Huitièmes Rencontres de Théo Quant

Besançon, 10-12 janvier 2007

http://thema.univ-fcomte.fr/

Calibration of households preference for open-spaces from an urban cellular automata model

Method and application to Dijon

Geoffrey Caruso*, Mohamed Hilal**

** Geography and Spatial Planning Research Centre, UR IPSE, University of Luxembourg*
Centre for Operations Research and Econometrics, Université catholique de Louvain, Belgium
geoffrey.caruso@uni.lu
*** INRA UMR 1041 CESAER Dijon, France*
hilal@enesad.inra.fr

KEY WORDS

Cellular automata
Calibration
Urban morphology
Neighbourhood externalities
Suburbanisation

ABSTRACT

We propose a calibration method for a residential growth model that is grounded on a micro-economic cellular automaton. This model can be seen as a spatial and dynamic representation of an urban economic model with neighbourhood externalities. A 2D spatial equilibrium of residential locations is obtained stepwise through time instead of being a one-shot instantaneous equilibrium. Previous research work showed that the model can produce more or less dense and fragmented urban patterns depending on the preference of households for open space. We propose here a method for calibrating those preferences from simulations of the model and land rents observed within the Dijon urban area. More precisely, equilibrium properties of the model are used to derive the elasticity of open-space preferences. Then simulations are used to estimate the neighbourhood distance to which open-space are valued by households. Our first results tend to support our residential behaviour assumptions and tend to be in accordance with results obtained with more classic methods. Further methodological improvements are however needed.

MOTS-CLÉS

Automate cellulaire
Calibrage
Morphologie urbaine
Externalités de voisinage
Périurbanisation

RÉSUMÉ

Calibrage des préférences des ménages pour les espaces ouverts en utilisant un automate cellulaire urbain. Méthode et application à Dijon.

Nous présentons le calibrage d'un modèle de croissance résidentielle fondé sur un automate cellulaire microéconomique. Ce modèle peut être vu comme l'expression spatiale et dynamique d'un modèle économique urbain avec externalités de voisinage. Un équilibre 2D des localisations résidentielles y est atteint graduellement, contrairement aux projections instantanées dans le temps des modèles économiques standards. Des travaux antérieurs ont montré que ce modèle produit des configurations spatiales plus ou moins denses et fragmentées selon les préférences des ménages pour les espaces ouverts. Nous proposons ici une méthode pour calibrer ces préférences à partir de simulations du modèle et de rentes foncières observées au sein de l'aire urbaine dijonnaise. Plus particulièrement, certaines propriétés d'équilibre du modèle nous permettent de définir l'élasticité liée à ces préférences. Les simulations permettent ensuite d'estimer l'extension spatiale du voisinage considéré par les ménages pour évaluer ces externalités vertes. Les résultats préliminaires obtenus semblent corroborer nos hypothèses de comportement résidentiel et vont dans le sens des résultats obtenus par des méthodes plus classiques. Néanmoins les développements méthodologiques sont à poursuivre.

Caruso G., Hilal M., 2008, Calibration of households preferences for open-spaces from an urban cellular automata model. Method and application to Dijon, in : Foltête J.-C. (dir.), *Actes des Huitièmes Rencontres de Théo Quant*, Besançon.
ISSN 1769-6895. Article mis en ligne le 12 juin 2008.

1. Introduction

In this paper, we present a calibration methodology for a cellular automata type model of urban growth that is founded on microeconomics. Residents are assumed to dislike denser neighbourhoods and trade-off this density with housing consumption and transport costs. Local density, and thus the open-space amenity, is endogenous and changes through time as a result of further development.

Theoretical experiments with a version of the model where households value local density both positively and negatively, can produce a variety of mixed urban/agricultural spatial patterns at the periphery of a city in the short and the long-run equilibrium (Caruso *et al.*, 2007; Caruso, 2005). The form of peripheral urban developments is more or less fragmented depending on the residential preference for local open-space amenities (low density) and the maximum neighbourhood distance at which these amenities are perceived.

An application test of this model to the southern commuting area of Brussels (Belgium) has shown that the model can generate spatial patterns that are reasonably similar to observed patterns when the model is constrained by actual transport networks and spatial planning restrictions (Caruso *et al.*, 2005). A comparison of observed and modelled fractal measures and fragmentation indices was used to derive preference parameters.

However, no test was undertaken to analyse whether the land rent profile generated by the application of the model is consistent with observed values. A calibration solely based on forms cannot ensure that the effects of the local externalities with respect to the transport cost are not over- or under-estimated.

Moreover, in the previous experiment, the model was unable to represent the growth of existing local settlements beyond simple contiguity effects. By assumption, therefore, it was unable to represent the attractiveness of local towns and sub-centres as providers of public goods and services.

In the present experiment, we try to respond to these two insufficiencies by (i) taking into account the accessibility to local centres (where social amenities are provided constantly, at given price and represent a fixed consumption) and (ii) by calibrating the model against observed land consumption rather than on the actual form of periurban developments. The local agglomeration effect (the liking for local density) being transformed into a transport cost, the number of unobservable parameters is reduced. Two parameters are to be found as part of the calibration exercise: the preference parameter, β, for local open-space and the radius of the neighbourhood where these open-space amenities are gleaned, x.

The calibration is based on geo-referenced transactions for the Dijon urban region. A highly schematic representation of existing spatial settlements is used. Although the spatial outputs are thus less realistic, our interest is not in replicating the global form of residential developments. Moreover, we now from the previous experiments that the morphology within the mixed suburban area is not influenced by the contours of existing settlements (except within a fringe of the size of neighbourhoods) but by the spatial structure of the transport costs and zoning constraints. It is the 'within morphology' that affects land values in the model.

In the paper, we present successively the model (section 2), the dataset (section 3), the calibration method and the first results obtained (section 4) so far.

2. The model

In this section, we first introduce the main assumptions and the functioning of the model. It is a particular case of the model proposed in Caruso *et al.* (2007) where a more detailed description of the assumptions can be found. We then formulate the residential choice and present the long-run equilibrium properties of the model that will be used to calibrate the model.

2.1. Model assumptions and dynamics

We consider a city region where space is divided up into cells. A cell is either occupied by agricultural use or residential use. Residents choose location according to the transport cost that characterizes a cell and according to the level of 'green space', *i.e.* the agricultural cells around.

The model can be seen as a discrete 2D application of the standard urban model with crowding externalities (Fujita, 1989). However, the treatment of the crowding externality is quite different because space is discrete and 2D. In the standard model, residents are allocated continuously along a distance line. Density at distance d is then formed by the housing consumption in d or if a neighbourhood of size x is considered, by the sum of housing consumption in the range $[d–x,d+x]$. In the case of our model, density for each location is taken from a 2D neighbourhood of radius x around the location (and radial symmetry is no longer assumed). The relationship is thus direct with applied models of environmental economics where locational attributes are measured using GIS neighbourhood functions (*e.g.* Geoghegan *et al.*, 1997).

The model is also dynamic: the density considered by a resident is lagged in time. Therefore, like in cellular automata models, the evolution of the spatial structure is driven by neighbourhood attributes. Compared to most cellular automata models of urban sprawl, however, the advantage of this model is to ground transition rules within a well established theoretical framework.

The dynamics is further governed by a migration process. The number of residents within the area increases at a given constant rate g. The migration is thus sequential. In theory, the time interval t is to be taken sufficiently small so that g=1. Such an asynchrony assumption ensures that each migrant maximizes utility when choosing location within the metropolitan area. However, in practice, depending on the cell size of real case applications, g>1 can be chosen to avoid computing burdens.

We further impose that no residential site can return to agriculture because of very high re-conversion costs. The model thus yields path-dependency, which is seen as an important characteristic for representing the spatial impact of migrations through time (in another context, dependence on history is seen as a part of the explanation of the location of firms agglomeration processes (Arthur, 1994)). Also, Yacovissi and Kern (1995) showed that dynamic urban models (see the review in Brueckner, 2000) are better at representing the flattening of density through time than static models. A renewed interest is seen in the urban literature for dynamic models and neighbourhood processes, with application to urban sprawl (*e.g.* Turner, 2005) or spatial segregation (*e.g.* Durlauf, 2004; Galster *et al.*, 2000).

