller parts, in order to keep them as homogeneous as possible. This subdivision has been made with the aid of planners; see [1].

pansion, and thus the balance between the rates of activity of the different processes is unlikely to change significantly. On the other hand, by varying Λ_A we can explore the parameter space in a consistent way and test the dynamics of the systems at different levels of activity. Figure 3 shows some of these results. The quantities in 3(a)–(e) are aggregated over all cells of the city. These plots clearly show, with the exception for the population, phenomena of saturated growth or consumption of the plotted variables. Moreover, as Λ_A grows it is possible to see a clear convergence to a stable trajectory. For (a), (b) and (c) we computed the confidence intervals at 95% probability for the parameters of a logistic growth / consumption model: while for the parameter corresponding to the saturation threshold we get < 1% error for $\Lambda_A \geq 3.4$, the best we can do to estimate the intercept of the logistic is an error of 12% for $\Lambda_A = 6.7$. This doesn't surprise us, since the city is already at a late state of growth and thus we cannot hope to give a good estimate for a parameter that is meaningful for an earlier period. The scatter plots (f)–(i) instead give substantial informations about the patterns of residential expansions occurred. They refer to $\Lambda_A = 10$, and each point is a cell. It seems clear that new houses are built in zones with a high residential density or in zones with a substantial presence of local infrastructures. Since we deal with Poissonian diffusion processes, agents explore a portion of the urban space, on average, only 3 jumps deep, starting from the most attractive cells. Thus the patterns expressed in (f)–(i) give good evidence that PCA, on which the urban and regional forces of attraction are defined, provides a good model of residential centrality.

4 Integration for environmental impact assessment

It is possible to simulate the dynamical behaviour of the MAS with an asynchronous algorithm that takes advantage of some properties of Poisson processes, namely that the time between two jumps is exponentially distributed, and that the sum of independent processes is still a Poisson process (a sketch of the algorithm is given in [10]). Since the assumption of independence holds only for a time step of length Δt , we recover an evolution schema familiar to that of a CA. This means that the model could be integrated very easily in an environmental decision support system (EDSS), even as a simple routine call. At each time step, either the configuration of the system may be updated by running the PCA, and thus producing new values of the force of attraction F and G , or the value of the dynamical state $v(c, t)$ of each cell c can be retrieved as an output, (or both). Moreover, interaction issues with other subsystems – such as the biophysical component of an EDSS – arising from a different time scale of simulation, can be reduced by decoupling those two operations. Integration at the data level can be done in a straightforward manner if the model has to produce only output values to be fed into other components. In this case it is worth to note that we explicitly model the dynamics of the population of the system (see figure 3(d)), as well as the number of workers in local facilities of each cell (not shown in the figures). This just to name a few examples. These can become the inputs to compute measurable indicators, and thus evaluate the significance of the growth of the city on the environment. A more pondered approach - instead - should be taken if data integration contemplates the possibility to feed variables describing the status of the environment into our model. A very simple operation would be to change (9) to allow a sum on more variables than those in V_α , and thus it would be only needed to apply PCA to an augmented data matrix $X(t)$ with a new column for each environmental indicator one would like to consider. As an example, for the event of buying a house, agents could then take into account

also informations about air and noise pollution in a cell.

5 CONCLUSIONS

The results we have discussed in this paper stem from a calibration that makes several highly idealized assumptions, namely the choice of the values for Λ_D^α , Λ_L^α , Λ_U^α , which result in homogeneous decision process of the agents with respect to the kinds of interaction α , and the dynamics of the global rate of activation Λ_A , which lead to a stationary dynamics of the system, which is motivated by the fact that the city in our case study is already in a late stage of development. Nonetheless, our model was able to show a meaningful urban dynamics of growth, with key factors such as saturation due to depletion of resources (see figures 3(a), (b), (c) and (e)), and realistic decision processes of urban agents. We thus believe that this model has the potential to produce quantitatively precise scenarios of growth – given enough data to perform a proper calibration. A shortcoming of our approach is the fact that PCA does not define a probability density model, and thus doesn't allow one to compute the likelihood of data, which can be problematic for the calibration of the forces of attraction F^α and G^α . We will have to reflect about a proper calibration technique for this part of the model. Some benefits of our approach are: (i) the dynamics of our MAS are mathematically well defined so that one could study effects due to bifurcations or phase transitions analytically, and not only through simulation, (ii) the asynchronous algorithm we use can simulate many thousands of agents at once, (iii) this approach doesn't have to create coherent identities for various social classes of interest, (iv) update rules of the CA have a clear urban meaning and are built up from simple events (e.g. building a house), (v) we generalize a CA to have a real valued multidimensional state space, and this allows a very easy integration in an EDSS and (vi) the cells of the CA correspond to real administrative zones which would make it easier to policy makers and other stakeholders to assess the model's outputs.

ACKNOWLEDGEMENTS

We would like to thank Carlo Bagnasco and Ilia Arcidiacono for having provided us the data from their study on Rome's eastern area (see [1], in Italian) and for many insightful discussions. Giovanni Luca Ciampaglia is supported by the Swiss National Science Foundation under project no. 20021-112420.

