

**LES CHOIX RESIDENTIELS
DANS L'AIRE URBAINE DE LYON**

**Résultats d'une modélisation des interactions entre
transports et urbanisme**

**Agence d'Urbanisme de Lyon
Laboratoire d'Economie des Transports**

SOMMAIRE

INTRODUCTION GENERALE	5
SIMULER LES MOBILITÉS POUR UNE AGGLOMÉRATION DURABLE : LE PROJET SIMBAD	7
COMMENT PRENDRE EN COMPTE LA LOCALISATION RESIDENTIELLE DES MENAGES ?	9
La microéconomie urbaine comme cadre théorique	9
Une modélisation en deux étapes : décision de mobilité résidentielle puis choix de localisation	10
DÉMÉNAGER OU PAS ? UN MODÈLE DE MOBILITÉ RÉSIDEN­TIELLE	12
OÙ SE LOCALISER ? UN MODÈLE DE CHOIX DE L'IRIS DE RÉSIDENCE	13
Les variables utilisées	13
Le modèle retenu : un fort impact des variables d'accessibilité	15
CONCLUSION : TESTER L'INFLUENCE DES POLITIQUES DE TRANSPORT ET D'URBANISME SUR LES LOCALISATIONS DES MENAGES	19
Les principaux enseignements sur les logiques de localisation résidentielle	19
Quelles pistes d'amélioration peuvent être proposées ?	20
ELÉMENTS DE BIBLIOGRAPHIE	21

INTRODUCTION GENERALE

Le fait que les transports et l'aménagement du territoire s'influencent mutuellement est une réalité généralement reconnue. La prise en compte et la compréhension de cette interaction entre transports et urbanisme est essentielle pour prendre des décisions sur un projet de nouvelle infrastructure de transport tout en ayant une idée de ses effets à court, moyen et long terme sur l'occupation du sol, en termes de localisation des ménages et des activités économiques. Réciproquement, il est important de mesurer les effets des projets de développement urbain sur les déplacements et les trafics induits.

Les enjeux des politiques publiques en matière de planification et de coordination entre le volet transports et le développement urbain sont donc multiples et apparaissent généralement communs aux agglomérations urbaines. Diminuer le temps dédié aux déplacements quotidiens de la population en assurant une cohérence entre la répartition de l'habitat et des emplois sur le territoire, limiter le plus possible l'étalement urbain, maîtriser l'usage de l'automobile et développer l'usage des modes doux, favoriser l'urbanisation autour des nœuds de transport (gares, stations de transport collectif en site propre...), lutter contre la ségrégation sociale dans le territoire à travers des dessertes plus importantes des quartiers sociaux, etc. : autant de questions qu'aborde l'Agence d'Urbanisme de Lyon et que l'on retrouve posées dans les autres grandes agglomérations françaises.

C'est dans ce cadre que l'Agence d'Urbanisme soutient le projet SIMBAD (Simuler les MoBilités pour une Agglomération Durable) développé par les chercheurs du Laboratoire d'Economie des Transports à Lyon, et également financé au niveau national par le Programme de REcherche et D'Innovation sur les Transports (PREDIT).

Un premier cahier a été publié en 2008 pour faire un état des lieux général de la modélisation des interrelations entre transports et urbanisme. Ce second cahier présente aujourd'hui les résultats des travaux de SIMBAD en matière de localisation des ménages au sein de l'aire urbaine de Lyon, travaux réalisés dans sa thèse par Marius Homocianu suivi scientifiquement par le LET et financé par l'Agence d'Urbanisme dans le cadre d'une bourse CIFRE.

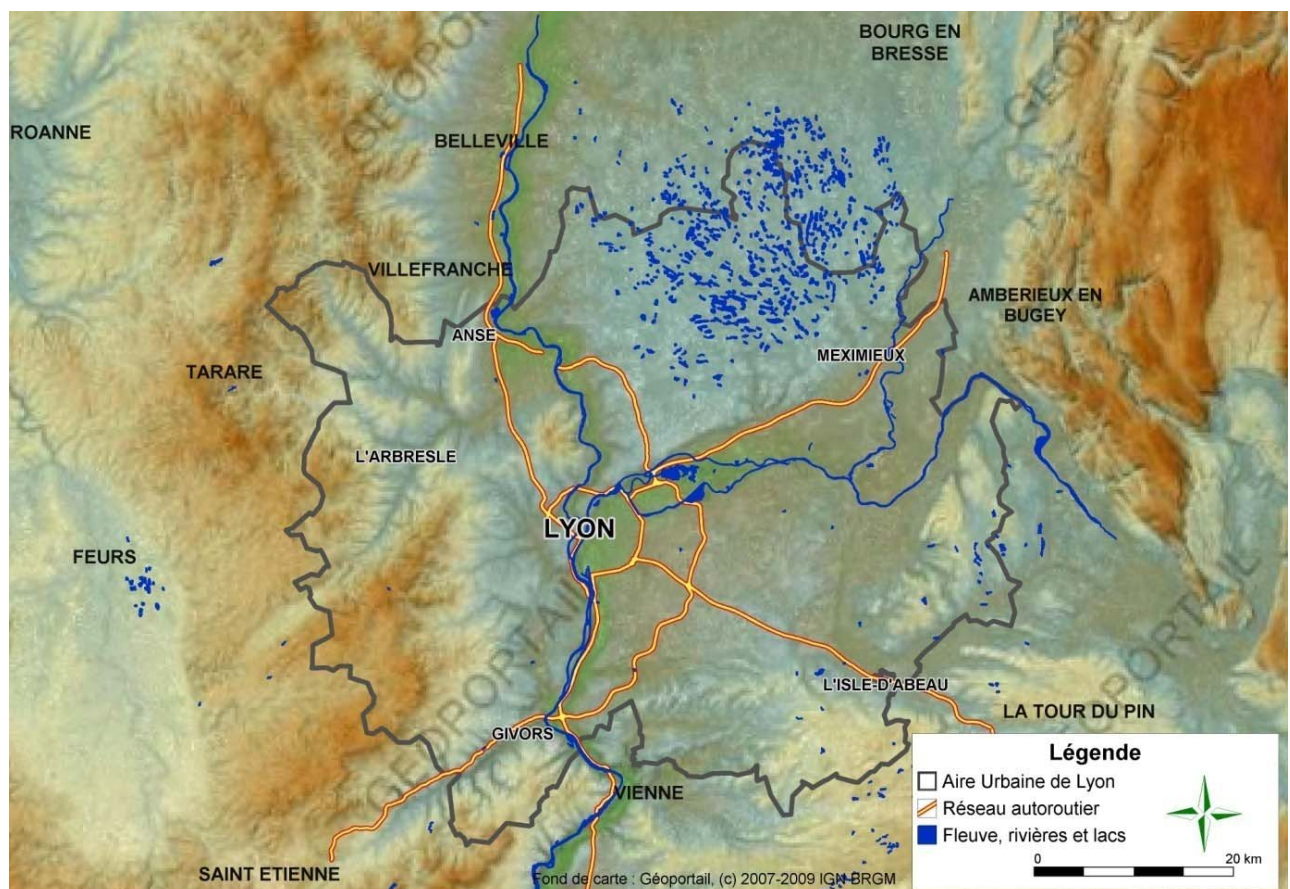
Après un rapide retour sur les objectifs et l'architecture du projet SIMBAD (partie 1), nous présentons plus précisément la manière dont la question de la localisation des ménages a été abordée (partie 2) et les outils qui ont été utilisés pour la traiter (partie 3). Les résultats sont ensuite présentés, en distinguant d'une part les logiques de mobilité résidentielles (partie 4) et d'autre part les facteurs explicatifs des localisations proprement dits (partie 5).

