Topic area F1: Integrated Land Use and Transport Planning/Implementation
Topic area F2: Urban and Regional Modelling

Simulating household location choice in the Lyon Urban Area

Marko Kryvobokov, Dominique Bouf, Nicolas Ovtracht, Pierre-Yves Peguy

Laboratory of Transport Economics (LET), University of Lyon, CNRS – Lyon, France

Abstract

A transportation-land use modelling framework UrbanSim is applied. It includes a system of models, which simulate the distribution of population, jobs, real estate development and real estate prices. The study focuses on the Household Location Choice Model, which is calibrated at spatial units of different geographical levels in order to find an appropriate one. The capacity to predict the future geographical distribution of population is seen as a criterion for the choice of a spatial unit. Thus, residential location choice is predicted for past years with the same model specification at the level of blocks, zones, and municipalities. Two groups of variables are taken into account: location attributes (real estate prices, residential vacancy, and employment accessibility) and household attributes (income and car-ownership status). In comparison with the block-based model, the zone-based and commune-based models not only demonstrate better prediction capacity, but also produce output comparable with evenly split population growth.

Keywords: transportation-land use modelling, UrbanSim, household location choice, spatial units.

1. Introduction

The increasing popularity of transportation-land use modelling frameworks is explained by their capacity to analyse and simulate urban development with a reasonably high level of representation of a complex reality of an urban area with interactions between land use, transportation, population and employment. The possibility to predict future development trends make these tools the important instruments for decision-makers and researchers in transportation and urban planning.

In the study, we apply a transportation-land use modelling framework UrbanSim, which has been under development since the late 1990s at the University of Washington (Waddell, 2002; Waddell *et al*., 2003). The popularity of this open-source framework among researchers and practitioners in different countries of the world is increasing (e.g. de Palma *et al*., 2005).

UrbanSim includes a system of models, which are used to simulate the distribution of population, jobs, real estate development, real estate prices and other parameters in future years on a yearly basis. Our focus is the Household Location Choice Model (HLCM), with which we simulate the geographical distribution of households across the area. The methodology of the HLCM in UrbanSim is a multinomial logit.

A brief overview of the literature on individual household location choices using discrete choice frameworks can be found in de Palma *et al*. (2007). The early sources on this modelling effort date to the 1970s. The theory of discrete choice analysis can be found in Ben-Akiva and Lerman (1985) and Train (2003).

UrbanSim has a capability to operate at a level of fine geographical detail, either gridcell (traditionally 150 metres by 150 metres) or land parcel (UrbanSim project, 2009). In the overview of Hunt *et al*. (2005), which analyses the selected contemporary frameworks representing the stateof-the-art in transportation-land use modeling, it is noted that UrbanSim is the most disaggregate of the frameworks reviewed.

The problematic aspect of this potential is that UrbanSim needs a very detailed data for each small spatial unit. In particular, UrbanSim requires a record for each household and each job, whereas both should be located in particular gridcell or parcel[1], which in its turn is also represented as a record with multiple attributes including the number of residential units and other land use and development characteristics. It is clear that not each city or region possesses such disaggregate data. Duthie *et al*. (2007) discussing the data needs of UrbanSim in the US context note that disaggregate data may take months or even a few years to refine to an acceptable level of reliability. At the same time, a powerful modelling potential of UrbanSim offers a strong incentive to apply its tools. The dilemma is discussed in detail in Pattersson *et al*. (2010), where in the prototype UrbanSim applications created for Brussels and Lyon, aggregate data were used. While in Brussels gridcells were created and available aggregate data were disaggregated to them, in Lyon available zone data were applied directly, i.e. each zone corresponded to one "gridcell". The zone version of UrbanSim would better suit the Lyon approach. The need for modelling at this geographical level is recognized by the UrbanSim development team (see UrbanSim project, 2009), and currently the zone version has been developing and testing. An example of a model of residential location choice at a commune level is those developed for the Paris Region (de Palma *et al*., 2005, de Palma *et al*., 2007).

In this study, we continue the zone approach started for Lyon in Pattersson *et al*. (2010). We analyse the Lyon Urban Area, which is the second biggest in France by population, with 1.7 million inhabitants in 2005. The current stage can be seen as a transition from the prototype model to the predictive one. The aim of the paper is to find which "zonal" level is appropriate for the HLCM in the Lyon UrbanSim application: ILOTs[2] (blocks), IRISes[3] (zones used for statistics and traffic analysis), or communes (municipalities). The motivation of this geographically-specific model calibration was unsatisfactory prediction results obtained at the level of ILOTs.

We focus on the question of the reasonableness of simulation results (see Waddell *et al*., 2007) given the existence of historical data for comparison. This back-casting is an important stage before simulating future development. To evaluate the reasonableness of simulation results, Waddell (2002) uses two benchmarks, namely the overall correlation between simulated and observed values and the differences between simulated and observed change (in respect to the HLCM it is population growth during the simulation period) by zone. In our study, beside this, the evenly split population growth is used as an alternative to the HLCM. The output of the HLCM is compared with the alternative method, while the measure of prediction capacity is the deviation of simulated population from actual population in spatial units. The criterion for the choice of an appropriate spatial unit is twofold: the output at this level should be not only better than those for other geographical levels, but also better than the output of the alternative method.

The rest of the paper is organised as follows. The next section describes the Lyon UrbanSim application. In the third section, the HLCM is specified and estimated starting from the level of ILOTs and continuing at the levels of IRISes and communes. The fourth section deals with the model simulation exercises with the three geographical units. The penultimate section describes an

---

[1] In fact, in the parcel version of UrbanSim households and jobs have to be located in buildings, which are situated in land parcels. However, buildings are not directly used in the Lyon application due to the lack of data.

[2] *Unité géographique de base pour la statistique et la diffusion du recensement* (base geographical unit for statistics and census).

[3] *Les îlots regroupés pour l'information statistique* (ILOTs united for statistical information).

alternative model of population distribution with evenly split growth, while the final section concludes.

2. The Lyon UrbanSim application

We exploit UrbanSim version 4.2.2 with graphical user interface. In UrbanSim, the analysed territory geographically consists of two levels: the zones of the transportation model at the upper level and gridcells or land parcels at the lower level. In our study, we apply the framework originally created for gridcells. Thus, each household belongs to a particular gridcell.

