Analysis and prediction of household location choice in Grand Lyon with urban land use simulation tool UrbanSim

Marko Kryvobokov, Nicolas Ovtracht, Valérie Hiebaut

Analysis and prediction of household location choice in Grand Lyon with urban land use simulation tool UrbanSim

Marko Kryvobokov, Nicolas Ovtracht, Valérie Thiebaut

Laboratory of Transport Economics (LET), Lyon, France

Abstract

Modelling framework UrbanSim is applied for analysis and prediction of residential location choice of households in Grand Lyon for understanding the mechanism of spatial distribution of population. The aim of the paper is the evaluation of the predictability of the residential location choice in the analysed urban area with the UrbanSim application. The social component of location choice is mirrored in different variables for different income levels of population as well as in attributes connected with car-ownership. The paper concludes that the predictability can be increased by adding other UrbanSim models to the application.

Keywords: household location choice, land use modelling, UrbanSim, GIS, Grand Lyon.

Introduction

Urban area is a complex system, where its components are interacting and forming interconnected and interdependent relations. Understanding the relations between population, employment, land use and transportation is a necessary precondition of efficient urban planning and management aiming at sustainable development. As Wegener (2004) points out, it is difficult to empirically isolate impacts of land use on transport and vice versa, and mathematical models of those impacts are the only method by which the effects of individual determining factors can be analysed by keeping all other factors fixed.

Urban simulation models are important tools for decision-makers as well as for researchers in urban planning and transportation. The overviews of the contemporary frameworks representing the state-of-the-art of transportation-land use modelling can be found in Wegener (2004) and Hunt *et al*. (2005). One of the focuses of this kind of modelling is location choice made by households in urban area.

Residential location choice is a field of interest of many studies of European and American cities. Sermons and Koppelman (2001) deal with a household location choice model for the San Francisco Bay Metropolitan Area; Bürgle (2006) applies such a model from UrbanSim for the analysis of the Greater Zurich Area. The two latter studies focus on the integration of land use and transportation models: Pinjari *et al*. (2008) create an integrated simultaneous choice model of residential location, vehicle and bicycle ownership and commute tour mode in the San Francisco Bay Area; de Palma *et al*. (2005) integrate a land use model from UrbanSim with a dynamic traffic model METROPOLIS for the Paris region.

In the current study, residential location choice is analysed in the central part of the Lyon Urban Area, which is the second biggest in France by population. The demographic dynamics in the Lyon Urban Area, which was +0.8% per year during the period of 1999-2005, was higher than the average in metropolitan France. In particular, the population growth in the area was higher than those in Paris, Marseille and Lille (INSEE Rhône-Alpes, 2007). We study this period in order to understand the mechanism of spatial distribution of population over the area with the focus on Grand Lyon. For this, urban simulation framework UrbanSim is applied. The study is part of the project PLAINSUDD (PLAteformes numériques INnovantes de Simulation Urbaine pour le Développement Durable – Innovative Numerical Platforms of Urban Simulation for Sustainable Development) sponsored through French ANR.

UrbanSim was developed at the University of Washington (Waddell *et al*., 2003) as an open-source project (UrbanSim Project, 2008). Its popularity among researchers and practitioners in different countries of the world and in Europe in particular is increasing (Waddell *et al*., 2008). The experience with UrbanSim in France includes the Paris application (De Palma *et al*., 2005; De Palma *et al*., 2007) and the Lyon application (Patterson *et al*., 2009).

UrbanSim includes the system of models, which are estimated for the base year and then are used for simulation of future years. In the current paper we use the model of household location choice.

Thus, a part of the land use component of UrbanSim is exploited. Travel times are generated with a transportation model from MOSART (MOdélisation et Simulation de l'Accessibilité aux Réseaux et aux Territoires – Modelling and Simulation of Accessibility to Networks and Territories), see Crozet *et al*. (2008). MOSART is a numerical platform for modelling included in the PLAINSUDD project.

At the stage of estimation, the actual geographical distribution of population is analysed using actual data for 1999. In simulation, we predict the location choice of households for six subsequent years on the base of the estimated parameters. The simulated results are compared with actual population distribution in 2005. The aim of the paper is the evaluation of the predictability of the residential location choice in Grand Lyon with the UrbanSim application. The prospective of the further use of UrbanSim is the simulation of distribution of population as well as other important attributes in future years.