SIMULER LES MOBILITÉS POUR UNE AGGLOMÉRATION DURABLE : LE PROJET SIMBAD

Montée des préoccupations environnementales locales et inquiétude croissante quant aux effets du réchauffement climatique, contraintes financières fortes, sensibilité des citoyens vis-à-vis du libre accès à la ville... Dans ses différentes dimensions, environnementale, économique et sociale, le thème du développement durable intervient aujourd'hui de manière récurrente dans les discours sur l'avenir des transports.

Dans ce contexte, le Laboratoire d'Economie des Transports, en lien avec l'Agence d'Urbanisme de Lyon, a développé un axe de recherche consacré au développement du modèle prospectif SIMBAD (Simuler les MoBilités pour une Agglomération Durable). Conçu à l'échelle du bassin de vie d'une agglomération (les enjeux de la croissance urbaine dépassent largement le cadre des centres-villes), son application a été développée sur le territoire de l'aire urbaine de Lyon (Carte 1) à l'horizon 2025. Il veut rendre compte des dimensions environnementales, économiques et sociales des trafics qui s'y réalisent. Ainsi différentes politiques de transport et d'urbanisme peuvent être simulées de manière à envisager leurs conséquences sur ces 3 axes du développement durable et à mieux prendre en compte les liens qui existent entre ces derniers. Par ailleurs, l'idée est de développer un outil d'aide à la réflexion plus qu'un outil d'évaluation des projets. L'approche est donc de type stratégique, permettant de rendre compte des impacts de politiques contrastées à l'échelle de l'agglomération plus que de projets particuliers à une échelle fine.

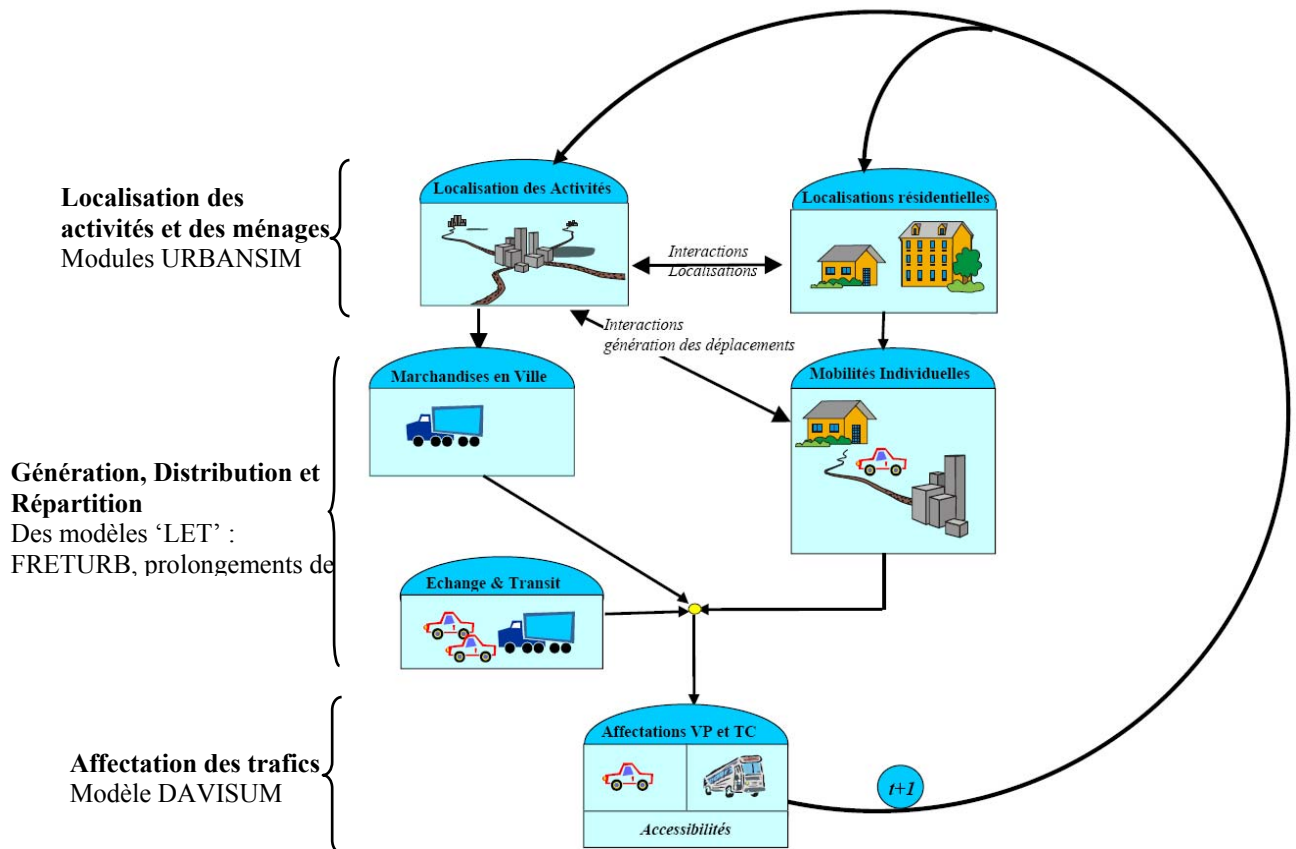
Carte 1 : L'aire urbaine de Lyon comme territoire d'application de SIMBAD



Source : à partir de Géoportail © 2007-2009 IGN-BRGM

Pour répondre à cet objectif de mesure de l'évolution d'indicateurs de mobilité durable, SIMBAD a été conçu pour enchaîner plusieurs grandes étapes de modélisation « localisation des ménages et des établissements » ⇒ « mobilités » ⇒ « affectation du trafic » qui se bouclent sur 25 pas d'un an (Nicolas et al., 2009 ; voir Figure 1).

Figure 1 : L'architecture du projet SIMBAD



Source : Nicolas, 2009

1. A la période initiale, les localisations des activités et des ménages sont données.
2. Le croisement entre la localisation des activités et les caractéristiques des ménages résidant dans chaque zone permet de déduire les besoins de déplacements entre chaque zone :
 - ✓ Une suite de modules prend en compte la mobilité des personnes en suivant une logique traditionnelle, avec la génération des déplacements, leur distribution entre les zones du territoire considéré, leur répartition modale puis horaire.
 - ✓ Un travail spécifique est réalisé pour les transports de marchandises en ville.
 - ✓ Enfin, l'introduction des flux d'échanges et de transit, calculés sur la base simplifiée de l'évolution économique générale, permet de compléter le tableau des trafics réalisés sur l'aire d'étude.
3. Les matrices origines-destinations établies par ces différents modules sont ensuite rassemblées par mode et période horaire pour lancer les calculs de l'affectation. Ceux-ci permettent de déterminer les temps généralisés entre toutes les zones de l'aire urbaine de Lyon.
4. Enfin, ces temps généralisés permettent d'établir des indicateurs d'accessibilité qui sont utilisés dans les modules de localisation des ménages et des activités.