We consider an open-city with a competitive rental land market where landowners allocate land to the highest bidder. New migrants thus compete with farmers on agricultural parcels. In order to settle in a previously agricultural cell migrants have to bid over the farmers. At each time step, the arrival of a migrant changes the level of open-space around the residential site chosen. The neighbours affected may therefore want to move (at no cost) if landowners do not adapt the level of the rent. By adapting the rent at each step, utility is kept constant among all households within the area. Each time step t is thus a short-run equilibrium where all households have utility U^t.

An important characteristic of the model is that landowners are put in competition by the arriving households, which creates a disequilibrium on the land market. Because of excess supply, migrants pay only the agricultural rent when they settle at time t. They therefore pocket a utility surplus which also accrues to all other residents already installed. As long as the surplus $U^t - \overline{U}$ is positive, the region is attractive and migration continues to occur. Once the utility surplus is null, a long-run equilibrium is reached where households have the same utility than elsewhere, $\overline{U}$.

2.2. Residential behaviour

All individuals are identical in terms of income and preferences. They obtain a wage (Y) from their work in the CBD and pay a transport cost T for their commuting and trips to local centres. Utility is Cobb-Douglas and maximized under the budget constraint:[1]

$$max \qquad U(Z,L,E) = kZ^{1-\alpha}L^{\alpha}E^{\beta} \tag{1}$$

$$s.t \qquad Z + RL = Y - T \tag{2}$$

with $\alpha \in [0,1], \beta \geq 0$, and $k = \alpha^{-\alpha}(1-\alpha)^{\alpha-1}$ (which simplifies the writing of the bid rent). E is the open-space externality. Z is the consumption of a composite good, acquired both in the CBD and in the closer local centre. L is the land consumption. The price of Z is unitary and R is the land rent.

The greenness externality is a convex decreasing function of density ρ. We thus suppose a decreasing marginal effect of increasing density. For a cell i, the externality is

$$E_i = e^{-\rho_i} \tag{3}$$

with ρ_i, the density of households in the neighbourhood of i:

$$\rho_i = \frac{\sum_{i' \in N_i} I_{i'}}{n} \tag{4}$$

$I_{i'}$ is a binary variable =1 if i' is in residential status and =0 otherwise. n is the size of the neighbourhood in number of cells. N_i is the set of cells i' belonging to the neighbourhood of i. i' is such that its distance to i, $x_{ii'}$, is $\leq x$, the radius of the neighbourhood.

At the optimum, demands for the two goods Z and L in i are

$$\hat{Z}_i = (1-\alpha)(Y - T_i) \tag{5}$$

$$\hat{L}_i = \alpha(Y - T_i)R_i^{-1} \tag{6}$$

and the indirect utility is given by

$$V_i = (Y - T_i)R_i^{-\alpha}E_i^{\beta} \tag{7}$$

Households that migrate into the city are ready to pay Ψ_i to locate in i and obtain a level U of utility. Ψ_i, the bid rent of the household, is determined by

$$\psi_i = (Y - T_i)^{1/\alpha}U^{-1/\alpha}E_i^{\beta/\alpha} \tag{8}$$

2.3. Market equilibrium

At time t, as long as a new migrant can find a site j where, paying the agricultural rent Φ, he gets a utility

$$U^t = (Y - T_j)\phi^{-a}E_j^{\beta} \geq \overline{U} \tag{9}$$

rent in the last site developed is given by

[1] Conversely to Caruso *et al.* (2007), land consumption is not fixed here and will be used to calibrate the model.

$$R^t_j = \phi \tag{10}$$

and by

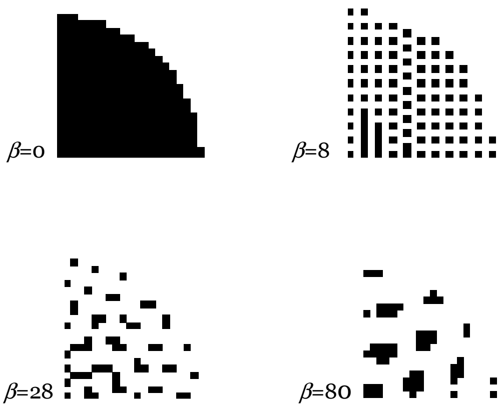
$$R^t_{j'} = (Y - T_{j'})^{1/\alpha}(U^t)^{-1/\alpha}E_{j'}^{\beta/\alpha} \tag{11}$$

in all other site *j*' already occupied.

Given that utility, is determined by another household, $R^t_{j'} < \phi$ is possible, even in the long-run when $U^t = \overline{U}$.

2.4. Long-run equilibrium properties

Figure 1 provides an example of spatial configurations that can be obtained in the long-run equilibrium using this model. A mixed area is created in the periphery of the city. The spatial configuration depends on the preference for the externality (β) and the spatial extension of the neighbourhood (*x*). These are the parameters we want to calibrate in the remainder of the paper.



CBD at bottom left corner. Standard equilibrium (β=0) and then equilibrium with increasing neighbourhood size with β>0.

Figure 1. Example of a monocentric output of the model for 3 neighbourhood sizes

One can also see from the figure that the standard urban fringe obtained when β=0 is the maximal extent of the mixed area obtained with β>0, whatever *x*. In fact, we know that at the maximum extent of the monocentric city at the long-run equilibrium, the reservation bid-rent of a household equals the agricultural rent. If the standard urban fringe is denoted by *f*, the maximum cost of transport, T_f is then obtained by equalling (12) to Φ:

$$T_f = Y - \phi^\alpha \overline{U} E^{-\beta} \tag{12}$$

If β>0 and ρ>0, one can see from (12) and (3) that T_f is inferior to T_f when β=0. Therefore, the city with the green externality cannot be further expanded than the standard city. Also, if a household wants to locate in a cell where the transport cost is T_f, the neighbourhood density must be null (ρ=0). We denote by T_{fc} the commuting fringe of the city, *i.e.* the maximal transport cost possible in the mixed area.

$$T_{fc} = Y - \phi^\alpha \overline{U} E^{-\beta}_{\rho=0} = Y - \phi^\alpha \overline{U} \tag{13}$$

Similarly, we know that the worst local condition correspond to a fully urbanised neighbourhood, ρ=1. The lower the transport cost, the better a household can compensate for the disamenity. There is a transport cost limit, below which, even with ρ=1 an household will always overbid farmers. We denote by T_{fu} the limit of the compact city:

$$T_{fu} = Y - \phi^\alpha \overline{U} E^{-\beta}_{\rho=1} \tag{14}$$

Finally therefore, we know that at the long-run equilibrium, a mixed area will exist between T_{fu} and T_{fc}, an agricultural plain beyond T_{fc} and a compact city where *T* is lower than T_{fu}.

3. Study area and dataset

3.1. Study area

The study area is situated in Burgundy (France) (Figure 2). It corresponds to the urban area of Dijon and made of an urban pole (Dijon and 14 suburban municipalities) and 199 periurban municipalities (see definitions in appendix). The study area represents a surface of 2 300 sq.km and 330 000 inhabitants (237 000 within the agglomeration of Dijon and 90 000 in the periurban municipalities. There are 138 000 employments of which 59% in the Dijon municipality and 28% in the suburbs (Hilal, 2005).

Between 1990 and 1999, 15 729 persons have moved from the urban pole to the periurban area, while 8 802 have made the reverse move. 82% of the workers who live in the periurban belt work outside of their own municipality. Given the intensity of the residential mobility and the intensity of commuting, we can make the assumption of self-contained labour and residential markets across the study area.

