



# ULTISIM

## Vers un modèle intégré transport-urbanisme européen.

### Première phase

## Rapport final



INSTITUT  
D'AMÉNAGEMENT  
ET D'URBANISME  
ÎLE-DE-FRANCE

en collaboration avec



**significance**  
quantitative research

Juin 2011



5.09.022 (numéro IAU)

09 MT CV 10 (numéro PREDIT)

## **ULTISIM : Vers un modèle intégré transport-urbanisme européen. Première phase**

### **Rapport final**

Institut d'Aménagement et d'Urbanisme de la Région d'Île-de-France (IAU îdF)

15, rue Falguière 75740 Paris cedex 15

Tél. 01.77.49.77.49 - Télécopie 01.77.49.77.69

E-mail : [dgcomm@iau-idf.fr](mailto:dgcomm@iau-idf.fr) –http : [www.iau-idf.fr](http://www.iau-idf.fr)

Directeur Général : **François DUGENY**

Directeur du Département Mobilité et Transports : **Alain MEYERE**

Rapport réalisé par :

**Dany NGUYEN-LUONG** (IAU îdF)

**Luca CANEPARO et Alfonso MONTURI** (Laq-TIP/Polytechnico di Torino)

**Michiel DE BOK et Eric KROES** (Significance),

**Denis VERRIER** (IAU îdF)

© IAU îdF juin 2011



## **Remerciements**

*Cette recherche a été financée par le MEDDTL dans le cadre du PREDIT 4 (GO6). Nous remercions Yves Crozet, Professeur d'Université et Président du GO6, et Gérard Brun, Secrétaire général du GO6.*

*Les auteurs remercient également le STIF (Laurence Debrincat, Olivier Nalin, Nicolas Pauget et Thierry Siméon) pour la subvention complémentaire et les informations sur le modèle ANTONIN 2.*



# SOMMAIRE

<b>1. Introduction .....</b>	<b>9</b>
<b>2. Etat de la pratique .....</b>	<b>11</b>
2.1. Généralités sur la modélisation LUTI .....	11
2.2. Origine d'ULTISIM .....	13
2.3. Logiciels de modélisation LUTI .....	15
2.3.1. Cube Land .....	16
2.3.2. DELTA .....	18
2.3.3. MARS .....	19
2.3.4. OPUS/UrbanSim .....	20
2.3.5. PECAS .....	23
2.3.6. Tranus .....	24
2.3.7. Autres logiciels .....	25
<b>3. Critères de développement d'ULTISIM .....</b>	<b>28</b>
3.1. Une absence de logiciel LUTI destiné au milieu professionnel .....	28
3.2. Modèle d'équilibre général versus modèle dynamique .....	29
3.3. Microsimulation versus simulation agrégée .....	30
3.4. Zonage .....	31
3.5. Structure de données .....	32
3.6. Modèle de demande de déplacements .....	32
3.7. Réseaux et affectation .....	32
3.8. Modes à prendre en compte .....	33
3.9. Ergonomie du logiciel .....	33
3.10. Positionnement d'ULTISIM .....	35
<b>4. Méthodologie de développement .....</b>	<b>37</b>
4.1. Outils de développement rapide .....	37
4.2. Ateliers de Génie Logiciel .....	38
4.3. L'off-shore .....	38
4.4. La méthode Agile .....	39
4.5. Open source versus logiciel propriétaire .....	39
<b>5. Données nécessaires .....</b>	<b>44</b>
5.1. La question des données .....	44
5.2. Eléments de réponse .....	46
5.3. Comparaison du cas Île-de-France avec deux cas étrangers .....	48
5.3.1. Pratiques de modélisation .....	48
5.3.2. Chiffres clés des aires d'étude .....	49
5.3.3. Tableau synthétique de comparaison .....	51
5.4. Données européennes .....	53
5.5. Nouvelles sources de données .....	54
5.5.1. Mouvement « Opendata » .....	54
5.5.2. Projet Belgrand-GEBD .....	56
5.5.3. Projet CASD (Centre d'Accès Sécurisé Distant) .....	56
5.5.4. Données issues de capteurs GPS et GSM .....	58
5.5.5. Mouvement « crowdsourcing » .....	60
<b>6. Spécifications fonctionnelles d'ULTISIM .....</b>	<b>62</b>
6.1. Degré de formalisme .....	62
6.2. Structure modulaire .....	63

6.3. Prise en compte des politiques urbaines et environnementales .....	65
6.4. Simulation dynamique .....	66
6.5. Accessibilité .....	66
6.6. Le découpage géographique .....	68
6.7. Le modèle de développement urbain .....	69
6.7.1. Objectif .....	69
6.7.2. Spécifications fonctionnelles .....	69
6.7.3. Base de données d'historiques .....	70
6.8. Interface entre ULTISIM et ANTONIN .....	71
6.8.1. Introduction .....	71
6.8.2. Description des données de réseaux du modèle ANTONIN .....	71
6.8.3. Description des modèles et des étapes de calcul d'ANTONIN .....	73
6.8.4. Passage de matrices de boucles journalières à des matrices de déplacement ...	75
6.8.5. L'affectation des déplacements sur les réseaux .....	76
6.8.6. Les résultats produits par le modèle ANTONIN .....	76
6.8.7. Echanges entrées-sorties entre ULTISIM et ANTONIN .....	77
<b>7. Conclusions, orientations, recommandations .....</b>	<b>79</b>
<b>Bibliographie .....</b>	<b>81</b>



# 1. Introduction

**ULTISIM** signifie « **Urban Land Transport Integrated SIMulation Model** ». La recherche vise à mettre au point un logiciel de simulation de l'interaction occupation du sol-transport répondant au cahier des charges suivant :

- un logiciel adapté aux conditions de planification européenne ;
- un logiciel générique, c'est-à-dire comportant une structure de données universelle et des modèles paramétrables, donc transposable d'une aire d'étude à une autre ;
- un logiciel complètement intégré et non un couplage *a posteriori* d'un logiciel de transport et d'un logiciel d'urbanisation ;
- un logiciel orienté « utilisateur final », adapté à une utilisation au quotidien.

La recherche est réalisée par l'IAU île-de-France en partenariat avec le Laboratoire de recherche italien Laq-TIP de l'Ecole Polytechnique de Turin et l'Institut de recherche privé néerlandais Significance. Elle bénéficie d'une subvention complémentaire du STIF dans le cadre du développement de son modèle de trafic ANTONIN.

Ce logiciel se positionne sur un marché encore balbutiant voire quasiment inoccupé, celui des **logiciels LUTI (Land Use Transport Interaction)** destiné au milieu professionnel : bureaux d'études, agences de transport et d'urbanisme, donneurs d'ordres (maîtres d'ouvrage et assistants à la maîtrise d'ouvrage), services techniques de l'Etat, services de transport et d'urbanisme des villes et des collectivités territoriales, gestionnaires, exploitants et autres autorités organisatrices des transports.

La recherche est prévue en deux phases, une phase de conception et une phase de développement. La présente recherche concerne la **première phase du projet ULTISIM**, d'une durée de 20 mois.

Conformément à la proposition de recherche retenue par le Predit en juillet 2009, cette première phase doit permettre une réflexion approfondie sur les données d'entrée ainsi que sur la méthodologie de développement, et **aboutir à une conclusion la plus « opérationnelle »** possible sur un futur modèle LUTI en Île-de-France. Elle comprend trois parties :

- une description actualisée de l'état de la pratique des modèles LUTI dans le monde ;

- une réflexion sur les données nécessaires aux modèles LUTI, comportant en particulier une comparaison des données disponibles de trois cas (Île-de-France, Aire urbaine de Turin en Italie et Randstad aux Pays-Bas) ;
- la définition des spécifications fonctionnelles du logiciel ULTISIM, avec en particulier des spécifications sur la structure de données, sur l'architecture du modèle, sur le lien avec le modèle ANTONIN du STIF et une proposition de méthodologie de développement du logiciel.