COMMENT PRENDRE EN COMPTE LA LOCALISATION RESIDENTIELLE DES MENAGES ?

Nous l'avons dit, ce cahier traite plus particulièrement des choix de localisation des ménages, tels qu'ils ont été envisagés dans SIMBAD.

La modélisation traditionnelle du système de transport s'appuie généralement sur une structure en quatre étapes (la génération des déplacements, leur répartition entre les zones du territoire, le choix de leur mode de transport, leur affectation sur les réseaux routier ou de transport collectifs) qui prend en compte l'impact des localisations des ménages et des activités sur les transports. Par contre, les choix de localisation sont considérés comme exogènes et les rétroactions du système de transports sur la structuration du territoire ne sont donc pas prises en compte. Cette question, qui constitue autant un enjeu pratique pour les aménageurs qu'un enjeu de représentation pour les chercheurs, a donc constitué une accroche forte entre l'Agence d'Urbanisme et le Laboratoire d'Economie des Transports. Nous nous concentrons ici uniquement sur la localisation résidentielle des ménages au sein du territoire urbain, sachant que l'autre composante de la localisation – celle des activités économiques –, a fait l'objet d'une modélisation parallèle au sein de SIMBAD.

L'objectif du modèle de choix résidentiels est d'expliquer et de prévoir les choix des ménages en fonction de l'évolution des facteurs explicatifs de leur comportement. Ce type de modèles vise à donner la probabilité qu'un ménage change de résidence puis la probabilité de choix de la nouvelle localisation. Il peut ensuite être appliqué pour simuler l'évolution du système urbain sous diverses hypothèses d'évolution du contexte économique et démographique général et pour tester les effets de différentes politiques de transport et d'urbanisme.

La microéconomie urbaine comme cadre théorique

La théorie de la microéconomie urbaine a servi de cadre d'analyse pour l'étude du comportement des agents économiques dans la ville. Elle repose sur le principe d'agents économiques rationnels cherchant à maximiser leur utilité compte tenu de leur revenu, des caractéristiques de leur zone de résidence (coût du logement, cadre de vie, etc.) et des coûts de transport pour atteindre les autres zones et les services qu'elles offrent (marché de l'emploi notamment). Même si les hypothèses comportementales posées par cette théorie sont relativement simples, les modèles qui en découlent fournissent des éclairages intéressants pour la compréhension des comportements des ménages, des déterminants de la localisation résidentielle ou de l'impact du transport sur les formes urbaines. Les apports de la nouvelle économie urbaine ont été décisifs en la matière, avec des auteurs comme Alonso, Mills ou Muth qui ont écrits durant les années 60-70. Ils décrivent une ville de type américaine, où les emplois se concentrent en centre-ville et les habitants se répartissent en cercles concentriques, les plus pauvres proches du centre pour diminuer leurs coûts de déplacements et les plus riches plus éloignés pour bénéficier de l'espace et du cadre de vie. Des travaux plus récents comme ceux de Fujita ou Thisse dans les années 80 et 90 ont complexifié ce premier modèle en prenant en compte le développement des emplois, des centres commerciaux ou d'espaces de loisirs, en périphérie ou dans des centre secondaires.

Par ailleurs, la technique des modèles de choix discrets fondée sur la théorie de l'utilité aléatoire développée par MacFadden, prix Nobel d'économie, a été privilégiée pour un certain nombre de raisons. D'abord, s'agissant d'une technique de type économétrique désagrégée, elle est bien adaptée pour répondre à notre objectif de modéliser les comportements résidentiels au niveau désagrégé du ménage, en prenant en compte les nombreux facteurs explicatifs de ses choix. Ensuite, il s'agit d'une approche qui reste dans le cadre théorique

cohérent de la microéconomie urbaine tout en introduisant la souplesse de la théorie de l'utilité aléatoire, qui a l'avantage d'éliminer certaines restrictions de l'approche déterministe en permettant de concevoir des variations de comportements qui peuvent être interprétées autant comme un relâchement de l'hypothèse de rationalité des agents, que comme la prise en compte d'une connaissance imparfaite, soit des éléments du choix pour les ménages, soit des facteurs de ce choix pour le modélisateur.

Une modélisation en deux étapes : décision de mobilité résidentielle puis choix de localisation

La décision des ménages en matière résidentielle est un processus unique qui intègre deux composantes, le choix de déménagement et le choix d'une nouvelle localisation, qui sont dépendantes l'une de l'autre. Il serait donc théoriquement plus réaliste et plus juste de modéliser les décisions en matière résidentielle des ménages (mobilité + localisation) par un seul modèle.

Pour procéder de la sorte, il faut avoir une base désagrégée des ménages, c'est-à-dire avec une description des caractéristiques de chaque ménage, y compris leur localisation précise. Or, on ne dispose pas d'une telle source de données au niveau de l'aire urbaine lyonnaise. L'Enquête Logement, qui est la principale source de données décrivant les comportements résidentiels des ménages au niveau national, fournit des informations sur un échantillon de ménages à un moment donné, y compris leurs caractéristiques au moment du déménagement, avec une périodicité de 4-5 ans mais sans maintenir le même échantillon de référence. Cependant, cette enquête ne fournit pas la localisation précise de chaque ménage. La seule information disponible est le département de résidence, découpage spatial insuffisant par rapport à notre tentative de modélisation.

Cette limite dans la disponibilité des données n'est pas propre au cas lyonnais, et les deux facettes de la décision résidentielle sont le plus souvent traitées séparément dans les travaux sur ce sujet. Nous avons ainsi considéré les choix de mobilité et de localisation des ménages comme deux étapes successives et indépendantes du processus de décision résidentielle et ils ont été modélisés en conséquence.

La modélisation des choix discrets

La théorie de l'utilité aléatoire qui a été privilégiée dans notre approche aboutit à une formalisation particulière qui mérite d'être décrite car elle conditionne la mise en forme des données utilisées et les résultats obtenus.