The modelling is started with estimation of the HLCM model parameters for the base year using the existing location of households and applying a random utility theory, which assumes that a household selects an alternative with the highest utility for him. Every simulation year, new and relocating households are created and placed with a multonomial logit. In our UrbanSim application, the alternative locations are either ILOTs or IRISes or communes with available residential units.

Data on population are available due to the last general census conducted in France in 1999 and the last demographic estimation for the Lyon Urban Area executed in 2005. Thus, our base year is 1999, the simulation period is the six subsequent years, and the year for comparison between simulated result and actual population is 2005. Figure 1 demonstrates population density in thousand of inhabitants per square kilometre in the base year. The highest concentration of population (as well as employment) is observed in the central part of the region, in the cities of Lyon and Villeurbanne, whose common boundary is shown. Population density decreases in the belt formed around the central part, whose boundary is also shown. Lyon, Villeurbanne and the urbanised belt around them with the total 1999 population of 1.1 million inhabitants are named the Greater Lyon. In the available synthetic population based on “Enquête Ménages Déplacements Lyon 2006”[4], households are distributed among 743 IRISes in the Lyon Urban Area.

Geographically, communes consist of IRISes, which, in their turn, consist of ILOTs, though in the fringe of the Lyon Urban Area many non-urbanised communes contain only one IRIS or even only one ILOT. In our application, each of ILOTs, IRISes and communes is analysed as one “gridcell”, which is located in the centroid of the respective geographical unit and has a size of 100 metres by 100 metres. This size was chosen, because the shortest distance between the centroids of ILOTs is a bit higher than this distance. As such, the centroids of zones are considered as centroids of gridcells. With this, an irregular network of gridcells was created with the same number of gridcells as the number of zones. The distribution of areas and other attributes of the three geographical units is presented in Table 1. For ILOTs, we limit the study area by the boundary of the Greater Lyon, because outside it ILOTs in most cases coincide with IRISes.

Each record of the households table represents one household. For each household, there are data on the gridcell in which it is located; number of people; number of cars; and income group. Income refers to 2006. There are three income groups: poor (lowest 20% income), medium (medium 60%), and rich (highest 20%).

The data on population of IRISes are aggregated to 304 communes[5] in the Lyon Urban Area. For the Greater Lyon, on the contrary, the data are disaggregated to ILOTs in the way as follows. The available data contains population in ILOTs. In each IRIS, we randomly distribute households among ILOTs keeping the ratio of the ILOT population to the IRIS population. As a result,

[4] Household Relocation Survey Lyon 2006.

[5] For Lyon, nine arrondissements are used, while Villeurbanne is considered as one commune.

households in the Greater Lyon are distributed among 5,296 ILOTs, while the difference between the disaggregated population and the actual population in ILOT is 6 people in the worst case.

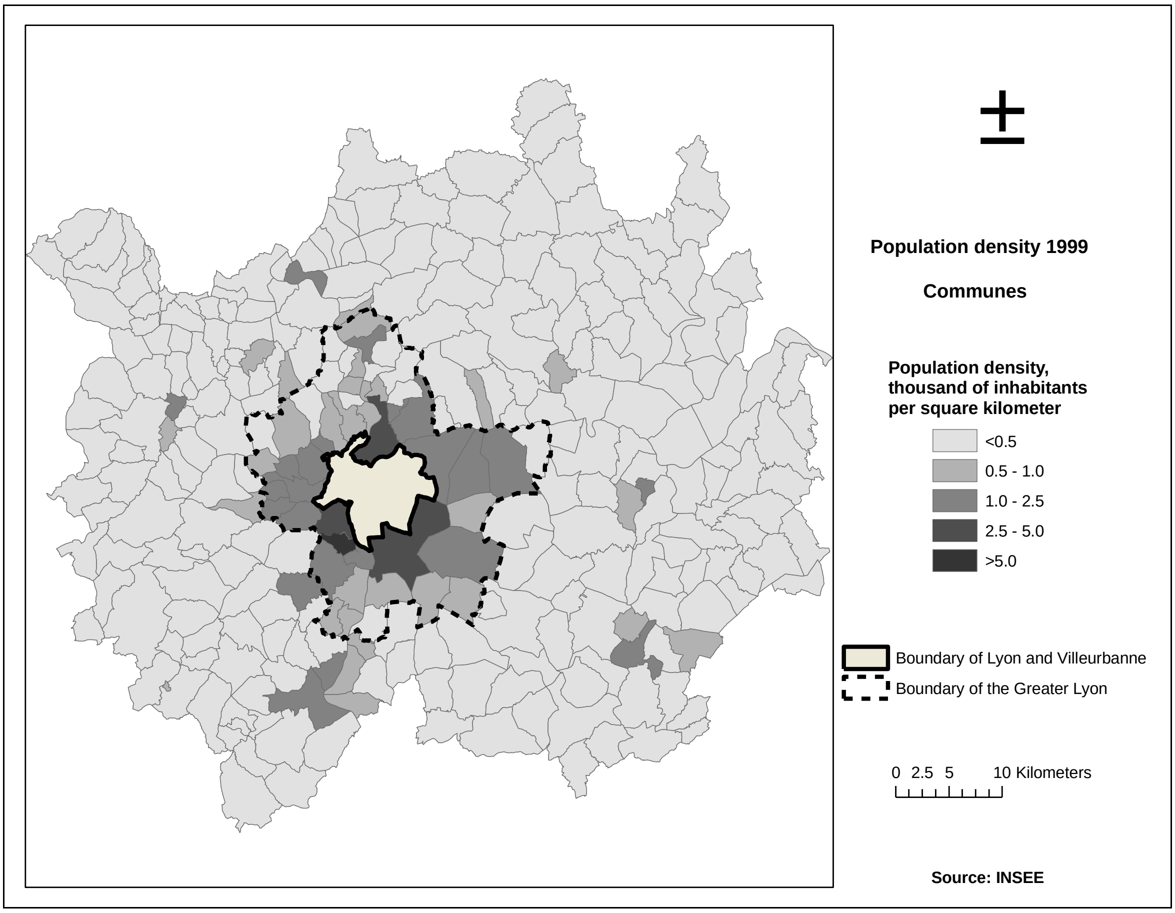

Figure 1. Population density 1999 in communes

Table 1. Description of the three spatial units

Attribute	ILOT	IRIS	Commune
Territory covered	The Greater Lyon	The Lyon Urban Area	The Lyon Urban Area
Total number of spatial units	5,296	743	304
Number of spatial units with dwellings	4,662	742	304
Area, km^2:			
Minimum	0.00002	0.01280	0.40235
Maximum	14.14738	39.99089	39.99089
Mean	0.09272	4.47694	10.94199
Std. dev.	0.54107	6.36776	6.52716
Population 1999:			
Minimum	0	1	96
Maximum	7,951	7,960	116,653
Mean	210	2,124	5,184
Std. dev.	422	1,288	12,033