Over the 1990's, the whole area gained 13 727 persons, *i.e.* a 0.5% annual growth rate. This growth was mainly due to natural growth (+15 683 persons). The migration balance with external zones was negative (-1 956 persons). Within the study area, population dispersion is observed, with higher growth rate further away from the central city: Dijon grew by 0.2% per year, the suburbs by 0.4% and the periurban belt by 0.9%. Although the population in Dijon and its suburbs increased, this was due to natural growth only. The migration balance was negative despite the growth of youth. Conversely, the periurban ring benefitted from an important population growth due to positive natural and migration balances.

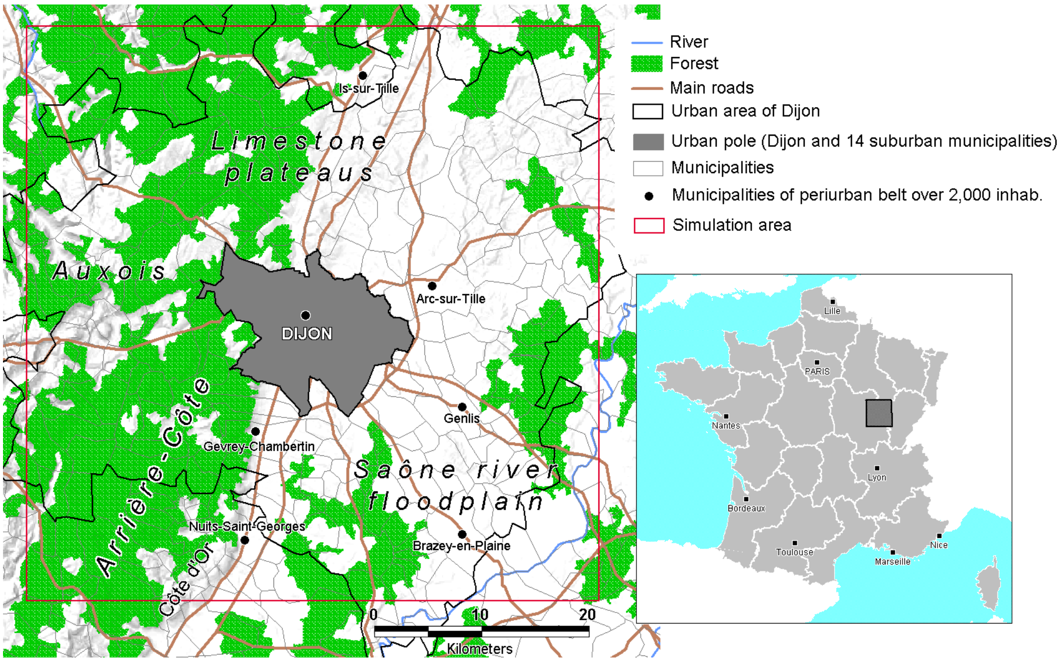


Figure 2. Study area location and geography

In 1999, there were 151 565 dwellings within the study area. This is 12.6% more than in 1990. Over the decades, population grows only by 4.4% however. The evolution is thus explained by decreasing family size, increasing single parenthood, and increasing number of one person households. Main homes correspond to 90% of dwellings and their increase was much greater in the periurban belt (+16 %). Individual houses count for 42% of main homes on average, but the proportion increases with distance from the centre: 18% in Dijon, 47% in the suburbs, and 91% in periurban municipalities. Urban spread re-inforces the volume of individual houses: they represent 8% of newly built dwellings in Dijon since 1990, 45% in the suburbs, and 90% in the periurban area. The proportion of owner-occupiers also increases with distance: 41% in Dijon, 59% in the suburbs and 80% in the periurban belt.

The study region covers four main geographic units differentiated by topography and land use. North of Dijon are limestone plateaus with large cereal farms employing little labour. South of Dijon lies a series of three strips: to the west is the Auxois livestock farming region with its landscape of hedge-lined meadows in the valleys and woods on the higher land; then comes the Arrière-Côte, a limestone plateau dissected by dry valleys with diversified farming (fruit, cereals, livestock); to the east is the Saône River floodplain with its forests and intensively farmed arable land (market gardening and arable crops). A sharp scarp separates the last two strips, along which run the vineyards for which Burgundy is of great repute.

Note also that the region is an area of clustered housing where views are rapidly masked by neighbouring houses. And finally, according to the Corine Land Cover classification, built-up areas cover 2.4% of the land, farmland 59%, and woodland and semi-natural land 38%.

3.2 Dataset

The model will allocate residents stepwise within the periurban belt part of the study area (about 5 to 30 km from the centre of Dijon). The simulation area is a 50 km × 50 km grid (see Figure 3, centred on Dijon. Spatial resolution is 50m, *i.e.* 10^6 cells. This is a very fine resolution compared to most cellular automata model. However, this is necessary as the recent literature on the value of open-space tends to indicate that the very close proximity of green spaces within mixed peripheries is more important than large open-spaces in the outskirts (*e.g.* Nechyba and Walsh, 2004, or work by Cavailhes and colleagues in Dijon and Besancon). For computation purpose, the study area is thus divided into 16 quadrants, each one containing 62500 cells. The different quadrants will be filled in with residents separately.

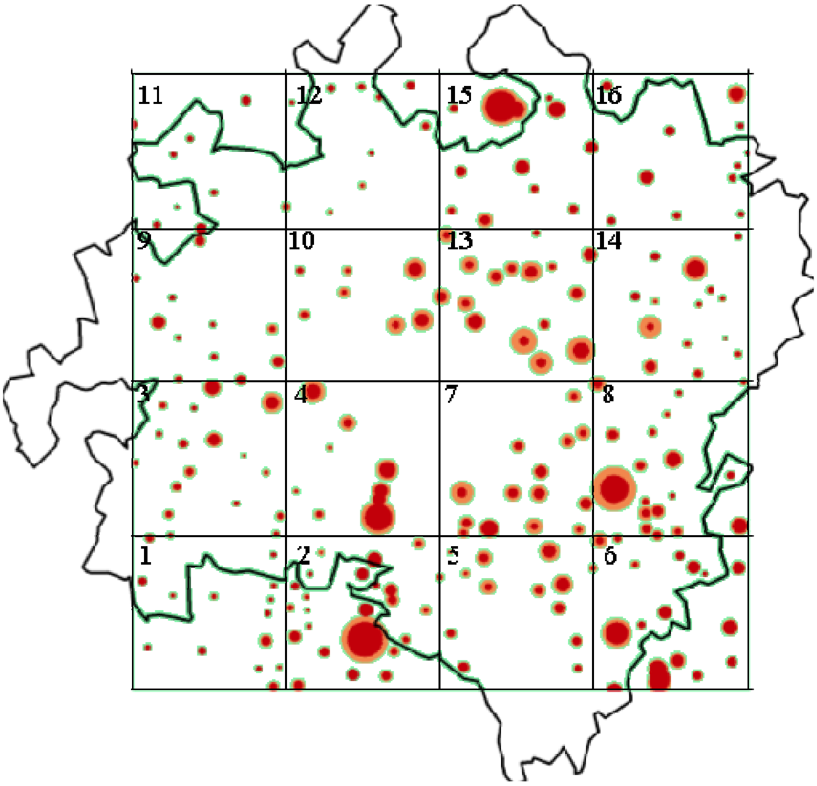


Figure 3. Study area quadrants: initial settlements 1968 (red) and development constraint for 1999 (orange)

Cells may be converted into urban use at time t based on transport costs characteristics (and neighbourhood density) calculated from t–1. Three GIS layers, following the same grid, are thus needed as input to the model:

– Layer 1 is the initial state of urbanisation (t_0). Initial date is fixed at 1968, which is the last census before the development of periurbanisation in France. A schematic representation of the built-up area at that period is created. For each municipality, the initial state of urbanisation is represented by a circular node. The geographic coordinates of the town hall is used as the centre of the node (with the church, it is always located in the centre of old villages). The surface of the nodes is made proportional to the number of dwellings given by the 1968 census (and coloured red on the map). Moreover the whole surface covered by the Dijon agglomeration is also considered as urbanised.