Les modèles basés sur cette théorie ont pour objectif d'estimer un choix à partir de la connaissance d'un ensemble d'alternatives mutuellement exclusives et collectivement exhaustives. La modélisation des choix discrets commence par l'association, à chaque alternative constituant l'ensemble étudié, d'une utilité qui dépend des caractéristiques des individus et des alternatives. On déduit après, pour chaque individu, la probabilité de choisir une alternative, en tenant compte du fait qu'il vise à maximiser son utilité. Cette utilité a deux composantes :

$$U_{in} = V_{in} + e_{in} \quad \text{où : } U_{in} - \text{utilité de l'alternative } i \text{ pour l'individu } n$$

$$V_{in} - \text{partie observable, déterministe de l'utilité}$$

$$e_{in} - \text{terme aléatoire}$$

La partie observable est conçue comme une fonction des attributs de l'alternative i et des caractéristiques de l'individu n :

$$V_{in} = \theta_1 Z_{in1} + \theta_2 Z_{in2} + \dots + \theta_t Z_{int} \quad \text{où : } Z_{in\tau} - \text{attributs de l'alternative } i \text{ ou de l'individu } n$$

$$\theta_\tau - \text{paramètre de la variable } Z_\tau$$

La probabilité que l'alternative i soit choisie par l'individu n est donnée par :

$$P_n(i) = P(U_{in} > U_{jn}), j \neq i \quad \text{où : } P - \text{probabilité}$$

$$j - \text{alternative } j$$

L'équation précédente est utilisée pour calculer la probabilité que l'individu n réalise le choix de l'alternative i sous l'hypothèse qu'on peut estimer la distribution du terme aléatoire e_{in} .

Un problème central est la spécification de V_{in} , c'est-à-dire la détermination des variables qui doivent entrer dans la fonction d'utilité ainsi que la forme fonctionnelle la plus appropriée. Dans la plupart des cas, on utilise des fonctions linéaires.

Ensuite, se pose la question de la distribution statistique du terme aléatoire e_{in} . L'hypothèse la plus souvent utilisée est que les composantes aléatoires des fonctions d'utilité sont indépendamment et identiquement distribuées suivant une loi doublement exponentielle (dite loi de Gumbel). Une variable aléatoire e suit une distribution de Gumbel de paramètres μ et η si celle-ci s'exprime par la formule :

$$F(\varepsilon) = \exp(-e^{-\mu(\varepsilon-\eta)}), \mu > 0 \text{ et la loi de densité de Gumbel est : } f(\varepsilon) = \mu e^{-\mu(\varepsilon-\eta)} \exp(-e^{-\mu(\varepsilon-\eta)})$$

On fait aussi l'hypothèse que la structure de la matrice de variance-covariance du terme d'erreur entre les alternatives est identique pour tous les individus. Cette hypothèse traduit le fait que la variation due aux facteurs non observés est identique entre les individus.

Cette distribution du terme d'erreur est fréquemment utilisée car d'une part elle est proche de la loi normale et d'autre part elle conduit à une forme mathématique simple quel que soit le nombre d'alternatives:

$$P_n(i) = e^{\mu V_{in}} / \sum_j e^{\mu V_{jn}}$$

C'est avec ce type de modèle qu'a été déterminée la probabilité pour un ménage de se localiser dans telle ou telle zone s'il a choisi de déménager.

Par ailleurs, le nombre d'alternatives possibles pour se localiser est extrêmement important puisqu'il correspond au nombre de zones ou de cellules du périmètre modélisé. Dans la pratique, la puissance de calcul nécessaire pour estimer un logit multinomial devient rapidement énorme et hors de portée. Pour dépasser cette limite, on peut réduire l'univers de choix des localisations en réalisant au préalable un tirage aléatoire d'un nombre réduit d'alternatives parmi toutes celles possibles. On s'éloigne ainsi des standards d'estimation du modèle logit, mais on peut montrer que les résultats obtenus sont convergents avec ceux de la méthode académique (Ben-Akiva & Lerman, 1985).

C'est la procédure que nous avons suivie à l'aide du modèle URBANSIM pour établir le choix de chaque individu statistique. Ainsi, chaque ménage qui doit être localisé est modélisé individuellement, un échantillon de cellules alternatives étant généré pour chacun des ménages. Après plusieurs essais de nombres de tirages différents, nous avons constaté une convergence très rapide des résultats, et nous nous sommes finalement arrêtés à 6 alternatives par ménage (Homocianu, 2009).

DÉMÉNAGER OU PAS ? UN MODÈLE DE MOBILITÉ RÉSIDENIELLE

Les données désagrégées au niveau du ménage de l'Enquête Logement 2002 ont servi à caler ce premier modèle. Les ménages de l'aire urbaine de Lyon ont été identifiés et on a pu estimer les coefficients des variables de la fonction d'utilité qui fournit les probabilités de déménager de chaque type de ménages.

La mobilité résidentielle décroît avec l'âge. A côté du nombre d'enfants du ménage, la variable âge de la personne de référence doit être considérée pour décrire les étapes du cycle de vie des ménages. Aux différentes étapes du cycle de vie d'un ménage correspond une trajectoire résidentielle particulière, avec des moments, dans certaines étapes, où a lieu un changement de résidence. Par rapport aux différentes tranches d'âge, qui correspondent aux différentes étapes du cycle de vie, on constate que les déménagements sont plus fréquents parmi les ménages plus jeunes, avec une tendance vers l'immobilité avec le vieillissement. L'âge de la jeunesse correspond à une étape de la vie dans laquelle, en général, le ménage n'a pas d'enfants, a moins de chances d'être propriétaire d'un logement, la mobilité professionnelle de ses membres est plus élevée, ce qui fait qu'il adopte un comportement résidentiel plus dynamique. L'avancement dans le cycle de vie tend à un renforcement progressif de ses diverses contraintes, qui fixent le ménage sur un lieu. Une fois ces variables contrôlées, il reste toujours une tendance à rester moins mobile chez les personnes âgées, que nous attribuerons simplement au facteur « âge ».

Les locataires déménagent plus que les propriétaires. Parmi les ménages plus mobiles, la part des ménages locataires est nettement plus importante que celle des ménages qui sont propriétaires de leur logement. C'est également une tendance qu'on retrouve dans l'étude bibliographique, qui s'explique notamment par le fait que l'une des raisons principales pour lesquelles un ménage change de logement est le désir de passer du statut de locataire de son ancien logement au statut de propriétaire du nouveau logement.

La mobilité résidentielle du ménage se réduit avec le nombre d'enfants. En comparant le profil du nombre d'enfants des ménages qui changent leur résidence avec celui des ménages qui ne bougent pas, on constate qu'à un ménage avec moins d'enfants correspond une mobilité résidentielle plus élevée, constat qui est en accord avec ce que préconise la littérature de la mobilité résidentielle. En effet, il est plus facile pour un jeune ménage sans enfant de changer de logement. Avec l'arrivée des enfants, le ménage a tendance à se stabiliser.