As exogenous data, UrbanSim needs control totals for population to simulate its spatial distribution in future years. According to INSEE Rhône-Alpes (2007), 38% of current inhabitants in the Lyon

Urban Area changed their accommodation during the previous five years. Though for them we know the percentage of immigrants from outside the Lyon Urban Area, there is no data on emigration. Thus, we use the reported figure assuming that the annual relocation rate, i.e. the probability that households will move from their current location, is 7.6%.

Average real estate price per square metre are calculated separately for ILOTs, IRISes, and communes. The initial data on individual observations of 4,308 apartment and housing sales for the period of 1997-2008 (from *Perval*) are recalculated to 1999 using price indices based on the hedonic price model from Kryvobokov (2009). At the next stage the prices in points are interpolated to raster[6] and zonal statistics is calculated with ArcGIS Spatial Analyst tools for the layers of respective spatial units.

Due to the lack of data on real estate development in simulated years we assume that the housing stock is not changing during the 6-year period, though we admit that this point is not realistic and thus an important path-dependent component of UrbanSim is ignored. Focusing on the HLCM, in this study we also do not update neither real estate prices nor travel times; thus, our independent variables are not endogenous.

Transportation model is external in UrbanSim. Travel times used in the HLCM are generated with a four-step transportation model from MOSART[7] based on GIS and transport modelling software Visum from PTV applying the NAVTEQ road database, see Crozet *et al*. (2008). Travel times by car are estimated for the a.m. peak in a week-day. They are calculated for IRISes and communes. The final goal is to build a transportation-land use model which will be linked to MOSART.

3. Model estimation

The HLCM is estimated with a 10% random selection of households. Using actual households and their actual location from survey data would be a better sampling strategy, but the only available data was synthesised households. The information about recent movers was not available either.

For each location decision, UrbanSim randomly samples 29 alternatives in addition to one more for actual location. This number of alternatives is kept in our application.

Selecting the variables for the HLCM, we account for two principles. First, these are the fundamentals of urban theory developed in the Alonso-Muth model (Alonso, 1964; Muth, 1969) and mirrored in the variables of employment accessibility, real estate prices and income groups of population.

Second, we should avoid the use of the “within walking distance” concept, which exists in UrbanSim, and be careful with scale variables. The reason of the former issue is that our interpretation of spatial units of different size and shape as “gridcells” leads to systematic error: the majority of our “gridcells” have few if any adjacent neighbours and few neighbours in the nearness. The latter issue is connected with the other type of distortion: all the spatial units are represented as if having the same size, but bigger zones can contain larger populations despite of their lower densities. This is also the reason why *average* real estate price per square metre in a spatial unit is used in our application instead of *total* improvement value for different real estate types in gridcells proposed in UrbanSim.

[6] Applying the Inverse Distance Weighted method with 12 neighbours, power 2, and cell size of 10 metres.

[7] *MOdélisation et Simulation de l'Accessibilité aux Réseaux et aux Territoires* (Modelling and Simulation of Accessibility to Networks and Territories)

Whereas comprehensive models of residential location choice count their variables in the dozens (Palma *et al*., 2005; Waddell *et al*., 2007), our parsimonious model specified for the aim of this study, includes a few variables. The same specification of the HLCM is applied at the three geographical levels (Tables 2). In the initial specification, not reported here, average real estate price in a spatial unit used as a one variable had a significant negative coefficient. To understand how real estate prices are perceived by different income groups, we use the first three variables accounting for an interaction between average real estate price and income group of a particular household. The fourth variable checks the influence of residential vacancy rate. Among residential units, 10.4% are vacant in the Lyon Urban Area, whereas for the Greater Lyon this figure is a bit higher: 11.0%. We assume that the Lyon Urban Area is non-monocentric, but instead of focusing on the accessibility to the CBD and several subcentres, we apply the index of employment accessibility calculated for each geographical unit as a sum of all the ratios of employment in IRISes or communes to the square root of travel times to them. For ILOTs, travel times are calculated at the level of IRISes. For IRISes and communes, travel times are estimated for the same level respectively. In the last two variables, the index of employment access interacts with car-ownership status of household.

Table 2. Model estimation

Number	Variable	Coefficient (*t*-value)		
		ILOT	IRIS	Commune
1	Log of average real estate price if high income household	0.505 (8.59)	0.526 (9.66)	0.247 (3.94)
2	Log of average real estate price if middle income household	0.138 (5.36)	0.127 (4.93)	-0.284 (-9.01)
3	Log of average real estate price if low income household	-0.338 (-9.96)	-0.638 (-18.59)	-0.973 (-20.56)
4	Log of residential vacancy rate	0.016 (4.24)	-0.144 (-18.62)	0.165 (17.31)
5	Log of index of employment access if household has a car	-0.599 (-32.53)	-0.310 (-23.67)	0.434 (29.54)
6	Log of index of employment access if household does not have a car	0.969 (28.93)	1.491 (53.64)	2.227 (68.97)
Model parameters				
Null log-likelihood		-168400.085	-225244.297	-225244.297
Log-likelihood		-167289.676	-222899.831	-220769.638
Likelihood ratio test[8]		2220.818	4688.932	8949.318
Number of observations		49512	66225	66225
Number of location choices		4662	742	304

While for ILOTs and IRISes rich households and to lesser degree middle income households can afford locations with more expensive accommodation, at the level of communes middle income households have a negative utility of real estate price. At all the three levels, the coefficient is negative and very significant for poor households.