UrbanSim

UrbanSim framework includes a system of models (Figure 1), which are used for simulation of distribution of people, jobs, real estate development and real estate prices on a yearly basis. The methodology applied in UrbanSim includes the multinomial logit models for location choice modelling and the Ordinary Least Squares regression models for residential land share and real estate prices modelling. UrbanSim models include the Household Location Choice Model, the Employment Location Choice Model, the Development Project Location Choice Model, the Residential Land Share Model, the Land Price Model, and the Real Estate Price Model. The outcome of models can be used as inputs in other models. An important feature of UrbanSim is a possibility to apply different scenarios of urban development. Transportation model is an external model in UrbanSim. GIS, being an external system in Figure 1, is an indispensable tool for data preparation and results interpretation.

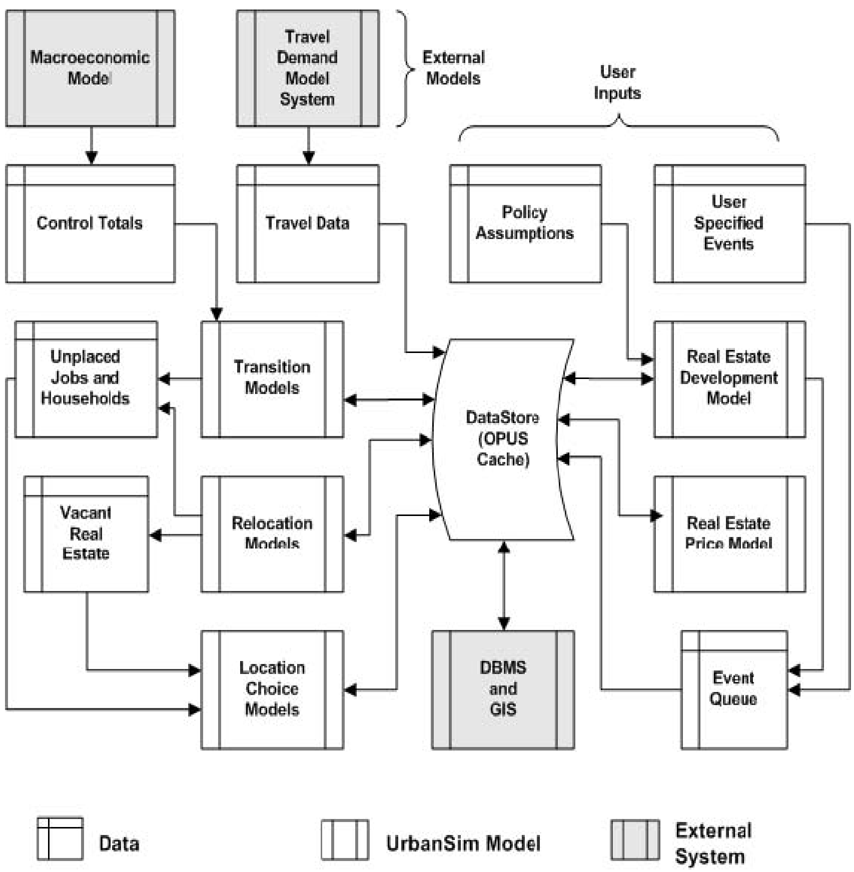

Figure 1. UrbanSim model components and data flow (source: UrbanSim Project, 2008)

For simulation of future years, the Household Location Choice Model needs exogenous data – control totals for population. The probabilities that households will move from their current location are user defined. Every simulation year, new and relocating households are created. The households are then placed with a multonomial logit applying a random utility theory, which assumes that a decision maker selects an alternative with the highest utility for him. The theory of discrete choice analysis can be found in Ben-Akiva and Lerman (1985). In our UrbanSim application, decision makers are households, and alternative housing locations are ILOTs.

Data

The area of analysis, Grand Lyon[1], is the core of the region. Grand Lyon includes 63 communes and city districts with the total area of 491 km^2 and the 1999 population of more than 1.1 million inhabitants. Grand Lyon consists of two parts. The central part (its boundary is shown in violet in Figure 2) is composed by the cities of Lyon and Villeurbanne with the area of 63 km^2 and the population of more than half a million inhabitants. These two cities, having a common planning structure and transportation network, make up the regional centre with the highest concentration of population and employment. The urban belt around the central part is almost seven times bigger, but much less densely populated; its population is only slightly higher than that of the central part. The belt is composed by densely built suburbs, some small villages, and also some parks, water bodies, agricultural land and other open space.