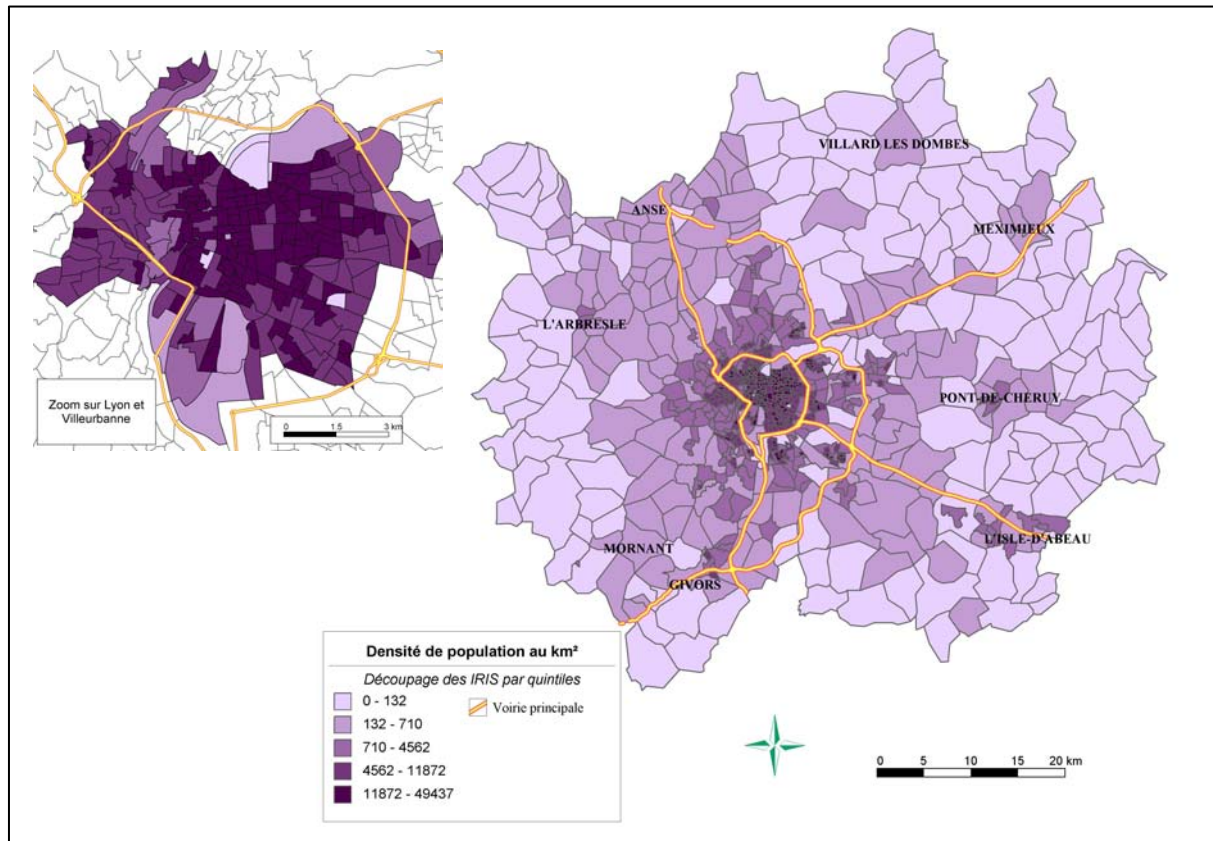
La mise en correspondance entre ce qui est observé et les prédictions du modèle montre que le pourcentage de ménages dont le choix a été correctement prévu est de 87%. Il y a cependant ici risque de confusion car 88,6% des ménages gardent leur lieu de résidence d'une année sur l'autre : un modèle ne prévoyant aucun déménagement obtiendrait de meilleurs résultats que le nôtre... Il est plus intéressant de noter que la prévision de notre modèle est correcte pour 41,2% des ménages qui changent de résidence dans l'année.

Ce sont des changements d'états qui sont principalement la cause d'un déménagement (changement d'emploi, naissance d'un enfant, mise en couple, etc.). Nos variables explicatives, statiques, sont plus indicatives de la probabilité de changement d'état que du changement d'état lui-même. Pour améliorer les résultats, il serait nécessaire de dynamiser l'approche avec, par exemple, un module de suivi démographique pour représenter le cycle de vie des ménages et des événements qui l'accompagnent. En l'état, avec les données statiques disponibles, prévoir correctement plus de 40% des ménages qui déménagent pour une année donnée alors que la moyenne globale est de 11,4% apparaît très satisfaisant et révèle une bonne prise en compte de la structure des variables explicatives sous-jacentes.

OÙ SE LOCALISER ? UN MODÈLE DE CHOIX DE L'IRIS DE RÉSIDENCE

Comme dans le cas du modèle de mobilité résidentielle, nous utilisons une formulation logit pour modéliser la localisation des ménages. Il s'agit cette fois du logit multinomial adapté, puisqu'on a un très grand nombre d'alternatives de choix (777 IRIS, qui ont été éclatés en 52 822 cellules de 250x250m dont 5 810 sont habitables et constituent des localisations possibles pour les ménages). Nous évoquons d'abord les variables utilisées avant de présenter la construction de ce modèle et les résultats obtenus.

Carte 2 : Localisation de la population dans l'aire urbaine de Lyon en 1999



Source : traitement LET, à partir du RGP Insee 1999

Les variables utilisées

Pour caler ce modèle, nous utilisons les données du RGP 1999, plus précisément une population synthétique des ménages dont on connaît, parmi d'autres caractéristiques, la localisation au niveau de l'IRIS. Le modèle est bâti uniquement à partir des ménages ayant déménagé dans l'aire urbaine de Lyon l'année précédent le recensement de 1999, soit 91 461 ménages sur les 662 249 de l'aire urbaine à l'époque. Grâce à l'INSEE, nous disposons d'une base exhaustive décrivant ces ménages à partir des variables suivantes :

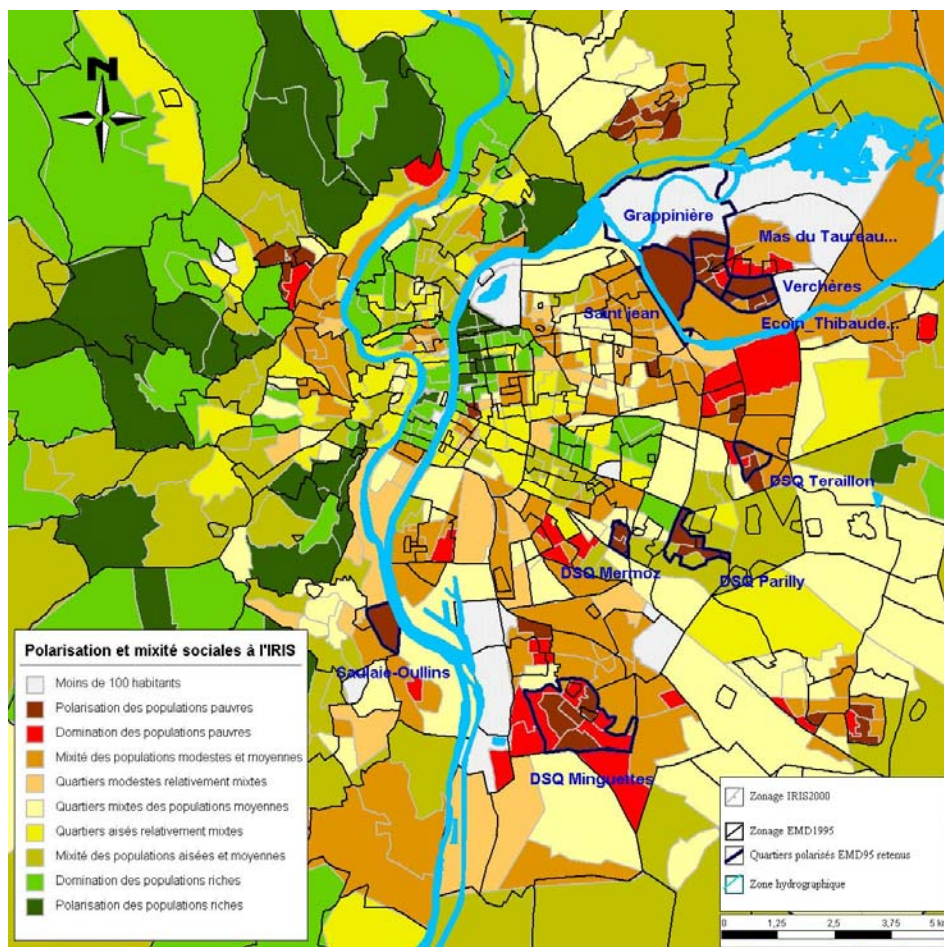
- ✓ IRIS de résidence au moment du recensement ;
- ✓ déménagement l'année précédent le recensement (oui/non);
- ✓ âge de la personne de référence (moins de 30 ans ; 30-44 ans ; 45-59 ans ; 60-74 ans ; 75 ans et plus) ;
- ✓ statut du chef de ménage (travailleur à temps plein ; travailleur à temps partiel ; chômeur ; retraité ; autres inactifs - étudiants et personnes au foyer notamment) ;