Residential vacancy demonstrates instability in its sign: being positive and significant for ILOTs, it becomes negative with higher significance at the level of IRISes and again changes its sign but remains highly significant when estimated for communes. One possible explanation of its positive sign can be the following: in the city of Lyon, there are a big number of small ILOTs as well as a small number of arrondissements (analysed as communes) with higher vacancy rates due to high real estate prices.

[8] Chi-Square critical value is 22.458 at the 0.001 level of significance.

The two variables of interaction between an employment accessibility index and the car ownership status of a household are highly significant. All the models highlight that employment accessibility has high utility for households without a car, whereas its influence is negative for car-owning households at the spatial levels lower than commune.

4. Model simulation

Figures 2-4 represent the deviation of predicted population from actual population in 2005 for the three geographical levels. The number of spatial units in each interval is shown in parentheses. To avoid distortions in deviation caused by low-populated spatial units, the differences not higher than 10 people are considered as zero.

At the level of ILOTs (Figure 2), the general tendency is the same as it was in the prototype: population is over-predicted in the central part of the area and under-predicted in the fringe. A remarkable difference exists between the simulation results for ILOTs (Figure 2) and IRISes (Figure 3): the latter model predicts population distribution much better. Moreover, the tendency mentioned for ILOTs is not observed at the level of IRISes. The extreme values of under-prediction to the north-east from Lyon and in the south of the city (Figure 3) refer to sparsely populated areas occupied by university campus and green space in the former case and by industrial area in the latter case. Modelling at the level of communes smoothes the differences to greater degree (Figure 4). For example, in the arrondissements of Lyon deviations varies from -10% in the 7th to +4% in the 4th and the 9th, while in the 1st and the 3rd it is equal to zero.

While in the IRIS-based and commune-based models there are no unplaced households, in the model with ILOTs for Grand Lyon, 75188 people (6% of population 2005) are unplaced. The analysis of simulated results is presented in Table 3, where the parameters of comparison are the correlation coefficient between simulated and observed population in spatial units, the percentages of deviation from actual population and *Moran's I* calculated for population difference and deviation. For the intervals of deviation, percentage of population and percentage of spatial units (in parentheses) are shown. *Moran's I* as a measure of spatial autocorrelation is calculated with the row standardised weight matrix of inverse squared distances.

Correlation of simulated to observed values is very high with the highest value for communes. In all the cases the correlation coefficient is higher than was reported by Waddell (2002) for Eugene-Springfield (0.811 for cells and 0.929 for zones). We admit however that correlation measuring *dependence* between values is not the best benchmark when we simulate population distribution using actual value only for total population: we do not calculate simulated population in territorial units as an explicit function of actual population in the same units.

Table 3 demonstrates that the IRIS-based and the commune-based models have much better prediction parameters in comparison with the ILOT-based simulation. In particular, 96% of IRISes and communes containing 99% of total population have predictions that deviate within 20% from actual population, while only 51% of ILOTs containing half of total population is within that deviation interval. At the same time, the result for ILOTs has low spatial autocorrelation, which is not the case at the upper geographical levels.