Un site internet bilingue a été mis en place pour favoriser le travail collaboratif entre les trois partenaires, faire connaître le projet et diffuser les résultats : [www.iau-idf.fr/ultisim](http://www.iau-idf.fr/ultisim)

## 2. Etat de la pratique

### 2.1. Généralités sur la modélisation LUTI

La recherche et le développement des modèles intégrés transport-urbanisme (ou transport-urbanisation ou transport-occupation du sol), plus connus sous le terme anglo-saxon de modèles LUTI (land-use transport interaction models) bénéficient d'un net **regain d'intérêt** depuis quelques années en France sous l'impulsion du programme Predit. Cette évolution de la recherche s'explique par la reconnaissance des **effets élargis** des transports, au-delà des gains de temps et gains environnementaux monétarisés. Ainsi quelques projets très prometteurs ont émergé de ce programme : *SIMAU*RIF en Île-de-France, SIMBAD à Lyon, MIRO, Mobisim. Dans le secteur privé, le groupe Vinci a développé de son côté le modèle Pirandello qui fait beaucoup parler de lui depuis trois ans et relance la question essentielle de l'**opérationnalité** des modèles LUTI dans le milieu hors académique.

Rappelons brièvement le principe de fonctionnement d'un modèle intégré. Les modèles classiques de prévision de trafic ne considèrent la relation transport-urbanisation que dans un sens. Les modèles intégrés (on parlera aussi de modèles interactifs ou d'interaction) transport-urbanisme visent à pallier à cette lacune en simulant les effets de la mise en service d'une nouvelle infrastructure de transport sur l'occupation du sol à moyen-long terme: effets de localisation résidentielle des ménages et des entreprises et effets sur les prix du foncier et de l'immobilier. Les anglo-saxons parlent de « wider benefits », c'est-à-dire des bénéfices économiques ou externalités positives au-delà des simples gains de temps obtenus après la mise en service de l'infrastructure. Par exemple, un modèle dynamique intégré tel que *SIMAU*RIF basé sur le logiciel OPUS/Urbansim associe un modèle de trafic classique (de type modèle agrégé à quatre étapes) et un modèle d'urbanisation (modèle « land use »). La répartition spatiale des populations et des emplois est endogénéisée<sup>1</sup>. A chaque pas de temps, cette nouvelle répartition génère des déplacements qui sont affectés sur le réseau multimodal de transport. Les coûts généralisés de transport et les mesures d'accessibilité sont, en retour, injectés dans l'expression de l'utilité de chaque modèle de choix discret. Les disponibilités foncières et immobilières et leurs prix sont modifiés à chaque pas de temps, ainsi que les réseaux de transport. Cette boucle de rétroaction annuelle

---

<sup>1</sup> Un des avantages des modèles LUTI est de laisser le système prévoir lui-même les projections locales de population et d'emplois à un horizon futur, alors qu'en modélisation classique des déplacements, ces données sont exogènes, la tâche étant réalisée par des démographes, des économistes et des urbanistes de manière plus ou moins quantitative, voire par les ingénieurs eux-mêmes lorsqu'ils ne bénéficient pas d'aide.

s'applique pendant 20 à 30 ans. La figure 1 schématise de manière simple le couplage et la structure rétroactive entre le modèle de trafic et le modèle d'urbanisation.

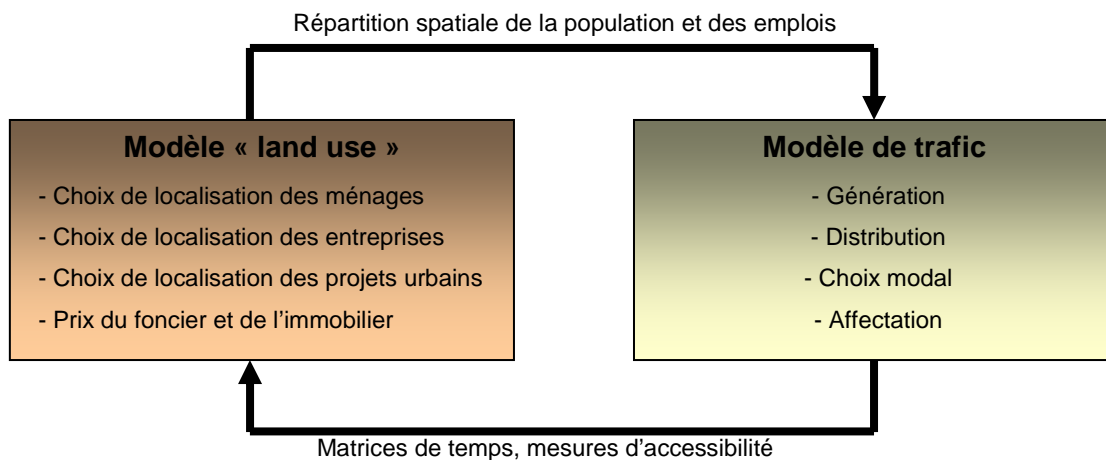


Figure 1 : schéma simplifié d'un modèle dynamique intégré

Les **questions typiques** auxquelles peuvent répondre théoriquement les modèles LUTI, contrairement aux modèles classiques de prévision de trafic, sont :

- Quelles sont les plus-values foncières et immobilières générées par une nouvelle infrastructure de transport en commun à proximité des gares ?
- Quels sont les effets sur les relocalisations des ménages et des emplois ?
- Quel est l'effet d'une politique de transport sur les émissions de CO<sub>2</sub> compte tenu de l'impact sur la répartition des populations et des activités ?
- Quel est l'impact sur les développements économiques régional et local d'un péage cordon autour du centre-ville ? Ce péage favorisera-t-il l'étalement urbain ou la densification ? Quels ménages se relocaliseront au centre-ville ?

Les **limites principales** des modèles LUTI sont :

- Ils « ne font que » de la répartition de la population et des emplois à partir de totaux exogènes sur l'aire d'étude fournis par un modèle macro-économique externe, mais ne modélisent pas les autres processus de la firmographie à savoir la création, la destruction et la croissance (De Bok, 2009). C'est un jeu à somme nulle par rapport à un scénario sans le projet de transport. La création nette d'emplois est introduite de manière exogène dans les hypothèses de totaux par année de simulation.
- Jusqu'à maintenant, les modèles LUTI dynamiques ne sont pas capables de prédire correctement l'impact d'une nouvelle infrastructure sur les prix de l'immobilier et du foncier. Les modèles de prix existants sont des modèles explicatifs et non prédictifs. Les modèles statiques d'équilibre général sont plus à même de simuler les prix.

- Ils prennent en compte les attributs d'agglomération (diversification, polarisation dans une zone donnée) mais ne modélisent pas les « effets d'agglomération »<sup>2</sup>, c'est-à-dire les effets économiques élargis calculés en termes de productivité des entreprises, de PIB supplémentaire et d'économie due à la polarisation (Charlot-Lafourcade, 2009). Les méthodes pour calculer ces effets sont toujours en débat. M. Jean Poulit en a proposé une, en partant de l'existence d'une corrélation entre suppléments de salaires ou de PIB et suppléments d'utilité liés aux destinations économiques commodément accessibles.