– Layer 2 represents the cells available for building. For each municipality, the surface of this layer is proportional to the number of dwellings in 1999. By subtracting the initial settlement surface (layer 1), we obtain a buffer of cells available for building. This second layer operates as a zoning constraint.

– Layer 3 is the transport cost. As seen previously, the area is broadly monocentric: 87% of the jobs are situated in the Dijon agglomeration and the historic centre is the part of the city that concentrates the more jobs (Hilal, 2005). The distance to Dijon is thus an important part of the transport cost. However, in this experiment we also aim at taking into account the presence of local centres as an agglomeration driver within periurban areas. So, the following generalised transport cost function (expressed in €) is used in the simulations.

$$GENC=(CLCKM+DIJKM)*0.30+(CLCMNT+DIJMNT)*0.16 \quad (15)$$

where *CLCKM* is the Euclidean distance between a cell and the centre (town hall) of the municipality, *DIJKM* is the road network distance (km) between the local town centre and Dijon, *CLCMNT* is the time distance between the cell and the local town centre (applying an assumed speed of 25 km/h on Euclidean distance), and *DIJMNT* is the road network time between the local centre and Dijon. The km costs are weighted by 0.30 €/km, which is the average price used in France by the revenue administration to calculate income taxes. The time costs are weighted by 0.16 €/minute, which is a time cost estimation by the French Ministry of Transport.

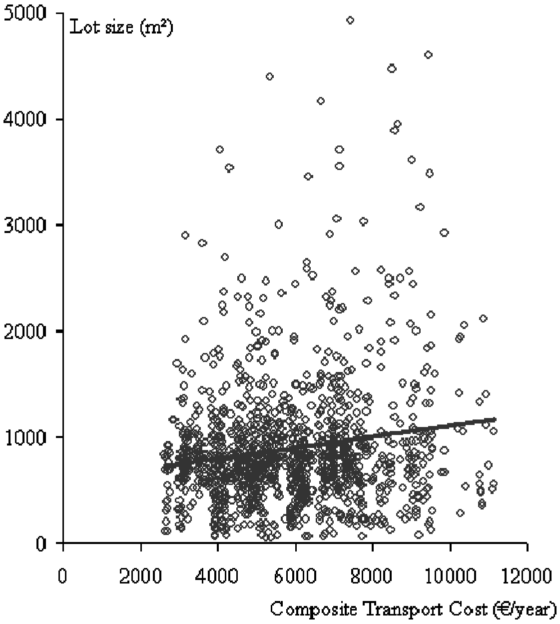
Since the model uses annual income of a household, the trip cost *GENC* is transformed into an annual transport cost, *T*:

$$T=GENC*2*1.5*230 \quad (16)$$

(Assumption of a return trip every of the 230 working days, by 1.5 persons by household)

Finally, lot size is used to calibrate the model. The dataset comes from the real-estate lawyers in charge of the conveyance of land property transactions in France ('notaires'). The database is made up of 1700 sales of building land between 1995 and 2002. The database records the location, the size of the lot, and the price of transactions that were made between private individuals.

Lot size within the periurban belt area (as defined by INSEE, see Annex A) is plotted against transport cost T on Figure 4).



1454 observations. r^2=0.0175

$\hat{L} = L_{T_{fu}} + a(T - T_{fu})$

$a = 0.0520$ (st err 0.0102)***

$L_{T_{fu}} = 726.6721$ (st err 37.6711)***

Figure 4. Observed lot size and trend

4. The calibration

4.1. Limits of the mixed belt and long-run utility

First, we assume (i) that the area is in long-run equilibrium and (ii) that the long-run mixed belt corresponds to the periurban belt of our study area.

Although the determination coefficient of the trend curve within the observed lot sizes is low, we use the above trend to derive lot size at the two limits of the periurban area. T_{fu}=2624 and T_{fc}=11180 being given by the lowest and highest T of the sample, we find $L_{T_{fu}} = 727$ sq.m and $L_{T_{fc}} = 727$ sq.m from the $\hat{L}$ fit (Figure 4).

We then determine the reservation utility $\overline{U}$. We use Y=32000 €/year as the average income and α=0.2 to represent the share of dwelling consumption in the total income net of transport costs. Utility is fixed from land consumption in T_{fc}, where it can be extracted independently of β as it is assumed that ρ=0 and thus (E=1) at the external limit of the periurban area. The bid-rent is the bid of the standard urban model with no externalities. Equalizing the demand function of L (equation 6) to $L_{T_{fc}}$=1171 and using T_{fc}=11180 we extract the capitalized value of the rent at T_{fc}. We therefore find the annual rent (with 0.05 interest rate) which is used in equation 8 with ρ=0 to extract $\overline{U} = 29683$.

By doing so, we force the $L_{|\rho=0}$ curve of the model to meet the observed trend $\hat{L}$ at T_{fc} as shown on the graph on Figure 5.

4.2. Determining the open-space preference

The calibration of β is also obtained from the long-run properties of the model without running simulations.

In fact, as previously seen (equation 13), the external limit of the mixed periurban area at the long-run equilibrium, T_{fc}, is supposed to be the intersection between the residential bid rent with null local density and the agricultural rent Φ. In practice Φ must be increased by the cost of servicing the land incurred by the developer or the public sector. Again therefore, we use equation 8 with ρ=0 as the limit rent, R_{fc}.

Assuming a constant agricultural rent throughout the periurban area, we can determine β following the definition of T_{fu}, the lower limit of the mixed belt, i.e. the upper limit of the compact urban area, where it is assumed that ρ=1. We thus extract β from equation 14 using R_{fc} instead of Φ. This is shown graphically on Figure 6. We find β=0.3443, the preference parameter for open-space amenities.

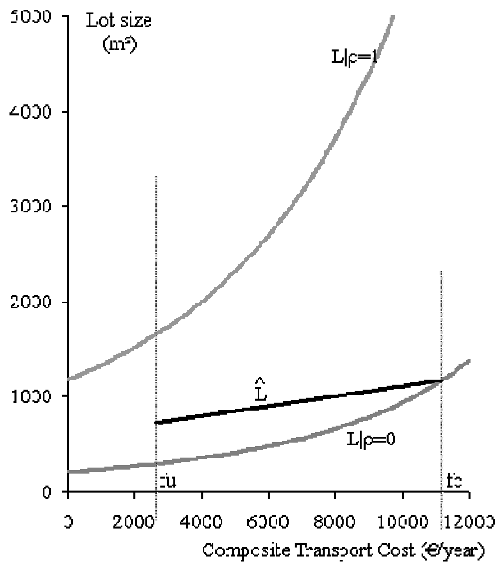


Figure 5. Model long-run equilibrium lot size curves

By doing so, we also fix the upper curve of land consumption ($L_{|\rho=1}$) as shown on Figure 5. All the housing demands generated by the simulation model will be comprised between the two curves. The trend $\hat{L}$ is, by construction, completely included within this range.

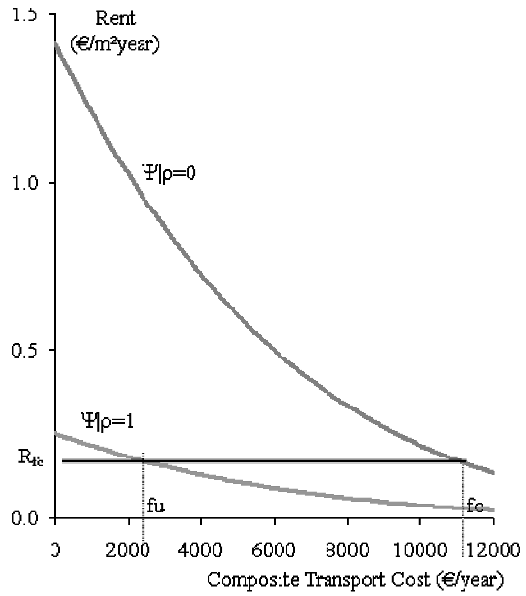


Figure 6. Model long-run equilibrium rent curves

4.3. Simulations and determination of neighbourhood size

Although β is now fixed, different spatial patterns can be obtained by varying the size of the neighbourhood where the local open-space amenity is perceived (x). As shown on Figure 1, we know that settlements are more clustered when x increases. Moreover we know that the model results in switch back and discontinuous land rents and thus land consumption profiles because the model is 2D and the density can vary from one cell to the other at a given distance (transport cost).