Geographically, the territory is divided by ILOTs[2] and traffic analysis zones (TAZs), which correspond to IRISes[3] from INSEE[4]. Communes consist of IRISes, which, in their turn, consist of ILOTs, though in the fringe some communes contain only one ILOT. In our application, ILOTs are analysed as “gridcells”, which are located in the centroids of ILOTs and have a size of 100 m by 100 m. While travel times are used in UrbanSim at the level of TAZs, the more detailed geographical unit of analysis is ILOT.

Grand Lyon is represented as 5,296 ILOTs and 479 TAZs. Figure 2 shows population density in thousands of inhabitants per square kilometre in TAZs. Each ILOT represented as a gridcell includes information on the following attributes: zone, to which it belongs; coverage by water, wetland, green area, open space; number of residential units; population density; residential vacancy rate; and average real estate price per square metre. Average residential vacancy rate in ILOTs in 1999 is 11.5%. Average real estate price was calculated in each ILOT using data on individual observations of 4,308 apartment and housing sales (from *société Perval*) and applying geographical interpolation with ArcGIS Spatial Analyst tools.

Data on population are available due to the last general census of the population in France conducted in 1999 and the last demographic estimation for the Lyon Urban Area executed in 2005. In the available synthetic population based on “Enquête Ménages Déplacements Lyon 2006”[5], households are distributed among TAZs. The available data contains population in ILOTs. In each TAZ, we redistributed households among ILOTs keeping approximately the same ratios of population in each ILOT to population in its TAZ. The best result was obtained with the random distribution of households belonging to the same TAZ, between ILOTs of that TAZ.

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 working people, number of cars, and income group. Population refers to 1999, whereas income to 2006. There are three income groups: poor (lowest 20% income), medium (medium 60%), and rich (highest 20%).

---

[1] We do not consider the communes of Givors and Grigny, which compose an enclave outside the urban belt since the beginning of the 2000s.

[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).

[4] *Institut national de la statistique et des études économiques* (National Institute for Statistics and Economic Studies).

[5] Household Relocation Survey Lyon 2006.