Il est donc admis que l'analyse classique coûts-avantages basée sur les gains de temps ne suffit plus à justifier un projet d'infrastructure de transport. Ainsi, la Société du Grand Paris (SGP) a lancé un marché d'études en avril 2011 dont l'objet est l'évaluation socio-économique du réseau de transport du Grand Paris. Ce marché comporte 7 lots, dont 3 concernent directement la **modélisation intégrée** :

- l'« Estimation des effets de la réalisation du réseau de transport du Grand Paris sur la localisation des entreprises et de l'emploi dans la région Île-de-France » (lot 2),
- le « Calcul des effets économiques élargis de la réalisation du réseau de transport du Grand Paris, appelés effets d'agglomération, en termes de productivité des entreprises et de PIB supplémentaire » (lot 3),
- le « Calcul de l'incidence globale de la réalisation du réseau de transport du Grand Paris sur la localisation des emplois et de la population, ainsi que sur les valeurs foncières compte tenu des disponibilités foncières actuelles et des perspectives dans l'utilisation des sols dans la région Île-de-France » (lot 4).

Il est apparu au cours des discussions que l'IAU<sup>3</sup> a pu avoir avec des bureaux d'études spécialisés en modélisation de trafic que tous étaient relativement **démunis** pour répondre aux trois lots dans le délai de réalisation de 6 mois, ne disposant pas d'outils adéquats.

## **2.2. Origine d'ULTISIM**

Une recherche bibliographique a été effectuée pour mieux connaître *l'état de la pratique* de la modélisation transport-urbanisation dans le monde. Nous ne parlons pas de *l'état de l'art* qui est du domaine de la recherche académique et qui s'attache essentiellement aux

<sup>2</sup> Ici le terme « agglomération » signifie le processus de concentration de constructions, d'activités ou de population sur un espace restreint, et le résultat de ce processus (Denise PUMAIN, dans Dictionnaire La Ville et l'urbain, Economica, 2006).

<sup>3</sup> L'IAU a été retenu pour travailler sur le lot 2, en partenariat avec le bureau MVA.

développements théoriques et aux estimations économétriques de modèles explicatifs non applicables en prédiction.

Au préalable, il est important de préciser ce que l'on entend par **état de la pratique**. Il est défini par l'utilisation dans le milieu professionnel de logiciels du commerce permettant de réaliser des simulations relativement facilement. Des connaissances de base en modélisation et en informatique devraient suffire. Ainsi, en modélisation de trafic classique, il y a aujourd'hui **quatre logiciels** dans le monde qui se partagent 90% du marché :

- Visem/Visum, de la société PTV (Allemagne)<sup>4</sup> ;
- Cube, de la société Citilabs (Etats-Unis) ;
- TransCAD, de la société Caliper Corporation (Etats-Unis) ;
- Emme 3, de la société INRO (Canada).

Ces quatre logiciels permettent un haut niveau opérationnel chez les praticiens des bureaux d'études, services techniques de l'Etat et des collectivités territoriales, grâce à une structure modulaire et ouverte, des algorithmes éprouvés et régulièrement améliorés après plus de vingt ans d'existence, une interface graphique utilisateur intuitive, des liens vers d'autres logiciels standards du marché (MSOffice, ArcGIS, MapInfo, Google Earth, etc). En pratique, ces logiciels sont parfois utilisés comme outils génériques par manque de moyens ou de temps (autrement dit, les paramètres de calage sur un modèle calé sur une aire d'étude sont utilisés tel quel sur une autre aire). Même si on le voulait, ce pis-aller ne serait pas possible aujourd'hui en modélisation LUTI. **A chaque étude de cas, il faut repartir de zéro** car il n'y a pas de logiciel permettant cette transposabilité et cette facilité de mise en œuvre par des praticiens de la modélisation.

Muni d'un de ces quatre logiciels, un ingénieur modélisateur peut réaliser, *seul*, une étude de trafic en deux à trois semaines selon son expérience et la durée de mise à disposition de données de calage et s'il dispose déjà de matrices origine-destination (même anciennes) et de réseaux codifiés. En pratique aujourd'hui, ce type de simulation s'effectue en mode quasi-industriel.

En revanche, en modélisation intégrée transport-urbanisation, force est de constater que la pratique quotidienne est encore à mille lieues d'égaliser celle de la modélisation de trafic. En effet, il n'existe pas de logiciel du commerce à l'égal de ce que l'on trouve en modélisation

---

<sup>4</sup> En France, il est commercialisé sous le nom de Davisum.

de trafic classique. Ceci s'explique par le fait que l'économie urbaine et l'économie des transports sont deux disciplines enseignées séparément, que les modélisateurs urbains et les modélisateurs de trafic ont deux cultures différentes et ont travaillé parallèlement pendant longtemps en s'ignorant complètement, et que c'est seulement depuis une dizaine d'années que l'intérêt d'associer les deux types de modélisation est apparu. En 2009, le bureau de consultants Cambridge Systematics a conclu dans un rapport remis à la Federal Transit Administration américaine que les modèles LUTI existants ne sont pas encore capables de remplir les objectifs d'évaluation socio-économiques des projets de transports en commun (Cambridge Systematics, 2009).

Il faut souligner que le terme de « modèle intégré » aujourd'hui est abusif. La plupart des modèles LUTI actuels sont en réalité des modèles connectés, c'est-à-dire résultant du couplage *a posteriori* d'un modèle de trafic et d'un modèle d'urbanisation. Il n'existe pas à ce jour d'outil réellement intégré, conçu **dès le départ** dans une approche globale de l'interaction entre système de transport et système urbain.

Parmi les modèles LUTI se trouvent les modèles dynamiques de micro-simulation. Les projets menés en France *SIMAU*RIF (Île-de-France) et SIMBAD (Lyon) à partir de la même plateforme Urbansim (logiciel « land use » américain) ont permis de tirer des enseignements intéressants sur les moyens à mobiliser pour mettre au point un modèle LUTI de micro-simulation et à l'appliquer sur une étude de cas. C'est à partir de ces enseignements et du constat de l'absence d'un logiciel LUTI ergonomique comparable aux quatre logiciels de modélisation de trafic qu'est née l'idée du modèle ULTISIM.

### **2.3. Logiciels de modélisation LUTI**

Supposons qu'une collectivité locale commande à un consultant une étude sur les effets de la transformation de l'autoroute A4 en boulevard urbain entre Paris et Marne-la-Vallée sur l'occupation du sol, avec un délai d'étude de six mois. Le consultant envisage une partie simulation. Si l'on prend comme critères « au moins une application récente en dehors du milieu académique » ou « commercialisé ou potentiellement commercialisable », seuls **six modèles** peuvent être retenus. Par ordre alphabétique : Cube Land, DELTA, MARS, OPUS/Urbansim, PECAS et Tranus. Il faut souligner que ce choix serait le même quel que soit le lieu de l'étude dans le monde.

### 2.3.1. Cube Land

Cube est un logiciel de planification des transports très répandu dans le monde. Il est composé de plusieurs modules dans un environnement Windows : Cube Base pour la gestion des modules, Cube Voyager pour la demande de déplacements et l'affectation, Cube Cargo pour les marchandises, Cube Dynasim pour la microsimulation, etc. Le dernier module disponible depuis début 2010 est Cube Land. Issu d'un modèle développé à l'Université de Santiago au Chili appelé **MUSSA** (Martinez, 1996), il a été racheté par Citilabs aux Etats-Unis et intégré à la plateforme Cube. Le modèle MUSSA est l'un des modèles LUTI **les plus documentés depuis 20 ans** dans la littérature scientifique en modélisation land use. Les sorties graphiques de Cube Land sont visualisables avec l'interface commune de tous les modules de Cube

Cube Land peut être classé parmi les **modèles dits de rente foncière ou d'équilibre général**. Il permet de simuler le marché immobilier de manière désagrégée, en se basant sur des variables qui décrivent les activités et les ménages à localiser, l'offre immobilière résidentielle et de bureaux et les valeurs immobilières et locatives. Cube Land peut prédire la localisation des ménages selon leurs caractéristiques et la localisation des entreprises selon leur type d'activité. Il est possible d'adapter Cube Land à n'importe quel zonage, et de prendre en compte des contraintes telles que les politiques de planification urbaine (aides au logement, niveau de taxes foncières par zone, ...).