Therefore, the second stage of the calibration consists in running simulations for systematic change in x. Moreover, it can be useful to change the value of β around the calibration made previously, so that the model results will depend less on the assumptions made when fixing the lower and upper limits of the mixed area.

This is work in progress, and so far, simulations have only been run for the south west part of the area (quadrants 1 to 4) and two different neighbourhoods. The results of these simulations are presented below as well as the quality of the observed and predicted lot sizes.

4.4. Dynamic patterns

Given the size of the grid, simulations have been run using a growth rate g=10. 10 migrants thus enter the area at each time step. Simultaneous entries can lead to higher level of neighbourhood density than what would be expected from the migration. Arguably, this can represent a sort of a lack of information available to households. In terms of forms, everything else being equal, simultaneity lead to more compact patterns (at the limit, with $g=\infty$ we find the standard static equilibrium at time step 1).

Two neighbourhood sizes are considered. The first consists in the 8 cells surrounding a cell (the Moore Neighbourhood), i.e the neighbourhood radius is x=1.42*50m=70.7m. This is the smallest neighbourhood distance that can be tested given the spatial resolution chosen. The second neighbourhood consists in the 28 cells that fall with a x=3*50m=150m of a given cell. Others are to be tested.

Figure 7 shows the evolution of urbanisation for the first SW quadrant adjacent to the Dijon agglomeration, i.e. quadrant 4 on Figure 3. Different spatial arrangements in the periphery of the clusters are obtained from the two simulations. The 8 cell neighbourhood showing sort of linear arrangements while the 28 cells is more clustered.

4.5. Long-run equilibrium patterns

The long-run equilibrium patterns for the four quadrants are shown on Figures 8 and 9. t^* is the time at which it is not possible for a household from outside the area to settle within the area and obtain at least the level of utility he would obtain in external places. From t^* onwards, the pattern is fixed. It is the end of urbanisation (as long as no external shock changes some of the parameters). In our simulations, t^* is not obtained at the same moment in the different quadrants as it depends on the accessibility and the zoning constraints applied. (t^* values are indicated on the figures). The different quadrants are thus not to be compared on their time value. Time is different per quadrant.

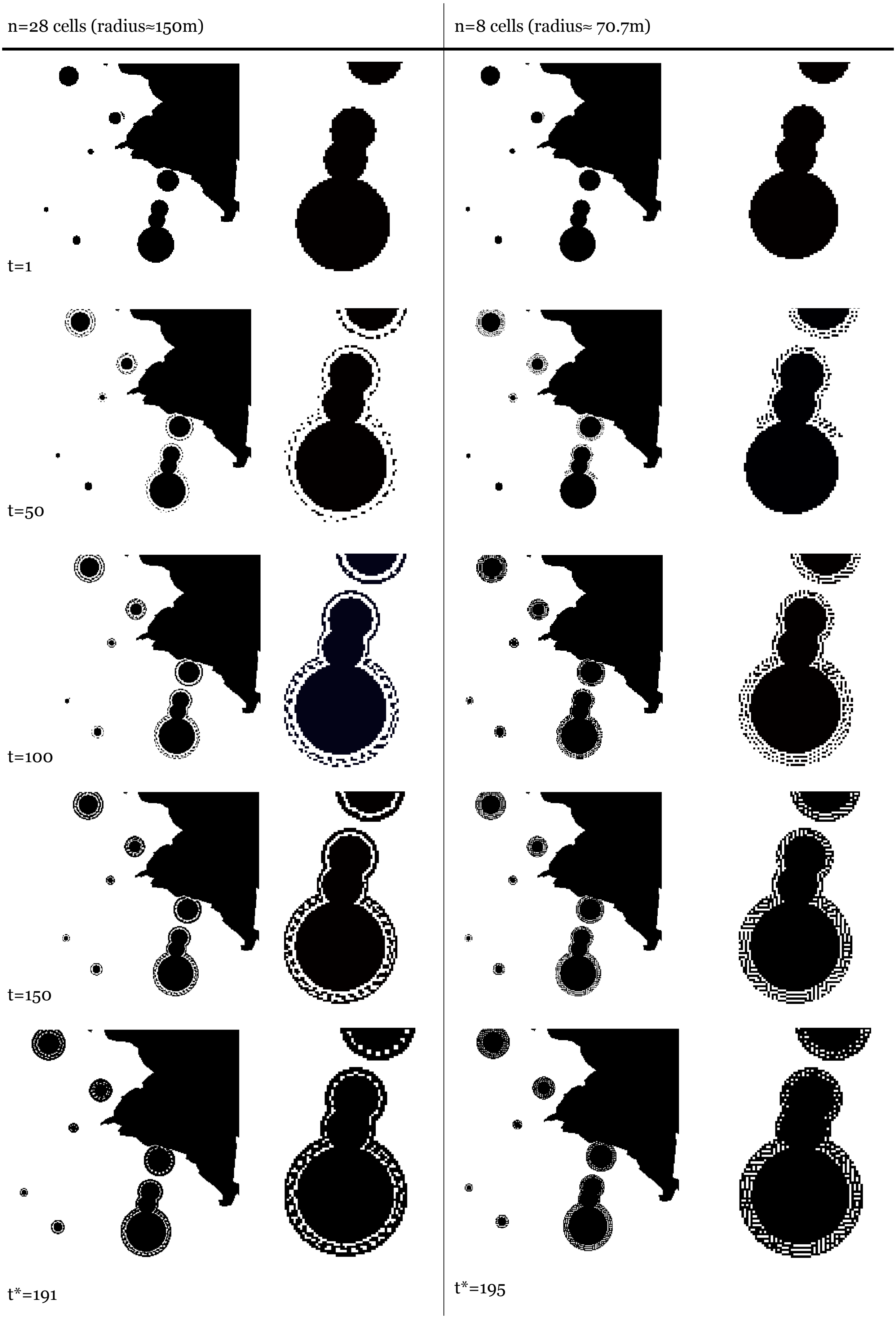


Figure 7. Consecutive short-run equilibria for quadrant 4, SW of Dijon agglomeration (+ zoom)