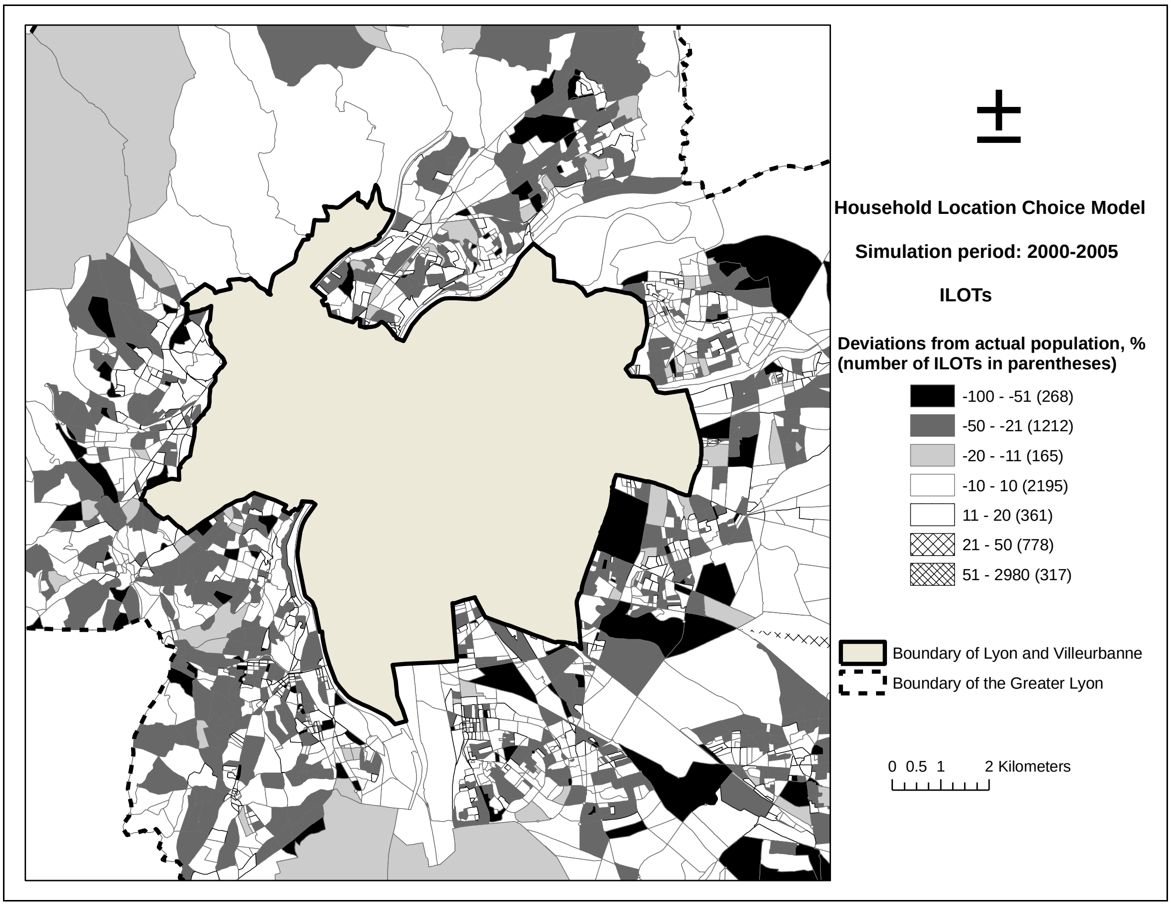

Figure 2. Deviations in Grand Lyon, ILOTs

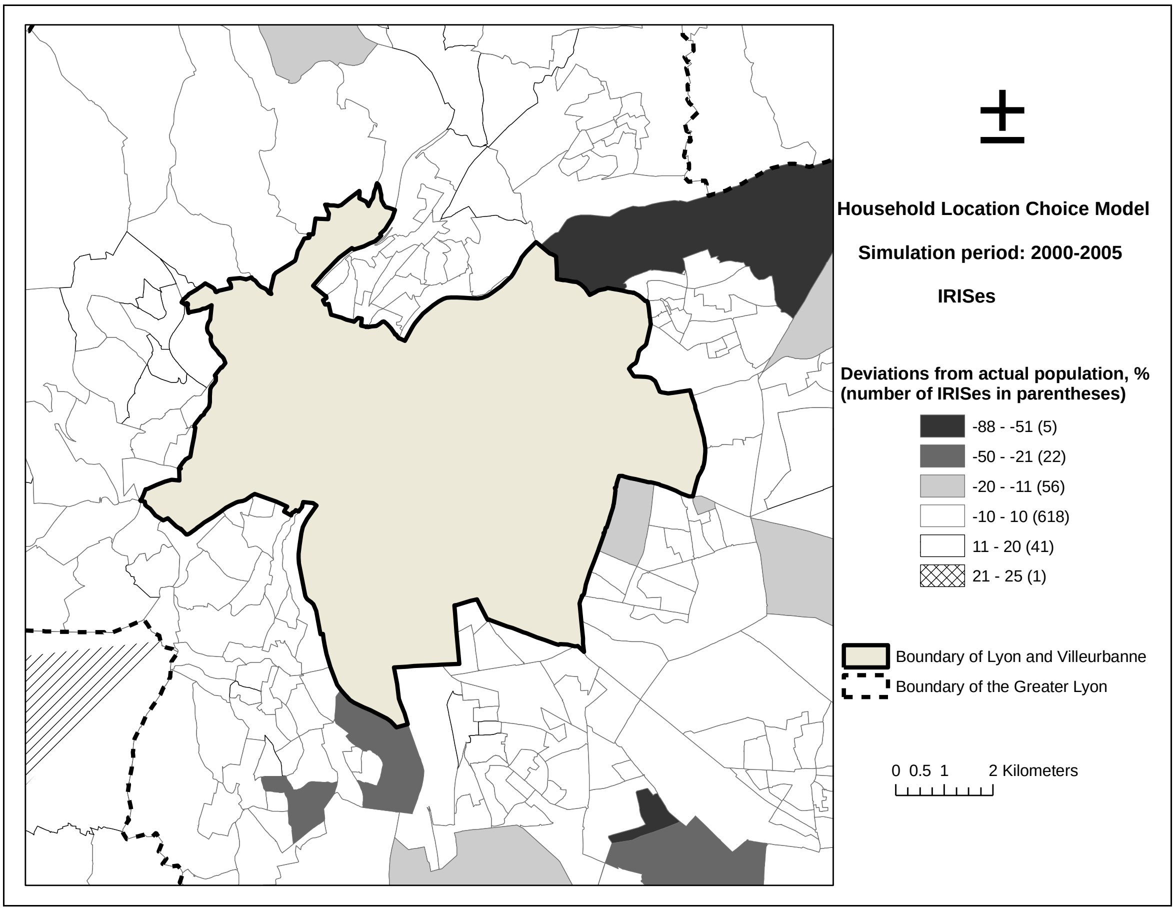

Figure 3. Deviations in the Lyon Urban Area, IRISes

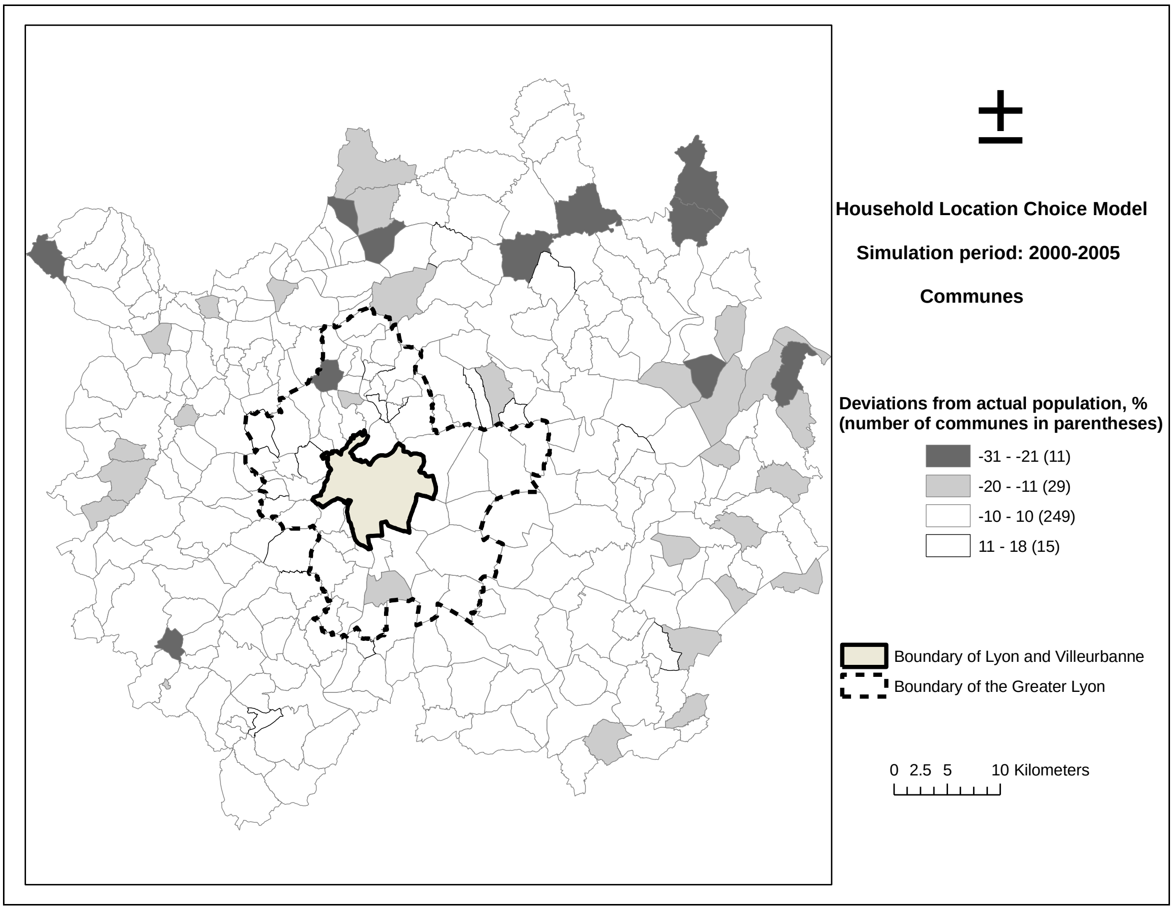

Figure 4. Deviations in the Lyon Urban Area, communes

Table 3. Analysis of simulated population 2005

Parameter	ILOT			IRIS	Commune
	ILOT	Aggregation to IRIS	Aggregation to commune		
Correlation coefficient	0.972	0.946	0.995	0.990	0.999
Within ±5% of actual population*	19 (36)	27 (25)	13 (24)	60 (57)	63 (54)
Within ±10% of actual population*	31 (41)	44 (41)	67 (52)	87 (83)	93 (82)
Within ±20% of actual population*	50 (51)	83 (79)	97 (86)	99 (96)	>99 (96)
Moran's I for difference	0.04	0.30	0.12	0.19	0.12
Moran's I for deviation	0.03	0.24	0.24	0.16	0.17

* Percentage of population and percentage of spatial units (in parentheses) are presented

Beside the simulation results obtained at the three geographical levels, Table 3 also analyses population simulated at the level of ILOTs and aggregated to IRISes and communes in the Greater Lyon. With aggregation, prediction gradually improves, though it is not as good as if started from the upper levels. Spatial autocorrelation increases with aggregation, but for communes it is lower than that for IRISes. Nevertheless, the total number of communes in the Greater Lyon is only 63 that is probably not enough for a proper calculation of *Moran's I*.