Cube Land résout un équilibre entre demande (ménages et entreprises), offre (marché immobilier) et prix. Un modèle de choix discret par maximisation de l'utilité (logit) est utilisé pour simuler les choix des agents contraints par leur propre budget. Cube Land peut être lié au module Cube Voyager pour évaluer l'effet des transports (par le biais des mesures d'accessibilité) sur l'utilisation du sol. Le calcul de l'équilibre est effectué en une fois (« one single shot »).

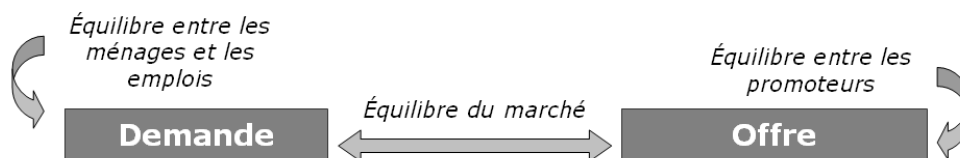


Figure 2 : schéma simplifié de l'équilibre offre/demande dans Cube Land

A la date de la Conférence WCTR à Lisbonne (11-15 juillet 2010), le Président de Citilabs, M. Mickael Clark, nous avait informés que depuis le début de la commercialisation en janvier



2010, une seule licence de Cube Land avait été vendue en Floride et une autre était en cours à Pékin. Un second contact a été pris avec M. Clark fin avril 2011, soit 9 mois après. Selon ses dires, entre 50 et 100 licences de Cube Land ont été vendues depuis, principalement aux Etats-Unis et en Asie, ce qui est remarquable si ces données commerciales s'avèrent réelles et confirme qu'il existe une réelle demande d'un logiciel prêt à utiliser. Il faut néanmoins tempérer cette déclaration, car il n'y a, à ce jour, pas encore de retour d'expériences de ce modèle pour faire un bilan, et surtout pas en Europe. Mais du point de vue du praticien, il est clair que Cube Land propose une solution qui semble **très intéressante** pour quatre raisons :

- Cube Land repose sur un système global paramétrable dans une interface orientée utilisateur final. Il bénéficie de son intégration dans un logiciel de trafic réputé et bien implanté dans le milieu professionnel, et donc pour un client de Cube, l'interface utilisateur de Cube Land lui semblera déjà familière et lui permettra de le coupler plus facilement avec un modèle de trafic fonctionnant déjà avec Cube Voyager ;
- issu du milieu universitaire, Mussa/Cube Land repose sur un modèle classique d'équilibre général suffisamment solide mathématiquement pour ne pas prêter le flanc à la critique, du moins pour la partie localisation résidentielle, même si l'hypothèse de marché parfait du logement est discutable ;
- de par sa conception, il est générique, c'est-à-dire que c'est un logiciel à paramétrer, en utilisant des scripts dans le macro-langage de Cube, mais il ne demande pas de programmation excessive si l'on en juge par le Guide de référence (Citilabs, 2010) ;
- étant un logiciel du commerce, il bénéficie du support d'assistance et de mise à jour que l'on est en droit d'attendre d'un produit acheté et non d'un produit « open source » venant directement du monde universitaire, peu habitué aux principes de travail du monde professionnel (délais, coût, réactivité, satisfaction du client, etc).

Nous pensons que Cube Land présente un **potentiel intéressant**, même si le marché, en France du moins, n'est pas encore mûr. Il n'est pas une « boîte noire » car il est une implémentation du modèle MUSSA sur une plate-forme très connue des modélisateurs de trafic. Au-delà des aspects théoriques, le reproche qu'on peut lui faire est qu'il a été ajouté à la plateforme Cube et non conçu au départ dans une approche systémique de l'interaction transport – urbanisation. Il est donc, lui aussi, un modèle « land use » couplé à un modèle de trafic existant et non un modèle réellement intégré. Il est intéressant aussi de noter que la société PTV, éditeur de Visem/Visum, a choisi une stratégie pragmatique, opposée à celle de Citilabs. Selon PTV, le marché des modèles LUTI n'est pas mûr et il n'y a pas urgence à intégrer un modèle land use dans leur logiciel de planification de transports. En revanche,

PTV a développé des interfaces spécifiques pour relier Visem/Visum à des logiciels d'urbanisation existants tel que Urbansim.

Le coût d'une licence avoisine les 15 000 euros TTC. Pour en savoir plus sur Cube Land, voir [www.citilabs.com](http://www.citilabs.com)

### 2.3.2. DELTA

Le modèle DELTA a été développé par M. D.Simmonds à partir de 1995 en Grande-Bretagne. DELTA signifie Development, Employment, status and communting, Location and property market, Transition and growth, and Area quality (Bramley, Simmonds et Dobson, 2008). C'est un modèle agrégé simulant la croissance économique, les changements d'occupation du sol et le marché immobilier. Il peut être appliqué à un niveau régional ou à un niveau local. Au niveau local, DELTA simule la localisation des ménages et des emplois, les changements démographiques, le taux de motorisation. Au niveau régional, DELTA dispose d'un sous-modèle qui simule les migrations résidentielles longue distance et d'un sous-modèle de croissance économique (investissement, production, commerces). Les choix de localisation sont basés sur des modèles logit.

Dans la famille des **modèles spatiaux input-output**, DELTA est un modèle land use et non un modèle LUTI contrairement à Tranus. Il est connecté à un modèle anglais de trafic appelé SATURN pour former un modèle LUTI. La dernière version datant de 2008 présente des nouveautés : module de micro-simulation pour la localisation des ménages appelé SimDELTA, calculs d'indicateurs de développement durable (émissions de CO<sub>2</sub> par exemple). M. D.Simmonds souligne bien les deux grandes difficultés de l'application de DELTA : le calage global du modèle LUTI et le couplage avec un modèle de trafic existant (Simmonds, 2010). Ces deux difficultés se retrouvent d'ailleurs pour tous les modèles LUTI.

DELTA a été appliqué à plusieurs villes et régions de Grande-Bretagne (Ecosse, nord de l'Angleterre, Londres, Région South East), à la région d'Auckland en Nouvelle-Zélande et à Seoul (Corée). La dernière application en cours concerne le modèle LASER sur Londres et la Région South-East pour remplacer le modèle land use actuel qui utilise le très ancien logiciel MEPLAN. La documentation technique n'est pas disponible au grand public sur leur site [www.davidsimmonds.com](http://www.davidsimmonds.com). De même, il n'y a aucun document permettant de se faire la moindre idée de l'interface (pas même une copie d'écran) et aucune information de prix. Il est fort probable que DELTA n'est pas orienté utilisateur final et que seuls ses concepteurs

savent l'utiliser. Autrement dit, le choix de DELTA pour une application nécessiterait de faire appel à la société de consultants DSC (David Simmonds Consultants) installée en Angleterre. Nous avons eu l'occasion de discuter avec Mr David Simmonds au Colloque sur la « Modélisation de la ville : du modèle au projet urbain » qui s'est tenu à l'École des Ponts les 23-24 février 2011. Il a confirmé notre soupçon sur l'utilisation de DELTA : chaque application nécessite un développement spécifique pour adapter DELTA au contexte du client et donc l'implication du bureau d'études anglais. Il ne présente donc aucun intérêt opérationnel.

### **2.3.3. MARS**

MARS (Metropolitan Activity Relocation Simulator) a été développé en Autriche par Paul Pfaffenbichler à partir de 2000 dans le cadre de sa thèse à l'Université Technologique de Vienne (Pfaffenbichler 2003, Pfaffenbichler *et al.* 2010). C'est un modèle intégré dynamique qui fonctionne à un niveau très agrégé dit « stratégique ». L'idée de base était de mettre au point un modèle simple à comprendre par les décideurs et facile à utiliser. MARS est utilisé en 2010 dans 18 villes dans le monde (dont des capitales européennes – Edimbourg, Helsinki, Madrid, Oslo, Stockholm, Vienne –, Hanoi au Vietnam, Chiang Mai et Ubon Ratchathani en Thaïlande).