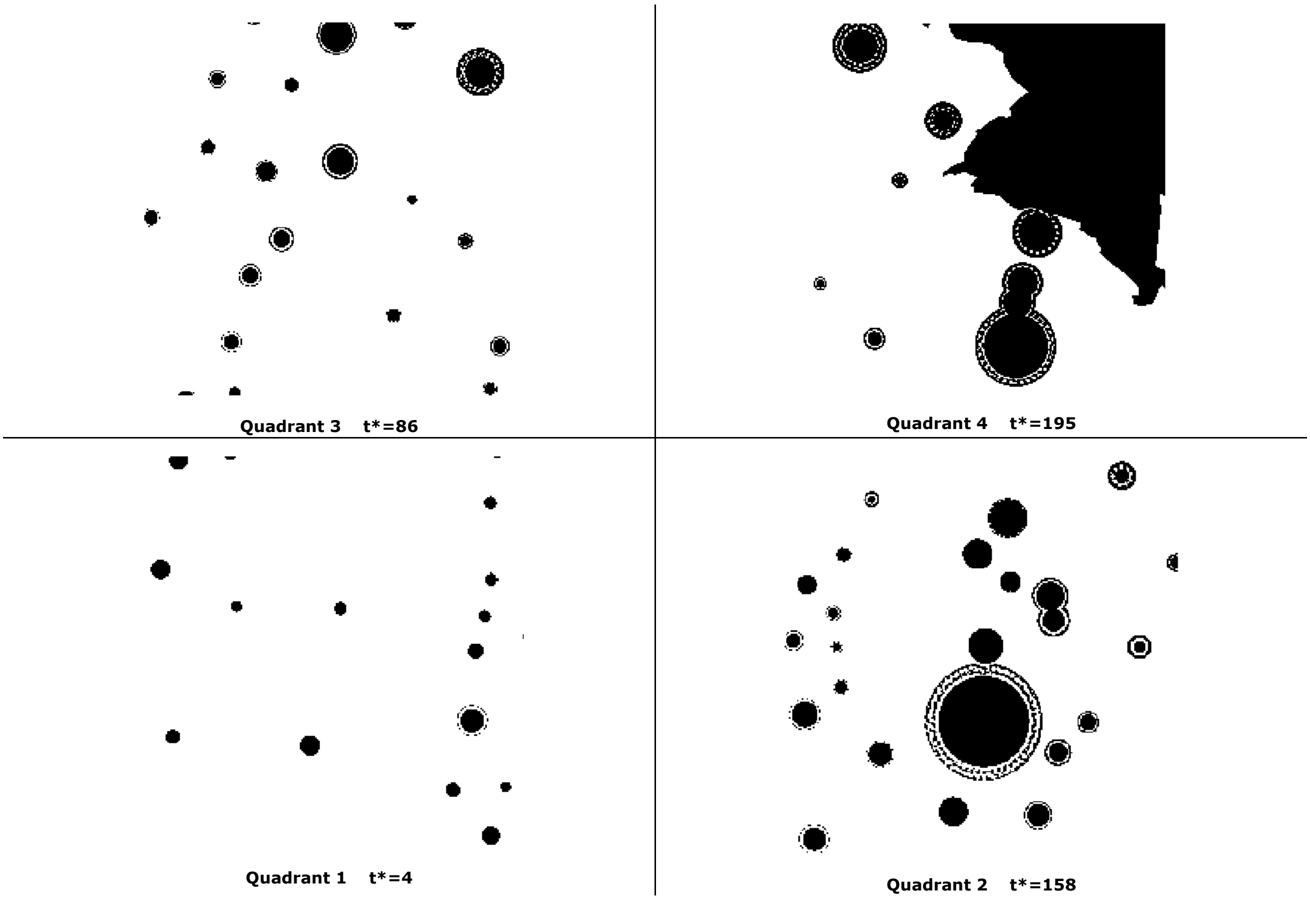


Figure 8. Long-run equilibrium pattern SW of Dijon (quadrants 1 to 4), with 28 cells neighbourhood.

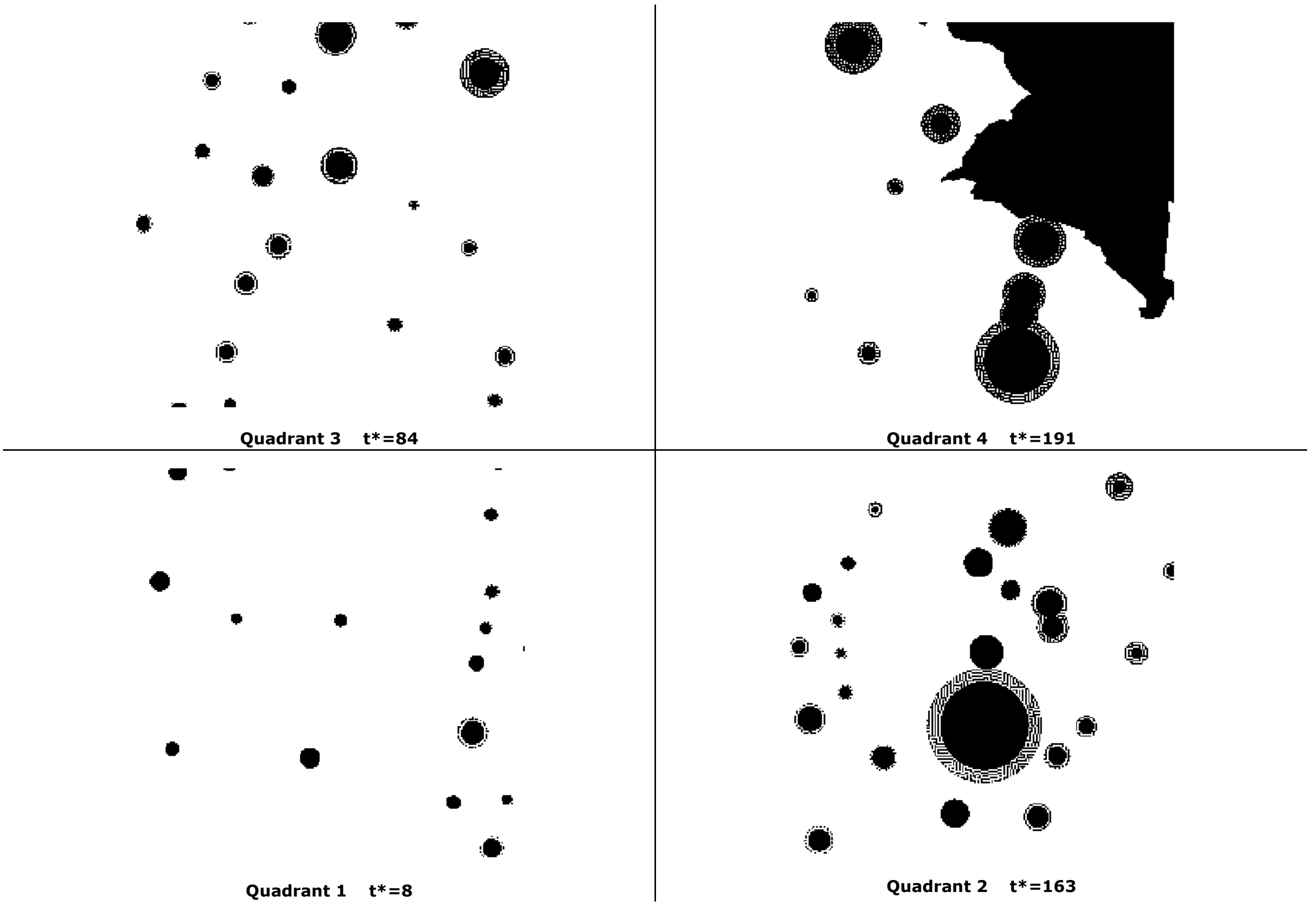


Figure 9. Long-run equilibrium pattern SW of Dijon (quadrants 1 to 4), with 8 cells neighbourhood.

4.6. Lot size adjustment at long-run equilibrium

We now need to compare the long-run equilibrium lot size obtained from the simulation model and the observed housing size for the two neighbourhood sizes, in order to decide which of n=8 or n=28 cells better accounts for the open-space amenity.

There are several difficulties for measuring the adjustment of the observed and modelled plots (Figures 10 and 11). First a point to point analysis is impossible because the model does not necessarily allocate an urban cell where one is effectively observed and sampled. Second aggregation of the data to a larger unit would erase the local variations that are typical and consists in the richness of the model. Third, comparing a linear (or not) trend within the observed and the modelled plots would also hide the local variations.

So far we have implemented a covariance analysis (combined regression and ANOVA), for which the coefficient table is found in appendix and the F-tests below (Table 1). The difference between the estimated slope coefficient is usually significant and thus the model does not seem to appropriately replicate the observed lot sizes for the two neighbourhood used. Only the results for the more remote quadrant (quadrant 1), where there is also less observations, is acceptable. The fit is slightly better for the 8 cells neighbourhood.

Although this finding is quite weak, the better performance of the 8 cells neighbourhood model (*i.e.* a 70 meter radius), is in accordance with econometric (hedonic prices) analyses of households preferences undertaken in the Dijon regions and that tend to indicate that green externalities are valued within very small neighbourhood distances (Cavailhès *et al.*, 2006). Moreover, a better performance of the model in the more remote periurban areas tend to indicate that the level of open-space externalities closer to the centre, at the limit of the urban and periurban area, is underestimated. Indeed, there is already a lot of open-spaces in the 'banlieue' area, which is not taken into account in the model and would bring extra-externalities and thus a better model fit.