References

- [1] Ilia Arcidiacono and Carlo Bagnasco. *L'analisi statistica. Aspetti teorici e applicazioni alla lettura del territorio*. Aracne, 2006.
- [2] Michael Batty and Bin Jiang. Multi-agent simulation: New approaches to exploring space-time dynamics within GIS. *CASA working paper*, 1999.
- [3] Itzhak Benenson. Multi-agent simulations of residential dynamics in the city. *Computers, Environment and Urban Systems*, 22(1):25–42, 1998.

- [4] Guy Engelen, Roger White, and Ton Nijs. Environment explorer: Spatial support system for the integrated assessment of socio-economic and environmental policies in the netherlands. *Integrated Assessment*, 4(2):97–105, June 2003.
- [5] Brian S. Everitt and Graham Dunn. *Applied Multivariate Data Analysis*. Hodder Arnold, February 2001.
- [6] J. T. Kent, J. M. Bibby, and K. V. Mardia. *Multivariate Analysis (Probability and Mathematical Statistics)*. Elsevier Limited, 2006.
- [7] John F. C. Kingman. *Poisson Processes (Oxford Studies in Probability)*. Oxford University Press, USA, 1993.
- [8] Sam T. Roweis and Zoubin Ghahramani. A unifying review of linear gaussian models. *Neural Computation*, 11(2):305–345, 1999.
- [9] Alberto Vancheri, Paolo Giordano, Denise Andrey, and Sergio Albeverio. Continuous valued cellular automata and decision process of agents for urban dynamics. In *Computers in Urban Planning and Urban Management (CUPUM 2005)*, 2005.
- [10] Alberto Vancheri, Paolo Giordano, Denise Andrey, and Sergio Albeverio. A model for urban growth processes with continuous state cellular automata, multi-agents and related differential equations. Part 1: Theory. *Environment and Planning B: Planning and Design*, 2008. (Submitted).
- [11] Roger White and Guy Engelen. High-resolution integrated modelling of the spatial dynamics of urban and regional systems. *Computers, Environment and Urban Systems*, 24:383–400, 2000.

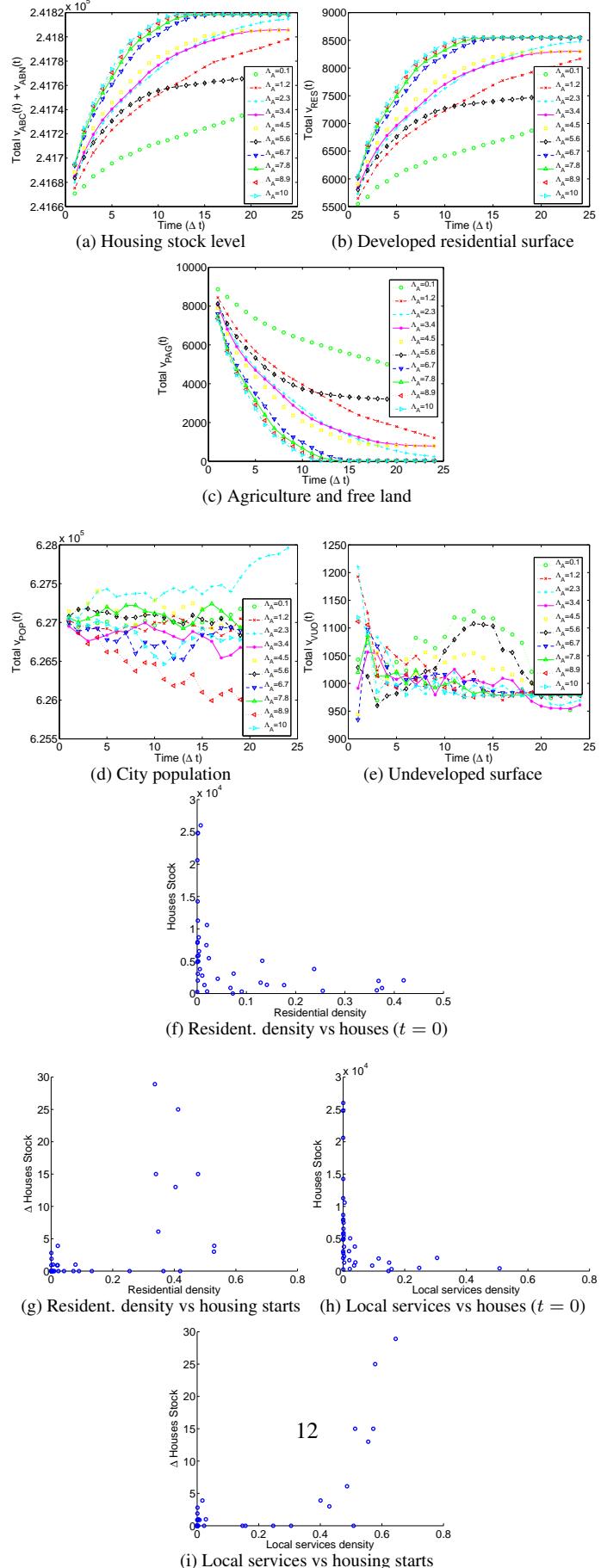


Figure 3: (a)–(e): comparisons of time series of global variables; different curves in each plot correspond to different values of the parameter Λ_A , the global intensity of ac-