- ✓ nombre de personnes du ménage (0, 1, 2, 3, 4, 5 ou plus) ;
- ✓ nombre d'actifs dans le ménage (0, 1, 2 ou plus) ;
- ✓ statut d'occupation du logement (propriétaire ou locataire) ;
- ✓ nombre de voitures à disposition du ménage (0, 1, 2 ou plus).

Un revenu par unité de consommation en trois classes a par ailleurs été imputé pour déterminer si le ménage avait un revenu modeste (les 20% des ménages aux plus bas revenus), moyen (les 60% médians) ou élevé (les 20% aux plus hauts revenus). Pour réaliser cette imputation, nous avons eu recours aux données de l'Enquête Ménages Déplacements de Lyon 2006, au sein de laquelle les variables descriptives précédentes sont toutes disponibles et où 7 000 ménages ont indiqué leur revenu. Nous avons ainsi estimé un modèle explicatif en introduisant toutes les variables potentiellement explicatives du revenu, en essayant de ne retenir que les variables significatives du point de vue statistique. Trois variables ont finalement été retenues :

- ✓ le taux d'activité du ménage ;
(nb actifs/nb de personnes, en 5 classes : 0,]0 ; 0,33],]0,33 ; 0,66],]0,66 ; 1], 1) ;
- ✓ le nombre de VP du ménage ;
- ✓ le statut d'occupation du logement.

Carte 3 : Polarisation et mixité sociale des quartiers de l'aire urbaine de Lyon en 1999



*Typologie basée sur le pourcentage de la population du quartier IRIS dans chaque quintile de revenu.
Bouzouina, Nicolas, 2009, à partir des données INSEE-DGI, IGN*

En plus des caractéristiques des ménages, cinq grandes familles de variables explicatives qualifiant les zones ont été retenues et testées pour la modélisation des localisations. Les données INSEE de population (RGP) et d'établissements (fichier SIRENE) ont servi de base pour qualifier chaque IRIS, et les temps d'accès entre zones ont complété ces premières informations en permettant d'estimer des indicateurs d'accessibilité plus complexes :

- ✓ la proximité à la population a été considérée à travers la densité et le nombre de ménages de l'IRIS, en distinguant l'âge de la personne de référence et la classe de revenu du ménage pour mettre en évidence les logiques de ségrégation et d'entre-soi ;
- ✓ la proximité des emplois a été exprimée à travers les densités et les nombres d'établissements et d'emplois dans chaque IRIS ;
- ✓ l'accessibilité joue un rôle essentiel dans les localisations, a été traduite sous diverses formes. Des accessibilités isochroniques ont été établies, avec un nombre d'opportunités (population, emplois, réseaux de transport, etc.) contenues dans un temps donné à partir de l'IRIS considéré. Des accessibilités gravitaires ont également été calculées, avec une forme simple de type : $AO_i = \sum_{j=1}^{777} O_j / tps_{ij}$
 Avec AO_i = Accessibilité aux opportunités (population, emplois...) de l'IRIS i ;
 O_j = nombre d'opportunités contenues dans l'IRIS j ;
 tps_{ij} = temps d'accès entre les IRIS i et j
- ✓ le temps d'accès au centre, indicateur plus simple du modèle monocentrique, a été estimé par le temps d'accès à la préfecture en heure creuse et en heure de pointe ; la distance a également été testée, mais avec de moins bons résultats ;
- ✓ le prix de l'immobilier, enfin, a été estimé à partir des données fournies par l'Observatoire des Transactions Immobilières et Foncières du Grand Lyon.

Le modèle retenu : un fort impact des variables d'accessibilité

La qualité du modèle est satisfaisante, avec 62% de ménages correctement classés. D'autres tests ont été réalisés, en intégrant notamment la localisation ou le statut du chef de ménage, sans obtenir de meilleure estimation. Dès lors, les probabilités d'affectation à l'une des trois classes de la variable explicative du revenu par unité de consommation ont été calculées pour chaque type de ménage (en fonction des différents niveaux des trois variables explicatives on a $2*3*5=30$ types de ménages). Elles ont ensuite été utilisées pour affecter à chaque ménage du recensement 1999 une des trois classes de revenu.

Les variables caractéristiques des ménages sont ensuite introduites dans le modèle en construisant des effets croisés avec des caractéristiques des IRIS. Par exemple, les variables de profil du ménage comme l'âge de la personne de référence ou le revenu sont croisées avec des variables indiquant le pourcentage des ménages du même type dans l'IRIS concerné, afin d'évaluer le degré de mixité sociale ou, au contraire, de ségrégation dans le territoire de l'aire urbaine.

Dans le premier modèle retenu, tous les coefficients sont significatifs et le pseudo- R^2 de McFadden est de 0,40, ce qui indique une bonne qualité du modèle et un pouvoir explicatif important des variables introduites.

Les coefficients correspondants aux termes d'interaction entre les caractéristiques des ménages – âge de la personne de référence, revenu et nombre de personnes – et les pourcentages des ménages ayant les mêmes caractéristiques dans l'IRIS choisi montrent que, en général, les ménages ont tendance à se localiser dans une zone où il y a plus de ménages du même profil. Les ménages à bas revenu et les ménages de taille élevée (ayant des coefficients

négatifs) font exception. Ils se localisent plutôt dans des zones où il y a moins de ménages résidents ayant le même niveau de revenu et respectivement de taille (en fait, il s'agit de ménages dont le niveau de vie est plutôt réduit, et il n'y a aucun intérêt pour ces ménages à s'installer dans des zones où habitent majoritairement des ménages ayant le même profil).