	Slope difference		Intercept difference	
Quadrant	F	(Prob>F)	F	(Prob>F)
Quadrant 1,n=28	0.1236	(0.7252)	0.0473	(0.8278)
Quadrant 2, n=28	4.1423	(0.0419)	11.4145	(0.0000)
Quadrant 3, n=28	31.4936	(0.0000)	38.5582	(0.0000)
Quadrant 4, n=28	27.8192	(0.0000)	202.8075	(0.0000)
Quadrant 1, n=8	0.5222	(0.4700)	0.2425	(0.6225)
Quadrant 2, n=8	3.5848	(0.0584)	1.3734	(0.2413)
Quadrant 3, n=8	6.3594	(0.0117)	12.9886	(0.0003)
Quadrant 4, n=8	32.8642	(0.0000)	194.8084	(0.0000)

Table 1. Covariance analysis

Finally, the weakness of the results obtained so far might simply say that the model is not a good model, or that the neighbourhood sizes are too small or too big. However, from the observation of the plots, we also see that this relative failure is largely due the plot comparison method. A method needs to be developed in the future where the pattern of the two plots are compared rather than any aggregation of the plots or a point to point analysis. Point pattern matching methods will be envisaged in the next stages of the research.

5. Conclusion

We have presented a methodology for the calibration of a microeconomic cellular automata model of urban spread. The aim was to provide a way of calibrating the preference parameters for neighbourhood open-space amenities. The method assumes that the mixed area under consideration is in the long-run equilibrium and therefore, properties of the theoretical model can be used. The method can provide values for the elasticity coefficient and for the extent of the neighbourhoods considered by residents. So far, we have run a limited number of simulations from which conclusions are still uncertain. Our results so far indicate the importance of green externalities at very short distances from houses both in near and remote periurban zones. Further simulations are needed and calibration methodology improvements should be conducted in the future in order to gain more insights on the residential choice assumptions made in the model.

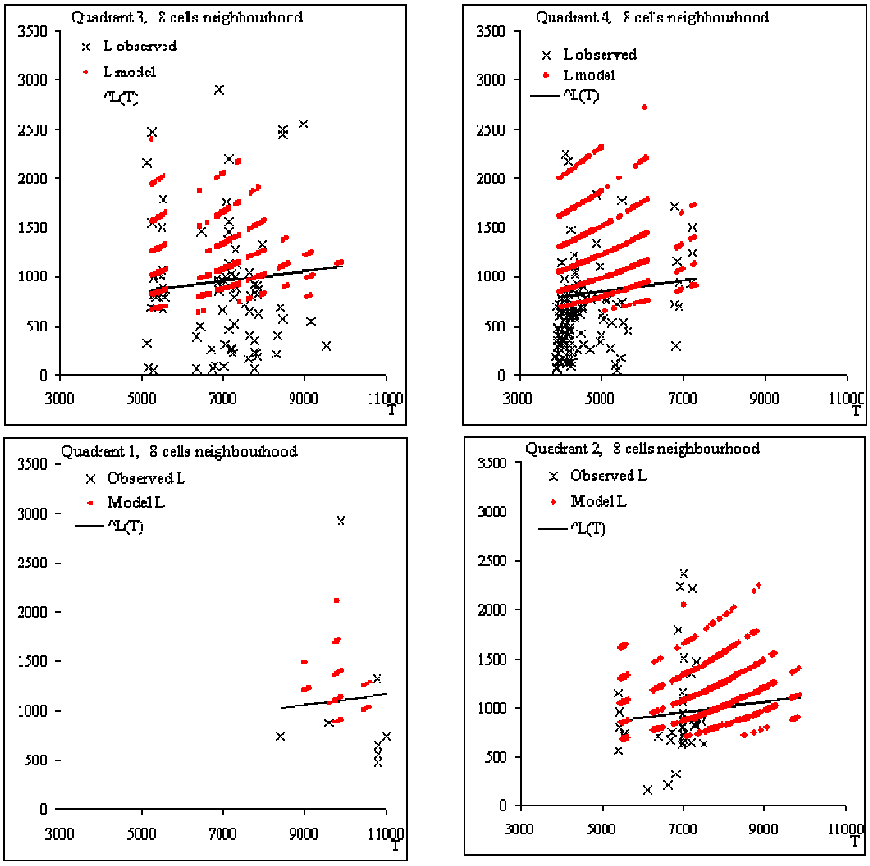


Figure 10. Long-run equilibrium lot size with 8 cells neighbourhood

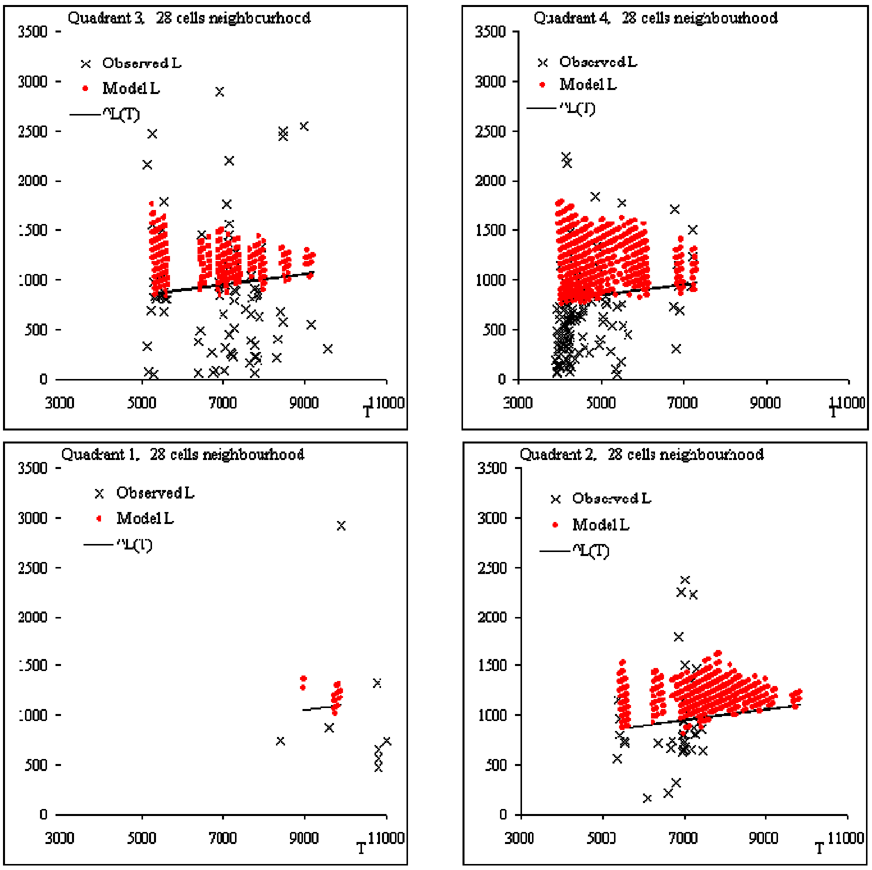


Figure 11. Long-run equilibrium lot size with 28 cells neighbourhood

6. References

Arthur W. B., 1994, Increasing Returns and Path Dependence in the Economy, University of MichiganPress, Ann Arbor.

Brueckner J. K., 2000, Urban growth models with durable housing: An overview. In: Thisse J. F., Huriot J. M. (eds.), *Economics of cities: theoretical perspectives*, Cambridge University Press, Cambridge, 263–289.

Caruso G., 2005, Integrating urban economics and cellular automata to model periurbanisation: Spatial dynamics of residential choice in the presence of neighbourhood externalities, Doctoral thesis, Department of Geography, Catholic University of Louvain.