Issu d'un travail universitaire, MARS a essayé d'évoluer vers un produit professionnel orienté utilisateur final. Ainsi, on peut trouver de la documentation et des modules à télécharger gratuitement sur le site de l'Université de Vienne : interface de gestion des données d'entrée sous Excel, logiciel Vensim qui est la plateforme de simulation dynamique sur laquelle est programmée MARS, interface pour la visualisation des résultats.

<http://www.ivv.tuwien.ac.at/forschung/mars-metropolitan-activity-relocation-simulator.html>

Du point de vue du praticien, cet outil paraît intéressant et peut inspirer certaines spécifications fonctionnelles d'ULTISIM. Il répond partiellement au cahier des charges d'ULTISIM : nombreuses applications antérieures dans des villes européennes, ergonomie et simplicité. En revanche, on reste insatisfait quant à la partie modèle de trafic qui n'est quasiment pas décrite. On sait seulement qu'il n'y a pas de codage de réseaux de transport, pas de module d'affectation et donc que les matrices de temps sont calculées de manière rudimentaire comme dans les modèles stratégiques de déplacements (voir par exemple le modèle Mostra développé par la SEMALY dans les années 90 et qui tourne simplement sur Excel). Le site internet n'est manifestement pas maintenu, la dernière mise à jour datant de

2007. Le logiciel est gratuit mais selon Mr Simon Shepherd (Université de Leeds) qui a bien voulu répondre à nos questions en avril 2011 et qui est l'un des auteurs de l'article Pfaffenbichler *et al.* (2010) décrivant l'application de MARS sur la région de Leeds (Grande-Bretagne), lorsque des projets utilisant MARS ont pu bénéficier de subventions de la part du client, c'était pour recruter des thésards à l'Université de Vienne. L'histoire de MARS montre que le seul modèle économique permettant de pérenniser un logiciel LUTI est celui de Cube Land qui s'est adossé à un éditeur informatique professionnel.

#### 2.3.4. OPUS/UrbanSim

L'équipe dirigée par le Professeur Waddell à l'Université de Washington a commencé à concevoir UrbanSim en 1997 et à le développer à partir de 1999 (Waddell, 2002). Compte tenu de son importance stratégique, l'équipe a reçu en 2001 une subvention de 5 millions de dollars de l'Etat fédéral américain pour poursuivre son développement. Immédiatement, il a suscité l'intérêt de toute la communauté scientifique des modélisateurs de trafic.

UrbanSim est l'aboutissement de nombreuses recherches commencées dès les années 60, d'abord en Amérique latine, puis aux Etats-Unis. Le Professeur P. Waddell a su tirer des expériences de l'utilisation d'autres logiciels existants tels que MEPLAN, MetroSim, ITLUP, DRAM/EMPAL pour concevoir son logiciel. Il intègre donc toutes les fonctionnalités d'un logiciel de prévision de l'occupation du sol, en particulier un modèle de prix du foncier et des modèles de choix discret de localisation des ménages et des emplois. Il faut reconnaître au Pr P. Waddell sa volonté de sortir UrbanSim du milieu académique et d'essayer de le déployer dans le milieu professionnel.

Mais UrbanSim n'est pas un modèle LUTI, c'est un modèle « land use ». Pour devenir LUTI, UrbanSim doit être couplé avec un logiciel de trafic. Ce couplage est encore aujourd'hui artisanal.

Le logiciel fait partie de la famille des **modèles dynamiques désagrégés basés sur des choix discrets et fonctionnant en micro-simulation**. Il ne cherche pas à atteindre un quelconque équilibre : chaque ménage, chaque emploi et chaque unité géographique (cellule carrée) sont représentés dans le modèle, chacun avec ses caractéristiques et ses comportements de choix. Toutes les hétérogénéités sont ainsi prises en compte. Le modèle intégré va faire passer chaque agent dans le processus de micro-simulation (par exemple

probabilité pour un ménage ou pour un emploi de se localiser dans telle cellule, probabilité pour une cellule vacante de passer à la catégorie habitat).

L'équipe a su fédérer l'ensemble des utilisateurs potentiels dans une communauté virtuelle en créant une liste de diffusion internationale et un site Internet [www.urbansim.org](http://www.urbansim.org). Pour sa diffusion, la stratégie du logiciel dit « libre » (sous licence publique GNU) a été choisie. Le logiciel est développé en langage Java et dans sa dernière version en langage Python. Le code source est disponible de sorte que tout utilisateur peut l'utiliser, le modifier et le redistribuer. UrbanSim a connu entre 2000 et 2008 de nombreuses évolutions et plusieurs phases de réingénierie. Dans sa dernière version 4.3.1., UrbanSim a rejoint une plateforme universelle de simulation urbaine appelée OPUS (Open Platform for Urban Simulation).

Le projet *SIMAUFRIF* sur l'Île-de-France, basé sur le couplage d'UrbanSim et de Davisum et mené entre 2003 et 2007, a été riche d'enseignements (Nguyen-Luong, 2008). Ainsi, il en est ressorti que :

- **UrbanSim n'est pas adapté au cas des villes européennes.** Par exemple, les variables expliquant tel phénomène aux Etats-Unis (par exemple le prix du terrain pour le choix résidentiel) ne sont pas déterminantes dans le cas de la région Île-de-France. La structure du modèle pose des difficultés pour pouvoir prendre en compte dans le cas des villes françaises les contraintes de politiques de régulation (aux Etats-Unis, l'urbanisation est peu contrôlée). De même la variable d'accessibilité par le mode véhicule particulier apparaît sur-représentée dans les différents modules d'UrbanSim, et a contrario on ne trouve pas de variables décrivant l'accessibilité et la desserte par les transports en commun. Il est possible cependant de modifier cette structure par défaut et de définir nos propres variables mais la souplesse n'est pas évidente. Enfin, c'est le modèle de prix qui apparaît le moins adapté aux villes européennes. Urbansim propose un modèle de prix du foncier alors qu'en Europe, c'est plutôt le prix de l'immobilier qui prime.
- **Les différents sous-modèles d'UrbanSim ne sont pas très clairement liés les uns aux autres.** L'approche systématiquement conduite par tous les utilisateurs qui consiste à construire une base de données pour chaque sous-modèle, à estimer les différents sous-modèles indépendamment les uns des autres et dans un but explicatif, ne favorise

pas l'opérationnalité. Ils se retrouvent au final avec des sous-modèles incapables d'interagir ensemble<sup>5</sup>.