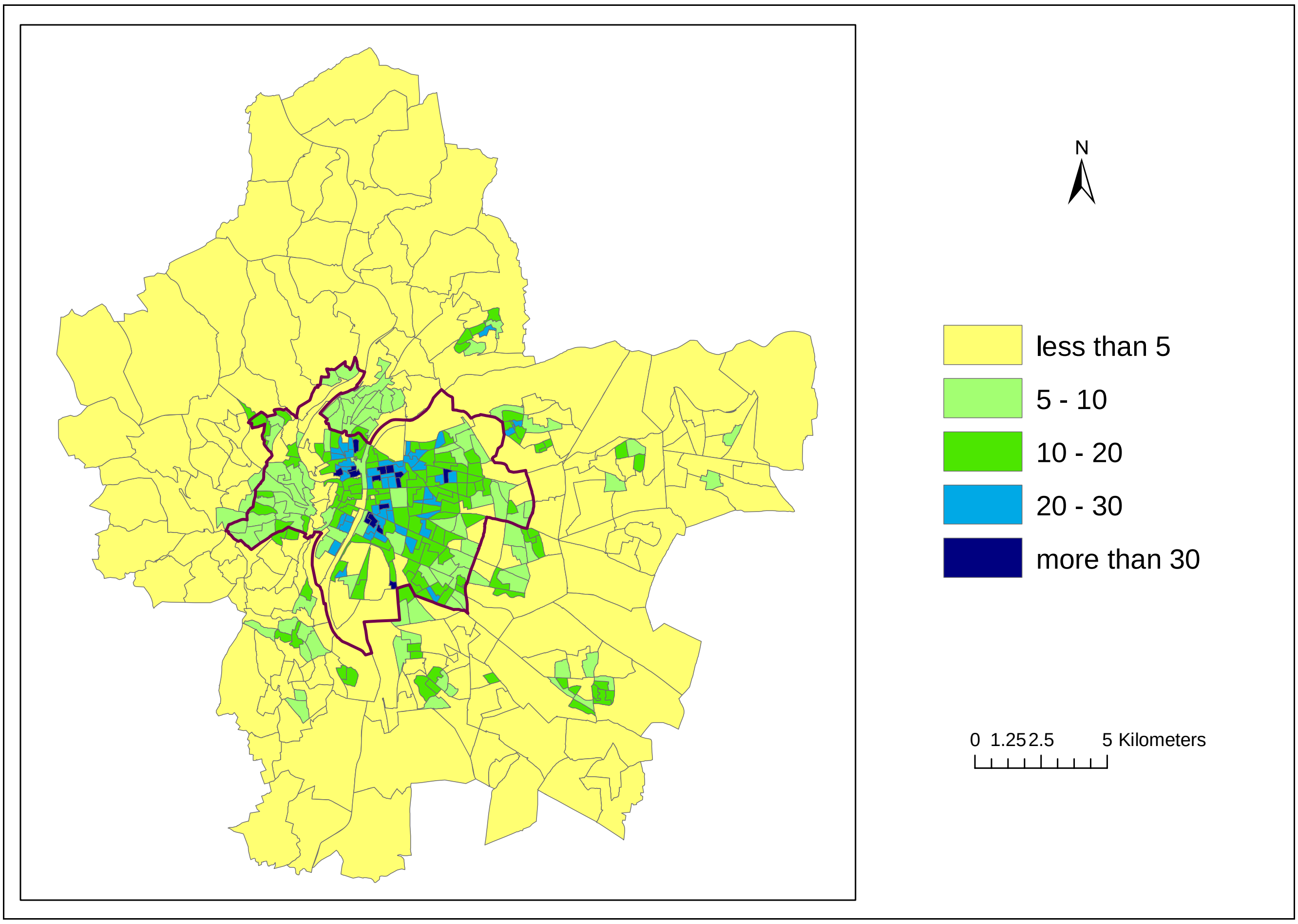

Figure 2. Population density in Grand Lyon in 1999, thousands of inhabitants per square kilometre

According to Insee Rhône-Alpes (2007), 38% of current inhabitants in the Lyon Urban Area changed their accommodation during the last five years. Of them, 13% were living in different *départements* or *regions*. As we do not have data about emigration, we use the first mentioned figure for households, i.e. annual relocation rate for households is 7.6%.

In general, in UrbanSim we operate only with control totals for population in each of future years, i.e. in this respect there is no difference between new families, emigrants and borne children – they all are considered as increase in population. Population growth during the period of 1999-2005 is assumed to have a constant rate.

Transportation model is external in UrbanSim. Travel times are generated with a four-step transportation model from MOSART based on GIS and transport modelling software Visum from PTV. NAVTEQ road database is applied. Travel time estimated for public transport for the a.m. peak in a week-day is based on the schedule of regional train and on the frequency of more than one hundred lines of buses, trams and metro. The travel times obtained with MOSART are then applied in UrbanSim.

Model estimation

UrbanSim Project (2008) grouped the independent variables of the Household Location Choice Model into the following categories: real estate characteristics, regional accessibility and urban design and scale. Empirical studies of household location choice (e.g. De Palma *et al*., 2005; Bürgle, 2006; Pinjari *et al*., 2008) report significance of a considerable number of interaction variables, which combine the attributes of households with the characteristics of location.

The Household Location Choice Model estimated for the Grand Lyon is presented in Table 1. For estimation, a randomly selected 10% sample of households is used. In the initial specification, real estate price used as a one variable had a significant negative coefficient. To understand how real estate prices are perceived by different income groups, three interaction variables (number 1-3) were created.

The outcome is the following: rich households and to lesser degree middle income households can afford locations with more expensive accommodation, whereas low income households live in less expensive locations, and the coefficient is more statistically significant for the high income and low income groups. The next variable (number 4) in the model is the logarithmic transformation of residential vacancy rate. Its positive sign can be explained by a big number of small ILOTs in the centre of Lyon with higher vacancy rates due to high real estate prices. The two last attributes (number 5-6) are the variables of interaction between the logarithm of the index of employment accessibility and the car ownership status of a household. The index of employment accessibility is calculated as a sum of ratios of zonal employment to the square root of travel times. These highly significant variables demonstrate that employment accessibility has high utility for households without a car, whereas its influence is negative for car-owning households.

Table 1. Model estimation

Number	Variable	Coefficient	*t*-value
1	Log of average real estate price if high income household	0.363	6.78
2	Log of average real estate price if middle income household	0.097	4.11
3	Log of average real estate price if low income household	-0.213	-6.92
4	Log of residential vacancy rate	0.010	2.62
5	Log of index of employment access if household has a car	-0.415	-26.98
6	Log of index of employment access if household does not have a car	0.622	23.77

Null log-likelihood: -168400.085
Log-likelihood: -167684.482
Likelihood ratio test: 1431.206
Number of observations: 49512

Simulation results

As there are no available data on residential development in the analysed period, in the current scenario of simulation there is no new real estate development. It means that the number of residential units in 2005 is the same as in 1999. This can be a reason of a considerable number of unplaced households, which compose 6% of population.