Caruso G., Peeters D., Cavailhes J., Rounsevell M., 2007, Spatial configurations in a periurban city: a cellular automata-based microeconomic model, *Regional Science and Urban Economics*, 37, 542-567.

Caruso G., Rounsevell M., Cojocaru G., 2005, Exploring a spatio-dynamic neighbourhood model of residential dynamics in the Brussels periurban area, *International Journal of Geographical Information Science*, 19 (2), 103–123.

Cavailhès J., Brossard T., Foltête J. C., Hilal M., Joly D., Tourneux F. P., Tritz C., Wavresky P., 2006, Seeing and being seen: a gis-based hedonic price valuation of landscape, Document de travail, INRA-CESAER, CNRS-ThéMA, Dijon, Besançon.

Durlauf S. N., 2004, Neighborhood effects, in: Henderson V., Thisse J.-F. (eds.), *Handbook of Urban and Regional Economics*, vol. 4, Elsevier, Amsterdam.

Fujita M., 1989, *Urban Economic Theory: land use and city size*, Cambridge University Press, Cambridge.

Galster G. C., Quercia R. G., Cortes A., 2000, Identifying neighborhood thresholds: an empirical exploration, *Housing Policy Debate*, 11 (3), 701–732.

Geoghegan J., Wainger L., Bockstael N., 1997, Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS, *Ecological Economics*, 23, 251–264.

Hilal M., 2005, Localisation des emplois et déplacements domicile-travail dans le grand Dijon et sa périphérie, INSEE-Bourgogne, Dossier 42, Février.

Le Jeannic T., 1997, Trente ans de périurbanisation : extension et dilution des villes, *Économie et statistique*, 7, 21–41.

Nechyba T. J., Walsh R. P., 2004, Urban sprawl, *Journal of Economic Perspectives*, 18 (4), 177–200.

Schmitt B., Perrier-Cornet P., Blanc M., Hilal M., 1998, *Les campagnes et leurs villes*, INRA-INSEE Contours et Caractères, Paris.

Turner M. A., 2005, Landscape preferences and patterns of residential, *Journal of Urban Economics*, 57, 19–54.

Yacovissi W., Kern C. R., 1995, Location and history as determinants of urban residential density, *Journal of Urban Economics*, 38, 207–220.

7. Appendices

7.1. Definition of urban units and urban areas

Two classifications have been produced by the 'Institut national de la statistique et des études économiques' (INSEE), reflecting two distinct approaches that are nevertheless linked: the first, that of urban units, refers to contiguous construction; the second, that of urban areas, aims to account for territories linked to the city, not by contiguous construction but by the volume of residents working in the city.

An urban unit is an agglomeration of inhabitants, defined as a group of dwellings such that none is separated from the nearest by more than 200 metres, and accommodating at least 2,000 people. If an agglomeration of inhabitants extends over several municipalities, the group of municipalities forms an urban agglomeration. If the agglomeration only extends across one municipality, it is an isolated city.

All municipalities belonging to one urban unit are regarded as being urban. Other municipalities are classified as rural.

A 'centre' has been defined for each of the multi-municipality agglomerations. If a municipality represents more than 50% of the population of the urban unit, it is the only town centre. Otherwise, all municipalities with a population greater than half that of the largest municipality are town centres. Municipalities that are not town centres consist of the suburbs of the urban unit.

An urban area (Le Jeannic, 1997; Schmitt *et al.*, 1998) is a group of connected municipalities, with no enclosed territories, made up of: an urban centre, which is an urban unit providing at least 5,000 jobs; a periurban belt composed of rural municipalities or urban units, of which at least 40% of the resident population in employment works within the remainder of the urban area (the centre or the municipalities within its influence).

INSEE's zoning into urban areas includes other elements besides the urban areas. Polycentric municipalities are therefore municipalities or urban units of which 40% or more of the active residents work in several urban areas, without this threshold being reached for one particular area. A polycentric urban zone is a contiguous group of several urban areas and the polycentric municipalities linked with it.

The predominantly urban area comprises all urban areas and polycentric municipalities. The predominantly rural area is made up of all municipalities not belonging to the predominantly urban area. This zone includes both small urban units and rural municipalities.

7.2. ANCOVA coefficient estimates tables

The initial fit models the *y* variable, lot size, as a linear function of the *x* variable, transport cost. Each group has a separate line (a1 vs a2, b1 vs b2). The coefficients of the two lines appear in the table (e.g. Quadran 1, *n*=28). The slopes are roughly -0.0379, with a small deviation for each group:

Observed L: $y=(1584–857.329)+(–0.0379+0.899)x+\varepsilon$

Model L: $y=(1584+857.329)+(–0.0379–0.899)x+\varepsilon$

Quadran 1 : *n*=28	Estimate	Std. Err.	*T*	Prob>|*T*|
Intercept	1584	2483.1	0.6379	0.5236
a1	-857.329	2483.1	-0.3453	0.7299
a2	857.329	2483.1	0.3453	0.7299
Slope	-0.0379	0.2557	-0.1483	0.8821
b1	0.0899	0.2557	0.3516	0.7252
b2	-0.0899	0.2557	-0.3516	0.7252

Quadran 2 : *n*=28	Estimate	Std. Err.	*T*	Prob>|*T*|
Intercept	890.9940	55.3919	853	0
a1	-164.3247	55.3919	9666	0.0030
a2	164.3247	55.3919	9666	0.0030
Slope	0.0359	0.0079	5381	0
b1	0.0161	0.0079	353	0.0419
b2	-0.0161	0.0079	353	0.0419

Quadran 3 : *n*=28	Estimate	Std. Err.	T	Prob>|*T*|
Intercept	1157.7	59.8299	9.3497	0
a1	-431.0223	59.8299	7.2041	0
a2	431.0223	59.8299	7.2041	0
Slope	-0.0017	0.0096	0.1804	0.8568
b1	0.0537	0.0096	5.6119	0
b2	-0.0537	0.0096	5.6119	0

Quadran 4 : *n*=28	Estimate	Std. Err.	T	Prob>|*T*|
Intercept	1064.1	37.7259	8.2054	0
a1	-337.4045	37.7259	8.9436	0
a2	337.4045	37.7259	8.9436	0
Slope	0.0101	0.0079	1.2779	0.2014
b1	0.0419	0.0079	5.2744	0
b2	-0.0419	0.0079	5.2744	0

Quadran 1 : *n*=8	Estimate	Std. Err.	T	Prob>|*T*|
Intercept	1438	949.2896	1.5148	0.1300
a1	-711.3524	949.2896	0.7494	0.4538
a2	711.3524	949.2896	0.7494	0.4538
Slope	-0.0182	0.0971	0.1872	0.8515
b1	0.0702	0.0971	0.7226	0.4700
b2	-0.0702	0.0971	0.7226	0.4700

Quadran 2 : *n*=8	Estimate	Std. Err.	T	Prob>|*T*|
Intercept	847.3605	55.7627	1958	0
a1	-120.6912	55.7627	1644	0.0305

a2	120.6912	55.7627	1644	0.0305
Slope	0.037	0.0079	6771	0.0004
b1	0.015	0.0079	8933	0.0584
b2	-0.015	0.0079	8933	0.0584

Quadran 3 : *n*=8	Estimate	Std. Err.	T	Prob>|*T*|
Intercept	940.6214	60.9413	4349	0
a1	-213.9521	60.9413	5108	0.0549
a2	213.9521	60.9413	5108	0.0549
Slope	0.0278	0.0096	8940	0.0038
b1	0.0242	0.0096	5218	0.0117
b2	-0.0242	0.0096	5218	0.0117

Quadran 4 : *n*=8	Estimate	Std. Err.	T	Prob>|*T*|
Intercept	1103.2	40.5994	7.1735	0
a1	-376.5567	40.5994	9.2749	0
a2	376.5567	40.5994	9.2749	0
Slope	0.003	0.0085	0.3485	0.7275
b1	0.049	0.0085	5.7327	0
b2	-0.049	0.0085	5.7327	0