- **UrbanSim est trop exigeant en données.** La construction des bases de données en entrée et des bases de données d'estimation a été chronophage dans le projet. Il faut revenir à une structure de données plus simple permettant de réaliser des applications plus rapidement (six mois au maximum et non quatre ans) pour rester dans les temporalités des maîtres d'ouvrage. De plus, une architecture plus simple permettrait plus de transposabilité du modèle d'une étude de cas à l'autre.
- Au fil des années, **le logiciel est devenu de plus en plus complexe et difficile d'utilisation.** On peut malheureusement parler de « complexité non essentielle ». Il nécessite dans sa dernière version des connaissances informatiques pointues (comme le langage Python), quasiment inexistantes dans le milieu des modélisateurs transports. Le développement d'une application d'UrbanSim a des caractéristiques négatives que l'on retrouve dans tous les projets dans le monde : d'une part le calendrier n'est pas maîtrisable (entre 2 et 4 ans pour une application), d'autre part il y a toujours un « effet tunnel »<sup>6</sup>. Ceci n'est pas compatible avec une demande de modélisation rapide par un ingénieur-modélisateur de trafic. Il faut dire que les chercheurs d'UrbanSim à l'Université de Washington n'ont pas tous le sens des objectifs fonctionnels dans la mesure où des applications opérationnelles ne sont pas leur principal centre d'intérêt ;
- Le fait qu'UrbanSim soit **seulement un modèle land use** et non un modèle LUTI nécessite de le coupler avec un modèle de trafic existant. L'utilisateur se trouve alors confronté à la mise en cohérence de deux modèles mis au point indépendamment l'un de l'autre. Par exemple, il n'y a pas de cohérence entre les zonages. Il n'est pas pratique de travailler simultanément sur un zonage TAZ (« Traffic Analysis Zoning ») de 1000 zones et sur un zonage LUZ (« Land Use Zoning ») de 50000 zones. Il n'y a pas de cohérence non plus entre d'une part les motifs choisis pour générer les déplacements et d'autre part les modes d'occupation du sol à simuler. Par exemple, il n'y a aucun intérêt

---

<sup>5</sup> Dans le roman "Vendredi ou les Limbes du Pacifique" (1967) de Michel Tournier, à un moment, après avoir lu la Bible, Robinson décide de rejoindre le Chili en construisant un bateau qu'il nomme L'Evasion. Robinson consacre toute son énergie et son imagination à construire ce bateau. Mais le jour où le bateau est terminé, Robinson se rend compte douloureusement qu'il ne peut pas flotter car il est trop lourd. La même leçon peut s'appliquer à la modélisation LUTI : si votre modèle est trop sophistiqué, vous ne pourrez pas l'appliquer en prévision car les variables retenues lors de l'estimation économétrique ne sont pas disponibles.

<sup>6</sup> Commencer une application d'UrbanSim, c'est comme entrer dans un long tunnel. Il fait très beau au moment où on entre. Quand on en ressort, il pleut et on découvre enfin ce qu'UrbanSim peut faire.

à modéliser un motif achat si l'on ne modélise pas dans UrbanSim les surfaces commerciales. C'est bien cette réflexion globale et en amont qui fait défaut aujourd'hui dans la plupart des projets d'application du modèle UrbanSim.

### 2.3.5. PECAS

PECAS (Production, Exchange and Consumption Allocation System) a été développé à l'Université de Calgary (Canada) par Hunt et Abraham à partir de 2002. Il s'inspire largement d'un modèle plus ancien, MEPLAN, développé dans les années 70 par Echenique et al, qui a été appliqué dans plusieurs régions dans le monde mais qui est complètement dépassé aujourd'hui (il fonctionne sous MS-DOS). Aux Etats-Unis, PECAS est actuellement porté par le milieu universitaire : l'Université Davis en Californie et Caltrans. Il est en cours d'application dans plusieurs aires urbaines en Amérique du nord (Ohio, Oregon, Sacramento, Chicago, Calgary et Edmonton au Canada)

PECAS comprend deux sous-modèles. Le *premier* est un modèle d'interaction spatiale de type input-output pour le court-moyen terme. La structure de données définit des activités (secteurs économiques et ménages) qui produisent ou consomment des commodités (biens, services, travail, capital et surfaces immobilières). Le modèle agrégé simule trois types de choix : choix de localisation des activités, choix du producteur et du consommateur selon sa localisation (quantité de commodités à produire ou à consommer), choix du lieu d'échanges (vendre ou acheter) de ces commodités selon la localisation et les quantités produites et consommées. Les flux d'échanges de producteur à consommateur dans le zonage sont alloués par des modèles logit maximisant des utilités incluant des variables de prix et des variables liées au coût de transport. Ils sont convertis en demande de déplacements puis affectés sur les réseaux de transport, ce qui va déterminer les coûts de transport de zone à zone et influencer à son tour sur l'allocation des flux d'échanges. Les prix des commodités sont ajustés à chaque échange, conduisant au final à un équilibre pour chaque type de commodité.

Le *deuxième* sous-modèle est appelé modèle de développement spatial. C'est un modèle long-terme de choix de localisation de projets urbains qui simule le comportement des promoteurs immobiliers. Les différentes catégories d'espace urbanisé sont des commodités qui doivent être consommées dans les zones où elles sont localisées. Les choix des promoteurs créent à chaque pas de temps de nouvelles contraintes quantitatives à prendre en compte dans les échanges.

PECAS est classé dans les **modèles spatiaux input-output**. Il est lui aussi un modèle land use et non un modèle LUTI. Ses deux sous-modèles tournent dans une boucle de rétroaction annuelle avec un modèle de trafic externe. Le zonage de PECAS peut être un maillage en cellules carrées comme dans UrbanSim.

De nombreuses applications sont en cours de développement en Amérique du nord (Zhong, 2007) et une seule application en dehors, à Mumbai (Inde). Comme pour les applications d'UrbanSim, celles de PECAS nécessitent de trois à cinq ans de développement et une équipe pluridisciplinaire. Par exemple, à Mumbai, l'application de PECAS est toujours en cours de développement depuis 2008 par une équipe d'étudiants de l'Institut de Technologie de Bombay.

De même que pour DELTA, il n'y a pas de site internet (comme pour Cube, Tranus et OPUS/Urbansim) dédié au modèle PECAS. On peut là aussi suspecter que PECAS, qui est un pur produit issu du monde académique, n'a pas encore atteint le stade de la maturité pour être commercialisé et prêt à utiliser.

### 2.3.6. Tranus

Tranus est un logiciel qui a été développé au Vénézuéla par M. de la Barra. Commercialisé en Europe par la société Stratec (Belgique) depuis une quinzaine d'années, il a été le premier logiciel du commerce au monde se présentant sur le marché des modèles de trafic comme un modèle LUTI. Tranus est classé dans la famille des **modèles spatiaux input-output**. Cette famille utilise le concept des matrices input-output de Leontief mais introduit en plus un élément de choix spatial avec des modèles logit. Tranus simule au sein d'un même outil la localisation des activités, le fonctionnement du marché immobilier et le système de transport multimodal. Le mécanisme de localisation des activités repose sur la théorie de la base, les matrices input/output, l'équilibre de l'offre et de la demande par les prix, et des choix discrets de localisation. La localisation des ménages et des emplois génère des déplacements qui sont affectés sur le réseau multimodal de transport. Les coûts généralisés de transport sont, en retour, injectés dans l'expression de l'utilité. Les disponibilités foncières ainsi que les réseaux de transport sont modifiés à chaque pas temporel. Le modèle calcule donc, à chaque pas, la localisation des ménages et des emplois, l'évolution du prix du foncier, les déplacements, et leurs affectations par mode sur les réseaux.



Tranus est le premier logiciel réellement LUTI. Il présente de plus un avantage pour les praticiens qui est son interface graphique de type Windows. De plus, il peut se coupler aisément avec un SIG (Système d'information géographique), ce qui permet de visualiser les effets sur l'occupation du sol des différents scénarios projetés. En revanche, son « vocabulaire » est exclusivement économique et peut paraître peu familier à des ingénieurs non spécialisés en économie.

Du point de vue théorique, les paramètres de TRANUS ne peuvent pas être paramétrés à l'aide de méthodes économétriques. Les paramètres sont ainsi « calibrés » par des procédures adhoc qui ne respectent pas les méthodologies statistiques, mais « à dire d'expert ». L'expérience montre que les prévisions dépendent fortement de ces paramètres définis arbitrairement. De plus, l'équilibre de la rente foncière ne prend pas en compte l'interaction entre les agents économiques (socio-économie de voisinage et économie d'agglomération) ce qui intervient à la fois sur les valeurs foncières et l'occupation du sol.