The comparison of predicted population with actual population in 2005 is presented in Figure 3. Relative differences between predicted and actual population are shown with number of ILOTs in each interval in parentheses. To avoid distortions in relative differences caused by low-populated ILOTs, the differences not higher than 10 people are considered as zero. As is seen in Figure 3, the level of under-prediction is higher due to existence of unplaced households. In general, population is mainly under-predicted in suburbs and over-predicted in the central parts of Lyon and Villeurbanne. In the arrondissements of Lyon and Villeurbanne the two-sample F-test shows the equality of variances between the simulated and actual population.

In relative figures of comparison between predicted and actual population, 37% of ILOTs (with 20% of population) have the differences within the interval of ±5%, 42% (31% of population) within the interval of ±10%, and 51% (50% of population) within the interval of ±20%.

Moran's I calculated with the row standardised weight matrix of inverse squared distances is equal to 0.05 for differences between predicted and actual population and 0.04 for relative differences. Thus, spatial autocorrelation of both measures is low.

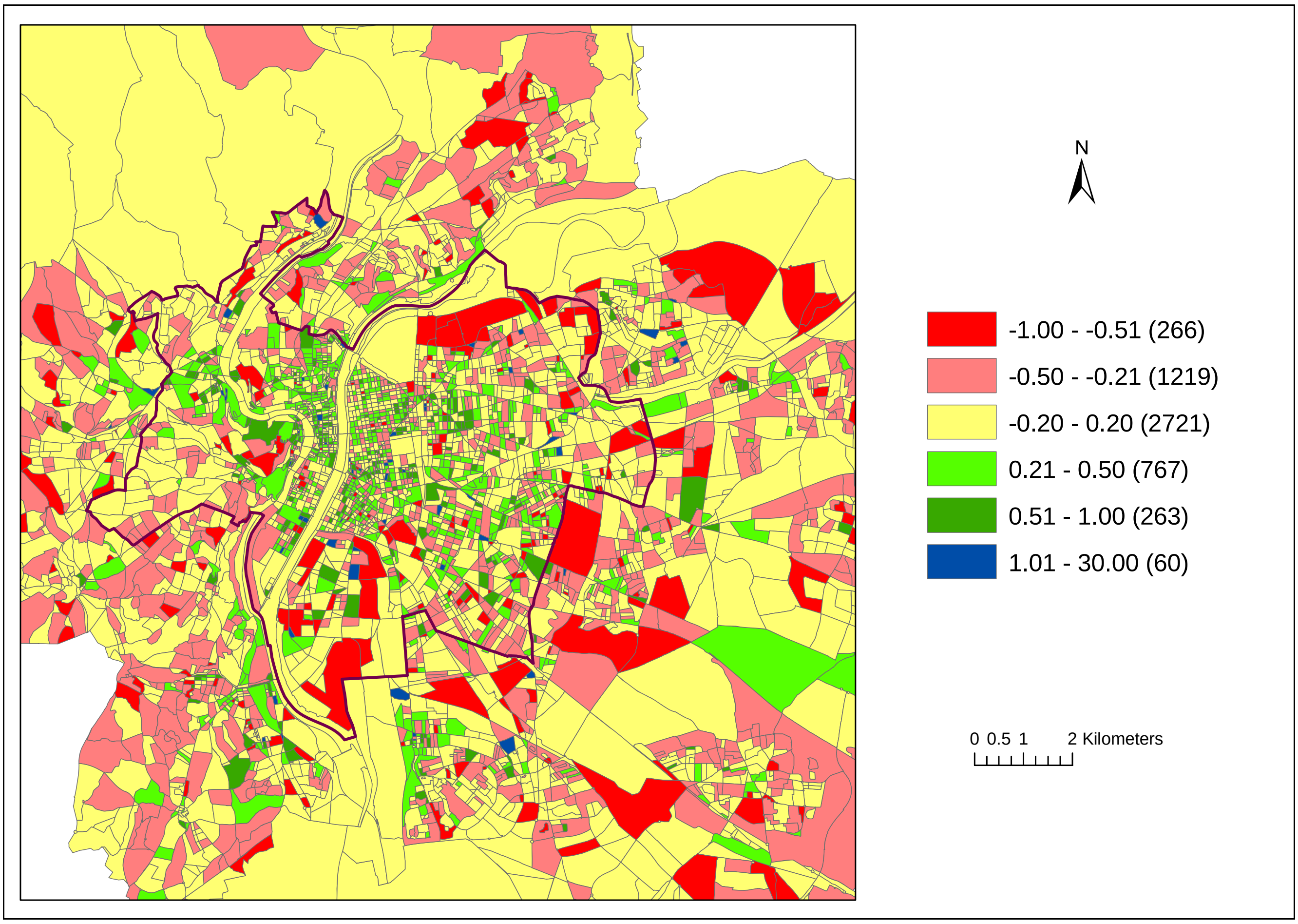

Figure 3. Relative differences between simulated and actual population in 2005

Conclusion

The paper is devoted to the UrbanSim application for Grand Lyon based on ILOTs considered as gridcells. The focus was the household location choice estimated for 1999 and simulated for six subsequent years. The residential location choice model specified with quantitative variables highlights the following findings. The most significant attributes are the employment accessibility interacted with car ownership. Accessibility to employment is especially important for households without a car. The interaction between real estate price and income levels of population shows that rich households and to lesser degree middle income households choose locations with expensive accommodation, whereas low income households live in less expensive locations. Residential vacancy rate has a positive influence, but its significance is low.

The simulation outcome contains a detailed geographical distribution of households. The predictability of simulation is measured by comparison with actual population. We observe the same general tendency as in the Lyon Urban Area model in Patterson *et al*. (2009): population is over-predicted in its very central part and under-predicted in the urban belt.