Difference between simulated and actual population 2005 by spatial units (which is the same measure as difference between simulated and observed change) is shown in Figure 5 with the same intervals as in Waddell (2002) fro his zones. While in general all the three distributions are not bad (note that the intervals are not equal), the best one comparable with that for Waddell's zones is observed for ILOTs. This is not surprising regarding smaller population at that level.

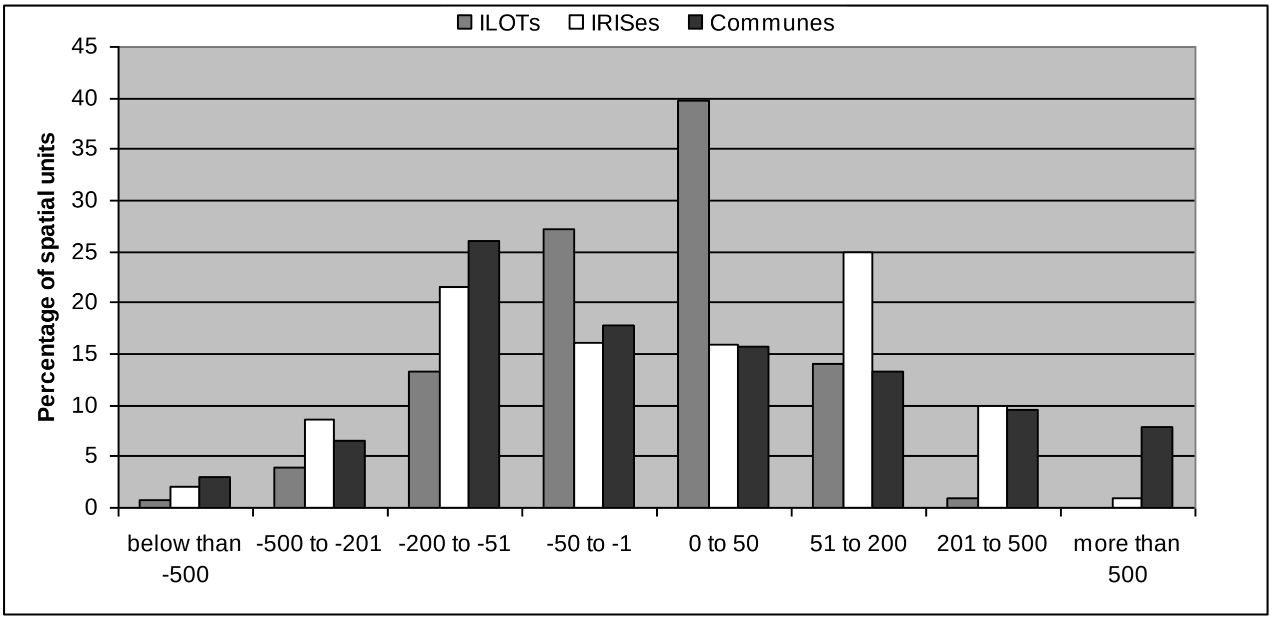

Figure 5. Difference between simulated and actual population 2005 by spatial unit

5. Evenly split population growth

As an alternative estimation of population 2005, it is calculated in each spatial unit as an even increase using the average growth rate for the period of 1999-2005 (Table 4). For the intervals of deviation, percentage of population and percentage of spatial units (in parentheses) are shown. The 6-year growth rate for Grand Lyon is 10.407%, while for the Lyon Urban Area is 10.680%. For the ILOT-based model, the evenly split population growth predicts population distribution much better than the HLCM. Among the three geographical levels, spatial autocorrelation in Table 4 is the lowest for ILOTs, similarly the HLCM results. Though correlation coefficient is not the best measure in this study, we can observe that for IRISes and communes it is practically the same as for the HLCM, while for ILOTs it is higher for the alternative method.