Tranus a été appliqué dans plusieurs villes d'Amérique du nord, d'Asie et d'Europe (Valence, Bruxelles) entre 1995 et 2005. En France, il y a eu une seule application à Lyon par le CERTU à la fin des années 1990. Elle n'a pas donné satisfaction pour des questions de calage mais aussi des questions de ressources humaines à mettre en place et à pérenniser dans ce type de projet. La conséquence est que cet échec relatif sur Lyon a constitué une douche froide pour Tranus en France. Après la tentative de Lyon, aucune autre ville ne s'est risquée à se lancer dans un projet d'application de Tranus en France dans les 10 années qui ont suivi. Un projet d'application est en cours actuellement à Grenoble dans le cadre du projet Aetic mené par l'Iddri (Institut de Développement durable et des Relations internationales de l'IEP de Paris).

Tranus était payant mais depuis peu, il est devenu gratuit et téléchargeable sur le site [www.modelistica.com](http://www.modelistica.com).

### **2.3.7. Autres logiciels**

Citons ici quelques autres logiciels, par ordre alphabétique. Certains ont fait une percée dans le passé mais sont devenus obsolètes, d'autres sont en cours de développement mais n'ont pas encore été appliqués en dehors de leur lieu de création, d'autres enfin sont de purs produits de la recherche encore cantonnés dans le milieu académique et appliqués à leur seule ville d'origine.

- **ILUTE** : logiciel en cours de développement depuis 2006 à l'Université de Toronto par l'équipe de M. Eric Miller. La stratégie de développement est la même que celle d'UrbanSim. C'est un produit issu du milieu académique, qui n'en est pas encore sorti, et qui fait travailler une dizaine de chercheurs et de thésards, chacun travaillant seul sur un aspect du logiciel, sans forcément avoir prévu des entrées-sorties communes avec les modules développés par ses collègues. Comme OPUS/Urbansim, c'est un modèle land use et non LUTI. Il s'interconnecte avec le logiciel canadien de trafic Emme/3. Il n'est pas orienté utilisateur final (environnement de programmation open source « R ») et il est fort probable que le produit final sera très difficilement adaptable vers le milieu professionnel.

- **IRPUD** : développé par M. Wegener dans les années 80 à l'Université de Dortmund, il est conçu comme un véritable modèle LUTI. Une seule application est décrite dans la littérature, celle de la région de Dortmund. On peut regretter que ce logiciel soit resté un produit de laboratoire, alors qu'il présentait des atouts pour se diffuser dans le milieu professionnel, peut-être en s'attachant les services d'un éditeur de logiciels tel que PTV, comme le logiciel chilien MUSSA racheté par Citilabs.

- **ITLUP** (ex DRAM/EMPAL) : développé par Putman et financé dans les années 90 par le DOT fédéral américain, il a été utilisé dans une quinzaine de Metropolitan Planning Organizations (MPOs) mais est devenu obsolète. Il est classé dans les modèles dits d'interaction spatiale.

- **LUM** : développé par Alex Anas et Associés, pour la région de New-York. Inutilisé et non connecté au modèle de trafic du MPO.

- **MEPLAN** : développé par M. Echenique and Partners Ltd. en Angleterre dans les années 80, il est utilisé dans une douzaine de régions dans le monde. Il a l'avantage d'être un vrai modèle LUTI, incluant en interne un modèle de trafic. Il appartient à la même famille que Tranus mais il n'a pas su évoluer avec la technologie en terme d'ergonomie et est devenu par conséquent obsolète.

- **METROSCOPE** : développé par la société In-house, pour la région de Portland. Il n'est pas utilisé dans les études d'évaluation socio-économique.

- **PIRANDELLO** : modèle français d'équilibre général développé par Vinci Concessions. Il y a eu une application sur l'Île-de-France en 2008, une autre est en cours à Lyon dans le cadre du projet de recherche ANR Plainsud. C'est un modèle en devenir, porté par un groupe

puissant. C'est le premier modèle en France réellement LUTI car il intègre un modèle de trafic, pour l'instant uniquement routier, et utilise le modèle ANTONIN du STIF pour la partie TC. Il permet aussi de calculer certains effets d'agglomération, ce qui est inédit. S'il veut se diffuser en France, Vinci devra sans doute investir dans la modélisation du mode transport en commun et surtout dans l'ergonomie du logiciel (programmé en Fortran, il n'est utilisable actuellement que par ses concepteurs), quitte à s'adosser à un éditeur de logiciels. Vinci Concessions a déjà créé une structure ad hoc pour poursuivre le développement du modèle : Pirandello Ingénierie.

- **RELU-TRAN** (« modèle de Chicago ») : prototype développé par le Professeur Alex Anas et appliqué en 2007 à Chicago après les modèles CATLAS et MetroSim devenus obsolètes. Il ne prend pas en compte le réseau de transport en commun, mais uniquement le réseau routier, se concentre davantage sur les emplois que sur l'habitat (Anas et Liu, 2007) et est basé sur un découpage stratégique d'une cinquantaine de zones. Même s'il présente des aspects théoriques très intéressants, son adaptation et calage à l'Île-de-France nécessiteraient un volume de travail difficile à évaluer et exigerait la participation du Pr Anas.

- **TIGER** : développé par l'Université de Louvain. Similaire à PECAS. Quasiment pas documenté.

- **TIGRIS XL** : modèle néerlandais développé par Significance (De Bok, Geurs, Zondag, 2010) en coopération avec MM. Wegener et Louter. C'est un modèle dynamique avec une représentation des marchés de l'immobilier et du travail, simulant les changements année par année. Il est basé sur la théorie de l'utilité aléatoire. Les fondements théoriques sont comparables à ceux des modèles UrbanSim et DELTA. Le modèle a été appliqué dans plusieurs études de cas aux Pays-Bas. Pour l'instant, Significance ne prévoit pas de le transformer en un logiciel de type commercial.

## 3. Critères de développement d'ULTISIM

### 3.1. Une absence de logiciel LUTI destiné au milieu professionnel

Revenons au problème de notre consultant à qui on a confié l'étude sur les effets sur l'urbanisation de la transformation de l'autoroute A4 en boulevard urbain entre Paris et Marne-la-Vallée.

Il se rend compte qu'il ne peut pas faire de simulations dans le délai de six mois qu'on lui a imposés. Il aurait fallu qu'il dispose d'un logiciel réellement LUTI, générique (c'est-à-dire paramétrable et ayant déjà été appliqué ailleurs), incluant un modèle de trafic, avec une interface utilisateur de type Windows, nécessitant des données de base dans un découpage suffisamment fin (mais pas trop non plus) et surtout homogène entre la partie trafic et la partie land use, avec une structure de données pré-définie et des modèles de choix à estimer dans une structure de paramètres pré-définie, tout en permettant une marge de flexibilité (par exemple le choix des motifs de déplacements par le praticien).

Ce qui vient d'être décrit est tout simplement l'esquisse du cahier des charges du logiciel ULTISIM. Il faut concevoir un « framework » (une charpente) supposé être la recette optimale pour le projet ULTISIM mais sans aller trop loin dans le niveau d'abstraction (il ne s'agit pas de décrire la maison). **Notre objectif est bien de faire entrer ULTISIM dans le monde du logiciel professionnel.** Pour cela, il faut aboutir à un logiciel convivial incluant des modèles simples, adaptés aux besoins et aux données disponibles. Attention, nous ne confondons pas simple et simpliste. Le simplisme ignore les complexités tandis que la simplicité résout les difficultés. Viser la simplicité est une tâche plus ardue que rechercher la complexité.

Les questions qui ont guidé nos réflexions concernent les sept points suivants : classe de modèle d'urbanisation, mode de fonctionnement, niveau du zonage, type de modèle de demande de déplacements, prise en compte des réseaux et de l'affectation, structure de données, niveau d'ergonomie du logiciel.