Le signe positif (apparemment contre-intuitif, et qui est d'ailleurs très faible) du coefficient du prix immobilier au m² traduit une relation positive entre le niveau de cette variable et la probabilité d'un ménage de se localiser dans une zone caractérisée par ce niveau du prix. L'explication peut résider dans le fait que le prix capitalise les externalités de la zone, qui peuvent être déterminées par des caractéristiques non introduites dans le modèle. Nous trouvons ici un résultat similaire à celui du projet parisien SIMAURIF, qui a également estimé un coefficient positif pour le prix immobilier dans son modèle de localisation des ménages (Nguyen Luong, 2007).

Par rapport aux variables d'accessibilité, le fait que tous les coefficients soient significatifs montre que cette composante joue un rôle important dans les décisions de localisation résidentielle des ménages, et confirme l'hypothèse de l'existence d'un impact des transports sur l'urbanisation. Le signe positif des coefficients pour les variables accessibilité gravitaire aux emplois, accessibilité gravitaire aux établissements secondaires, accessibilité gravitaire aux grands commerces, nombre d'échangeurs accessibles en moins de 15 min et nombre de stations de métro à moins de 1000 m exprime le fait que l'utilité pour les ménages de se localiser dans un IRIS augmente quand l'accessibilité correspondante à ces opportunités est élevée. Les signes des coefficients pour les autres variables d'accessibilité (qui sont apparemment contre-intuitifs) sont à prendre avec précaution, étant donné que les corrélations qui existent entre les variables explicatives d'accessibilité rendent difficile l'interprétation de ces coefficients.

Pour traiter ces problèmes de corrélation, nous avons cherché à synthétiser l'information que ce groupe de variables apporte, en construisant un indicateur agrégé d'accessibilité. En utilisant l'analyse en composantes principales, on a obtenu une première composante qui synthétise presque toute l'information qu'on obtiendrait si on utilisait les neuf variables initiales (85% de la variance globale est expliquée par cette composante). Nous retenons donc cette première composante comme indicateur synthétique d'accessibilité :

$$\begin{aligned} \text{Accessibilité agrégée} = & -0,119 * \text{temps d'accès au centre en heure de pointe} \\ & + 0,124 * \text{nombre d'échangeurs autoroutiers à moins de 15 mn} \\ & + 0,112 * \text{nombre d'arrêts de bus à moins de 1000 mètres} \\ & + 0,092 * \text{nombre de stations de métro à moins de 1000 mètres} \\ & + 0,117 * \text{nombre de gares à moins de 30 minutes} \\ & + 0,129 * \text{accessibilité gravitaire aux grands commerces} \\ & + 0,129 * \text{accessibilité gravitaire aux établissements secondaires} \\ & + 0,129 * \text{accessibilité gravitaire à l'emploi} \\ & + 0,130 * \text{accessibilité gravitaire aux emplois du secteur tertiaire} \end{aligned}$$

Cet indicateur synthétique a été calculé pour chaque zone. Une valeur élevée de l'indicateur signifie une forte accessibilité aux opportunités. Nous estimons donc de nouveau le modèle à l'aide cette variable en ne conservant que les variables dont la significativité (p-value) est supérieure à 5% (Tableau 1) :

Tableau 1 : Coefficients estimés du modèle de choix de localisation résidentielle à la cellule

Variable	Coefficient	t de Student	p-value
<i>indicateur d'accessibilité agrégé</i>	0.98	141.3	0.000
<i>prix moyen au m² de l'immobilier collectif ancien</i>	7 10 ⁻⁴	31.1	0.000
<i>Age_pr<30 x taux_men_age_pr<30</i> ^(a)	6.28	37.0	0.000
<i>Age_pr30-45 x taux_men_age_pr 30-45</i> ^(a)	1.33	15.5	0.000
<i>Age_pr>75 x taux_men_age_pr >75</i> ^(a)	3.45	5.0	0.000
<i>Men_revbas x taux_men_revbas</i> ^(a)	-2.863	-16.0	0.000
<i>Men_revhaut x taux_men_revhaut</i> ^(a)	5.19	26.6	0.000
<i>Men_taille1 x taux_men_taille1</i> ^(a)	2.85	32.2	0.000
<i>Men_taille2 x taux_men_taille2</i> ^(a)	6.71	23.6	0.000
<i>Men_taille>5 x taux_men_taille>5</i> ^(a)	-7.24	-15.9	0.000

^(a) Toutes les variables ménages sont multipliées par une variable indicatrice qui prend la valeur 1 lorsque le ménage possède les caractéristiques de la variable retenue et 0 dans les autres cas.

Source : Homocianu, 2009

L'introduction dans le modèle de l'indicateur synthétique d'accessibilité à la place des variables considérées antérieurement entraîne une légère baisse du pseudo-R², qui est maintenant de 0,33.

Les constats sont les mêmes que dans le cas précédent. En revanche, on peut maintenant voir clairement la relation positive entre la composante globale d'accessibilité pour une zone et l'utilité retirée par les ménages de la localisation dans cette zone.

CONCLUSION : TESTER L'INFLUENCE DES POLITIQUES DE TRANSPORT ET D'URBANISME SUR LES LOCALISATIONS DES MENAGES

Les résultats de cette modélisation des comportements résidentiels vont permettre de tester l'impact de politiques de transport et d'urbanisme sur la répartition des ménages au sein du territoire de l'aire urbaine de Lyon. Deux grands types de questions peuvent être posés :

- ✓ Compte tenu des logiques mises en évidence, risque-t-on d'assister à des phénomènes de ségrégation urbaine (aisés / modestes ; jeunes, actifs / retraités) dommageable pour l'équilibre et le fonctionnement harmonieux des territoires lyonnais ? Des politiques jouant sur les accessibilités urbaines pourraient elles contrebalancer ces tendances ?
- ✓ Dans la logique générale du projet SIMBAD, quelles conséquences ces phénomènes de ségrégation pourraient avoir sur les déplacements au sein de l'aire urbaine ? Seraient-ils d'une ampleur suffisante pour affecter de manière sensible les nuisances et le coût économique du système de transports ? Risqueraient-ils de renforcer les inégalités entre ménages en détériorant l'accès aux services urbains et/ou en renforçant le poids de la mobilité quotidienne dans le budget des ménages les plus modestes ?

Quelques grands enseignements peuvent cependant déjà être tirés de ce travail, qu'il convient en même temps de mettre en regard des limites propres à tout exercice de modélisation et des perspectives de recherche sur lesquels ils débouchent.

Les principaux enseignements sur les logiques de localisation résidentielle

Le modèle de mobilité résidentielle était censé répondre à la question : quels sont les ménages qui vont prendre la décision de changer de logement, en fonction de leur profil socio-démographique ? Les variables qui expliquent la variation de leur taux de déménagement sont l'âge de la personne de référence du ménage, le nombre d'enfants et le statut d'occupation du logement. Le modèle estimé a un pouvoir explicatif satisfaisant. Il montre une propension à déménager plus élevée pour les ménages jeunes, de taille réduite et locataires, tandis que les ménages plus âgés, de taille plus élevée et qui sont propriétaires de leur logement sont moins mobiles.