For the HLCMs for IRISes and communes, prediction is comparable with the evenly split population increase: percentages of population and percentages of spatial units within the intervals of deviation are practically the same, though sometimes marginally lower for the HLCM. Difference between simulated and actual population 2005 by spatial units portrayed by Figure 6 is in line with Figure 5, the only significant difference is better distribution for ILOTs for evenly split growth. Spatial distribution of the deviation from actual population is very similar for the two alternative models of population distribution. Figure 7 demonstrating the evenly split growth for communes, has in most cases the same locations of extreme values as Figure 4. It means that the communes, for which it was problematic to predict population with the HLCM, are found problematic in this respect for the alternative model too. Note that both models ignore real estate development during the simulation period.

Table 4. Analysis of population 2005 estimated as evenly split growth

Parameters	ILOT	IRIS	Commune
Correlation coefficient	0.993	0.991	0.999
Within ±5% of actual population*	66 (82)	60 (57)	63 (53)
Within ±10% of actual population*	88 (93)	88 (84)	94 (84)
Within ±20% of actual population*	96 (97)	99 (97)	>99 (97)
Moran's I for difference	0.01	0.18	0.08
Moran's I for deviation	0.04	0.16	0.18

* Percentage of population and percentage of spatial units (in parentheses) are presented

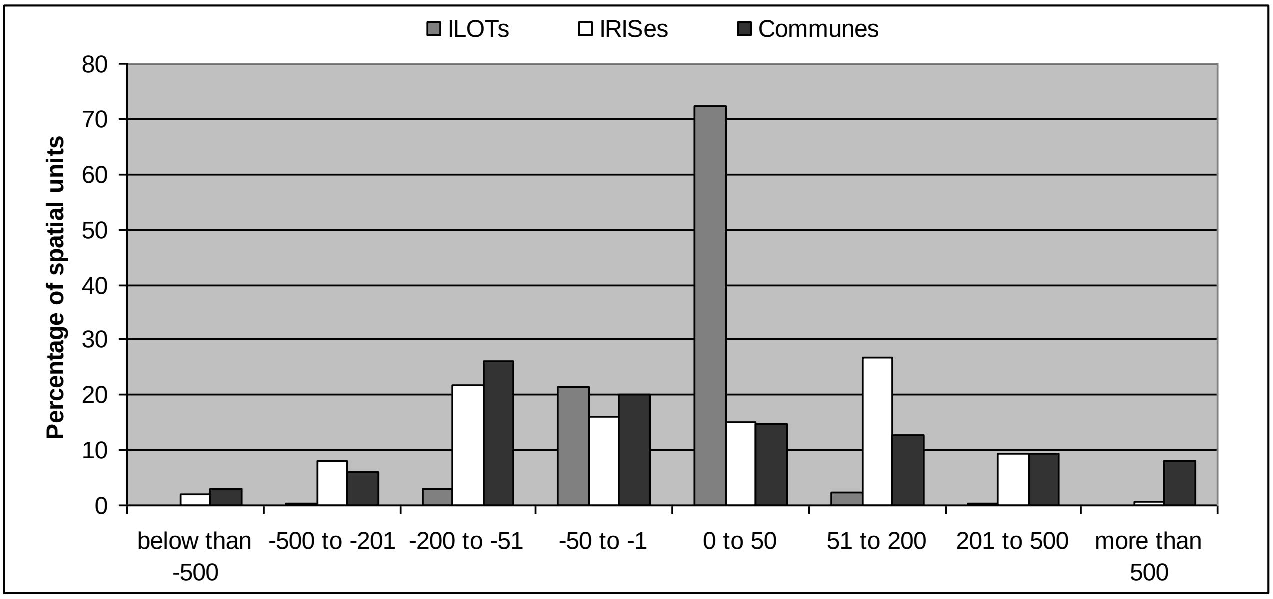

Figure 6. Difference between evenly split growth and actual population 2005 by spatial unit