The applied scenario did not include any real estate development in simulated years. This can be a reason of a considerable number of unplaced households and, consequently, of the higher level of under-prediction of population then its over-prediction. It underlines the fact that it is problematic to simulate residential location choice without links with other UrbanSim models, in particular the Development Project Location Choice Model and the Real Estate Price Model.

Despite low spatial autocorrelation of differences between actual and predicted population and good prediction in the central part of Grand Lyon in total, only half of population is predicted within the interval of ±20% of relative difference. Adding other UrbanSim models, the application can be developed further.

References

Bürgle, M. (2006). *Residential location choice model for the Greater Zurich area*. Conference paper. 6th Swiss Transport Research Conference, 19 p.

Crozet, Y., Marchal, F., Ovtracht, N., Thiebaut, V., Yvon, C., 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., Motamedi, K., Picard, N, Waddell, P. (2007). Accessibility and Environmental Quality: Inequality in the Paris Housing Market. *European Transport*, 36: 47-64.

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,* 4 p. Available at : http://insee.fr/fr/insee_regions/rhone-alpes/themes/syntheses/lettre_analyses/02079/02079_migration_aire_urbaine_lyon.pdf

Patterson, Z., Kryvobokov, M, Marchal, F., and Bierlaire, M. (2009). *Disaggregate Models with Aggregate Data: Two UrbanSim Applications*, working paper, 44 p.

Pinjari, A. R., Pendyala, R. M., Bhat, C. R., and Waddell, P. A. (2008). *Modeling the choice continuum: An integrated model of residential location, auto ownership, bicycle ownership, and commute tour choice decisions*. TRB 2008 Annual Meeting CD_ROM, # 08-2647, 33 p.

Sermons, M. W. and Koppelman, F. S. (2001). Representing the differences between female and male commite behavior in residential location choice models. *Journal of Transport Geography*, 9 (2): 101-110.

UrbanSim Project (2008). *Opus: The Open Platform for Urban Simulation and UrbanSim Version 4 – Reference Manual and Users Guide*, Technical Report, Center for Urban Simulation and Policy Analysis, University of Washington, Seattle, Washington, 188 p.

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., Wang, L. and Liu, X. (2008). UrbanSim: An Evolving Planning Support System for Evolving Communities. Paper forthcoming in *Planning Support Systems*, edited by Brail, R. Lincoln Institute for Land Policy, 31 p.

Wegener, M. (2004). Overview of Land-use transport models, in D. A. Heshner and K. Button (eds.), *Transport Geography and Spatial Systems*, Handbook 5 of the *Handbook in Transport*, Pergamon/Elsevier Science, Kidlington, the UK: 127-146.