En ce qui concerne la localisation des ménages, on constate une convergence des résultats des modèles estimés par rapport aux conclusions que l'on peut tirer sur l'impact de la composante accessibilité sur la localisation des ménages et donc sur l'urbanisation. Les préférences des ménages pour une zone ou une autre sont liées aux caractéristiques de celles-ci et notamment à l'accessibilité aux différentes opportunités et services, ce qui confirme que parmi les facteurs qui influencent le comportement de localisation des ménages on retrouve les accessibilités. L'hypothèse de l'existence d'un lien entre transports et urbanisme est donc bien confirmée. Ainsi, on constate que les ménages trouvent plus utile de se localiser dans une zone caractérisée par un niveau global d'accessibilité plus élevé, et l'utilité diminue au fur et à mesure que le niveau d'accessibilité de la zone baisse.

On a également montré que les caractéristiques des ménages comme l'âge de la personne de référence, le revenu ou le nombre de personnes ont une influence sur leurs choix de localisation. De même, les ménages ont tendance à se localiser dans une zone où il y a plus de ménages du même profil d'âge, de revenu ou de taille (sauf les ménages à bas revenu et les ménages de taille élevée, qui se localisent plutôt dans des zones où il y a moins de ménages résidents ayant le même niveau de revenu et respectivement de taille).

Quelles pistes d'amélioration peuvent être proposées ?

Trois grandes pistes d'amélioration pourront être explorées à plus ou moins long terme pour améliorer les résultats obtenus. La première concerne l'affinement de la modélisation actuelle, en renforçant la qualité des données utilisées et en explorant les améliorations qui pourraient être apportées aux deux modèles logit utilisés. La seconde touche à la modélisation de l'occupation des sols, qui repose pour l'instant sur des hypothèses externes. Enfin, une troisième piste, à plus long terme, concerne la modélisation du parcours de vie des ménages.

Affiner le modèle actuel. Le développement de SIMBAD se réalise progressivement et se poursuit encore. Le calage fin des modules d'affectation du trafic routier et des déplacements en transports collectifs vient d'être terminé. Il va permettre d'affiner les mesures d'accessibilité pour tester leur apport explicatif dans le modèle. Des indicateurs d'accessibilité gravitaire plus directement cohérents avec la théorie économique pourront aussi être testés. La méthode d'estimation des revenus des ménages peut aussi apporter une amélioration des résultats. Une prise en compte plus fine de la répartition géographique de ces revenus au sein de l'aire urbaine est ainsi en cours de réalisation, sur la base des travaux de thèse de Louafi Bouzouina (2008). Enfin, en parallèle de ces améliorations de la qualité des données utilisée, des modélisations alternatives sont également envisagées. Compte tenu des résultats déjà obtenus, une modélisation hiérarchique va être testée, dans laquelle un premier choix porterait sur le type d'IRIS en fonction de sa densité, puis un deuxième se ferait à niveau fin, sur la cellule au sein d'un type d'IRIS.

L'ensemble de ces améliorations sera réalisé et testé d'ici la fin de l'année 2009.

Concevoir un modèle de développement urbain. Une seconde piste concerne la détermination de l'univers des choix résidentiels. En effet, l'évolution des espaces disponibles dédiés au logement ou aux activités n'est pas modélisée. Le phénomène d'étalement urbain ou une éventuelle reconcentration liée à une politique volontariste en matière d'urbanisme et de transport sont pris en compte sur la base de scénarios prospectifs, avec des hypothèses posées sur les espaces disponibles dans chaque zone. Ainsi, le modèle de localisation mis en oeuvre ne fournit des indications que sur la composition de la population des zones.

Dans le cadre du projet SIMBAD 2 retenu par le PREDIT, le Laboratoire d'Economie des Transports et l'Agence d'Urbanisme travailleront sur la conception d'un tel modèle de développement urbain dans le courant des années 2010 et 2011.

Prendre en compte le parcours de vie des ménages. La troisième piste concerne la dynamique démographique des ménages. Pour l'instant, nous utilisons les projections démographiques réalisées par l'INSEE (modèle OMPHALE) sur l'aire urbaine de Lyon et notre modélisation s'assure année par année que le nombre total de ménages par classes d'âge de la personne de référence et par classe de tailles est bien respecté. Si une classe devient trop nombreuse par rapport à la prospective INSEE, les ménages en trop sont supprimés par tirage aléatoire ; au contraire, s'il en manque, de nouveaux sont créés. Ainsi, les caractéristiques de la population restent bien respectées à un niveau agrégé. Par contre, les caractéristiques d'un ménage (taille, âge de la personne de référence, revenu, etc.) sont données une fois pour toute. Le parcours de vie de chaque ménage n'est pas représenté dans notre modèle. Ses membres ne vieillissent pas, il n'y a pas de naissance, de décès ou de séparation, alors que ces éléments constituent des facteurs essentiels de choix de déménagement et de localisation. Nous envisageons donc à plus long terme de développer une approche qui prendrait en compte le cycle de vie de chaque ménage, ce qui permettrait de revoir le modèle de localisation sur une base renouvelée.

ELÉMENTS DE BIBLIOGRAPHIE

Homocianu M., 2009, *Modélisation de l'interaction transport-urbanisme – choix résidentiels des ménages dans l'aire urbaine de Lyon*. Lyon, Université de Lyon, thèse pour le doctorat en sciences économiques, 311 p.

Nicolas J.-P. (dir), Bonnel P. (dir), Cabrera J., Godinot C., Homocianu M., Routhier J.-P., Toilier F., Zuccarello P., 2009, *Simuler les mobilités pour une agglomération durable*. LET, projet Simbad, rapport final pour le compte de l'ADEME et de la DRAST. 173 p.

Bouzouina L., 2008, *Ségrégation spatiale et dynamiques métropolitaines*, Thèse de doctorat de sciences économiques, Université Lumière Lyon 2, 327 p.

Fujita M., Thisse J.-F., 2003, *Économie des villes et de la localisation*. Bruxelles, De Boeck Université, 560 p.

McFadden D., 1978, “Modelling the choice of residential location”. in Karlqvist A, Lundqvist L, Snickars F, Weibull J (eds), *Spatial Interaction Theory and Planning Models*, North Holland, Amsterdam, pp. 75-96.

Nguyen-Luong D., Motamedi K., Ouaras H., Picard N., Fernandes M., 2007, *SIMAUURIF : Modèle dynamique de SIMulation de l'interAction Urbanisation-transports en Région Ile-de-France*. Rapport final IAURIF/Théma pour le compte du PREDIT. 142 p.

Waddell P., 2002, “URBANSIM : Modeling Urban Development for Land Use, Transportation and Environmental Planning”. *Journal of the American Planning Association*, Vol. 68, No. 3, pp. 297–314.

Lien internet valide en mars 2009: <http://www.urbansim.org/papers/UrbanSim-JAPA.pdf>