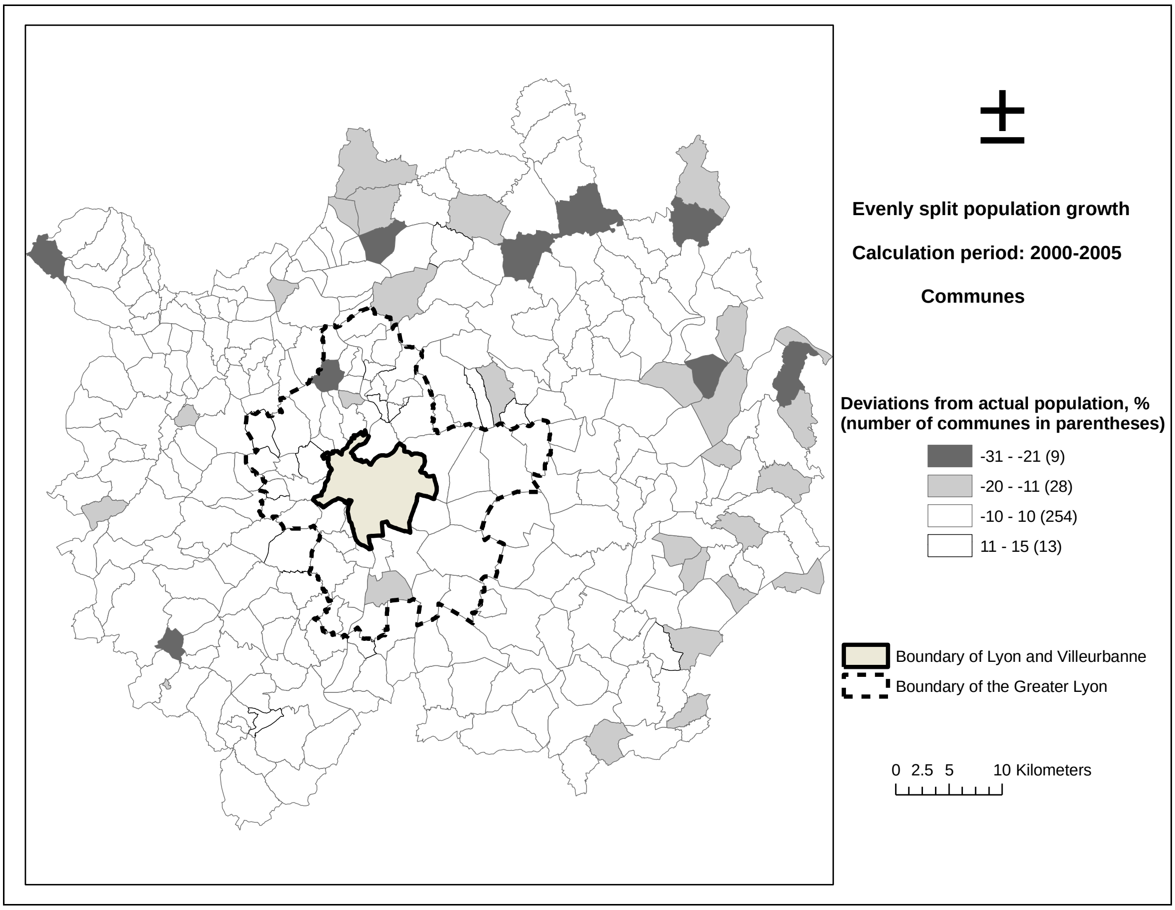

Figure 7. Deviations for evenly split growth, the Lyon Urban Area, communes

6. Conclusion and perspective

With larger spatial units the predictability of the HLCM gradually improves. We see two main explanations of this. First, a natural property of a multinomial logit is that it works better with smaller number of alternatives. Second, there are data errors and prediction errors at the lower geographical level, which are compensated at the upper levels. Data errors include also the errors of disaggregation of synthetic population from the level of IRISes to ILOTs. In this context we can repeat one of the conclusions from Pattersson *et al*. (2010) that only a limited amount of disaggregate information can be drawn from aggregate data. Aggregation of simulated population from ILOTs to upper levels demonstrates how the predictions errors are compensated *a posteriori*. When the model is exploited at the upper levels, the distortion from the heterogeneity of sizes of spatial units becomes smoothed to some extend, especially at the level of communes.

We should note however that spatial autocorrelation is the lowest at the ILOT level for both the HLCM and the alternative estimation with an evenly split population growth, whereas at the upper levels *Moran's I* is much higher. Differences between simulated and actual population by spatial units are also better distributed for ILOTs. These indicators highlight the fact that some information is inevitably lost when applying upper geographical levels.

The prediction parameters of the estimation with an evenly split population increase can be seen as a target for the HLCM. According to the analysis, only at the levels of IRISes and communes, the HLCM provides a result, which is in general not worse than an evenly split growth. Thus, the twofold criterion for the choice of an appropriate spatial unit is not satisfied for ILOTs and is close to being satisfied for IRISes and communes. The best prediction of population distribution is obtained at the level of communes.

In future study of household location choice in the Lyon Urban Area with multinomial logit, it seems logical to continue at the geographical level of communes. It might be easier to obtain and prepare new data at this level avoiding the errors of disaggregation. Our perspective is creation of a more comprehensive model of household location choice with many variables at the selected geographical level, which should better satisfy the chosen criterion.

Acknowledgements

The study is a part of the on-going project PLAINSUDD (Innovative Numerical Platforms of Urban Simulation for Sustainable Development) sponsored through French ANR (number ANR-08-VD-00). The authors are grateful to Alain Bonnafous and Florence Puech for valuable suggestions and Aurélie Mercier for calculation of the origin-destination matrix of travel times for communes.

References

Alonso, W. (1964). *Location and Land Use*. Harvard University Press, Cambridge, MA.

Ben-Akiva, M. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge (Massachusetts), London, 390 p.

Crozet, Y., Marchal, F., Ovtracht, N., Thiebaut, V., Yvon, C., and Bischoff, P. (2008). *Mise en place d'un outil d'aide à la décision et à l'aménagement du territoire : MOSART (MOdélisation et Simulation de l'Accessibilité aux Réseaux et aux Territoires)*, Conference paper. SAGEO Conference. Montpellier.

de Palma, A., Motamedi, K., Picard, N., and Waddell, P. (2005). A model of residential location choice with endogenous housing prices and traffic for Paris region. *European Transport*, 31: 67-82.

de Palma, A., Picard, N., and Waddell, P. (2007). Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region. *Journal of Urban Economics*, 62 (2): 204-230.

Duthie, J., Kockelman, K., Valsaraj, V., and Zhou, B. (B.) (2007). *Applications of Integrated Models of Land Use and Transport: A Comparison of ITLUP and UrbanSim Land Use Models*. Presentation paper. 54th North American Meetings of the Regional Science Association International, Savannah, Georgia, 16 p.

Hunt, J. D., Kriger, D. S., and Miller, E. J. (2005). Current Operational Urban Land-use-Transport Modelling Frameworks: A Review. *Transport Reviews*, 25 (3): 329-376.

INSEE Rhône-Alpes (2007). *La Lettre Analyses No 79 – Aout 2007*, Institut National de la Statistique et des Études Économiques, 4 p. Available at : http://insee.fr/fr/insee_regions/rhone-alpes/themes/syntheses/lettre_analyses/02079/02079_migration_aire_urbaine_lyon.pdf

Kryvobokov, M. (2009). *Is it worth identifying service employment (sub)centres for modelling apartment prices? The case of Lyon, France*, Presentation paper, ERES conference, Stockholm, 29 p.

Muth, R. (1969). *Cities and housing*. University of Chicago Press, Chicago.

Pattersson, Z., Kryvobokov, M., Marchal, F., and Bierlaire, M. (2010). *Disaggregate model with aggregate data: two UrbanSim applications*. 23 p. Forthcoming in *Journal of Transport and Land Use*.

Train, K. (2003). *Discrete Choice Methods and Simulation*. Cambridge University Press, 334 p.

UrbanSim Project (2009). *Opus: The Open Platform for Urban Simulation and UrbanSim Version 4.2.2. Users Guide and Reference Manual*. Center for Urban Simulation and Policy Analysis, University of Washington, Seattle, 296 p.

Waddell, P. (2002). UrbanSim: Modeling Urban Development for Land Use, Transportation and Environmental Planning. *Journal of the American Planning Association*, 68 (3): 297-314.

Waddell, P., Borning, A., Noth, M., Freier, N., Becke, M., and Ulfarsson, G. (2003). Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim. *Networks and Spatial Economics*, 3 (1): 43-67.

Waddell, P., Ulfarsson, G., Franklin, J. and Lobb, J. (2007). Incorporating land use in metropolitan transportation planning, *Transportation Research A*, 41 (5): 382–410.