### **3.2. Modèle d'équilibre général versus modèle dynamique**

Il existe plusieurs classifications des modèles d'urbanisation (Waddell, 2002, Wegener, 2003). Si l'on simplifie, les modèles d'urbanisation peuvent se classer en deux catégories : les **modèles statiques d'équilibre général** (Cube Land, Mars, PECAS, Tranus) et les **modèles dynamiques** (Delta, OPUS/Urbansim).

Les modèles d'équilibre général reposent sur la théorie de la rente foncière (« bid-rent theory ») développée par Alonso (1960). Cette théorie rend compte des affectations du sol autour d'un centre urbain en fonction des rendements décroissants des diverses utilisations de l'espace selon son éloignement du centre. Pour Alonso, les ménages par exemple, cherchent à maximiser leur utilité en choisissant le lieu de résidence en fonction de deux paramètres économiques : le coût du logement et le coût du transport. Il est donc possible d'exprimer combien les ménages sont prêts à payer pour résider dans les différentes zones. C'est l'équilibre entre offre et demande sur le marché immobilier qui fixe les prix et les usages du sol. Le nom de « bid-rent theory » traduit l'idée que l'affectation du sol aux différentes fonctions concurrentes résulte d'un processus d'enchères entre celles-ci. Chaque parcelle est utilisée par l'activité qui offre la plus haute enchère pour l'occuper et donc par l'activité qui en retire l'utilité maximum. L'avantage de cette approche est que **les prix sont endogénéisés**. L'inconvénient est que théoriquement il ne correspond qu'à un modèle de ville monocentrique. Un modèle d'équilibre général se résume au final à quelques équations qu'il faut résoudre numériquement. Son calibrage se fait sur une année de référence.

Les modèles dynamiques font appel à la théorie des choix discrets développée par McFadden (1975). Les modèles logit<sup>7</sup> permettent de simuler les choix des ménages, des entreprises et des promoteurs immobiliers. Il n'y a pas de recherche d'équilibre. Cette approche nécessite d'ajouter au modèle un sous-modèle de prix du foncier et de l'immobilier car les prix ne sont pas mis à jour automatiquement. Elle a l'avantage de pouvoir s'appliquer à n'importe quel type de territoire (monocentrique ou polycentrique) et est moins abstraite que le modèle de la rente foncière. Mais certains experts comme le Professeur Alex Anas aux Etats-Unis critiquent l'approche du non-équilibre et considèrent que les valeurs foncières et immobilières sont évaluées de manière erronée car non conforme à la théorie micro-économique (Anas et Liu, 2007). Le calibrage d'un modèle dynamique se fait sur une période passée (« backcasting ») et non sur une année.

---

<sup>7</sup> Les modèles logit sont aussi utilisés dans les modèles d'équilibre général.

### **3.3. Microsimulation versus simulation agrégée**

Les modèles land use peuvent aussi être classés selon leur mode de fonctionnement. Comme en modélisation de trafic classique, les modèles d'urbanisation agrégés regroupent les agents dans des classes de population selon leurs caractéristiques et des classes d'entreprises selon leur type d'activité. Tous les agents d'une même classe sont supposés avoir les mêmes préférences de choix. A contrario, les modèles de microsimulation décrivent les comportements des agents de manière individuelle. Chaque agent est décrit de manière désagrégée, avec ses caractéristiques et ses probabilités de choix. Par exemple, dans le modèle *SIMAU*RIF basé sur OPUS/UrbanSim, la table des ménages en 1999 comporte environ 5 millions de lignes, chaque ligne correspondant à un ménage.

L'avantage de la microsimulation est qu'il permet une **approche moins abstraite** que la simulation agrégée qui décrit de manière analytique les comportements homogènes des agents des différentes classes. Cependant, il faut être très prudent dans l'analyse spatiale des résultats. Basé sur la méthode de Monte-Carlo (générateur de nombres aléatoires), les sorties des modèles de micro-simulation ne sont pas valides à un niveau détaillé (tel que la localisation d'une centaine de ménages de type *i* dans la cellule *n*) à cause des variations stochastiques. L'analyse des résultats des modèles de microsimulation doit être effectuée à un niveau agrégé. Il existe un autre inconvénient d'ordre pratique de la microsimulation : les temps de calculs. Par exemple, une simulation de *SIMAU*RIF sur l'Île-de-France dure 7 heures sur un PC standard en 2009, ce qui peut limiter le temps consacré à la phase de calage global qui nécessiterait, au bas mot, une centaine d'exécutions pour ajuster les paramètres des différents sous-modèles, en particulier ceux des variables dummy de localisation agrégée.

Pour la partie modélisation de trafic, on en restera *dans une première étape* au modèle agrégé classique à quatre étapes (voir paragraphe 3.6.). En ce qui concerne l'affectation, la microsimulation de trafic qui permet de simuler les déplacements sur les réseaux seconde par seconde en heure de pointe<sup>8</sup> n'est pas en cohérence avec la modélisation land use qui simule les effets sur l'occupation du sol à moyen-long terme. En revanche, en ce qui concerne la demande, on peut envisager *dans une deuxième étape* de mettre au point un modèle de demande similaire à celui d'ANTONIN qui est un modèle statique désagrégé. Mais il faut bien garder à l'esprit que l'effort doit porter d'abord sur la partie modèle d'urbanisation et l'intégration d'un modèle de trafic, et non sur le modèle de trafic lui-même.

---

<sup>8</sup> Voir les logiciels de micro-simulation tels que Vissim (de PTV), Transims (logiciel américain), etc

En conclusion, dans un premier temps, le paradigme du « tout désagrégé » ne sera mis en œuvre que sur la partie land use.

### 3.4. Zonage

Il est pertinent et surtout pratique que la partie modèle de trafic et la partie modèle d'urbanisation reposent sur **le même zonage**, ce qui semble sensé et pourtant, on peut s'en étonner, cela n'a jamais été préconisé dans la littérature. Le logiciel ULTISIM devra être verrouillé sur cet aspect, c'est-à-dire ne pas autoriser deux zonages différents. La question est de savoir s'il faut retenir un zonage agrégé tel que le découpage communal ou un regroupement de communes dans le cas français, ou bien un zonage fin tel qu'un carroyage en cellules carrées de 300 à 500 mètres de côté. En fait cela dépend des résultats que l'on attend d'un modèle LUTI. Si l'on veut évaluer les effets autour d'une gare par exemple ou le long d'une voie rapide, il faut nécessairement un zonage relativement fin. Mais s'il est trop fin, les résultats ne sont pas fiables à ce niveau détaillé et surtout, la collecte et le formatage des données à ce niveau risquent d'être difficiles, chronophages et générateur de biais. Si le découpage est trop agrégé, plus rudimentaire que celui du découpage du modèle de trafic (le TAZ », Traffic Analysis Zoning), il risque d'être trop peu sensible aux gains d'accessibilité. En effet, les log-sum sont toujours calculées au niveau du TAZ et les zones du découpage land use héritent des log-sum, ce qui peut sérieusement diluer la capacité du modèle à estimer les bénéfices élargis.

Nous nous orientons pour ULTISIM vers un zonage intermédiaire existant, qu'on pourrait qualifier de **mésoscopique**. En Île-de-France, ce serait le découpage **IRIS**<sup>9</sup> de l'INSEE (il y a 5100 zones IRIS sur l'ensemble de la Région, contre 1300 communes et 50000 cellules carrées de 500 mètres dans le modèle *SIMAURIF*). Reste à savoir si dans les villes européennes, il existe des découpages analogues. D'où l'intérêt dans notre recherche d'analyser les données disponibles dans deux autres cas que l'Île-de-France (Aire urbaine de Turin et la Randstad, voir paragraphe 5.3).

Il faut cependant mettre en garde l'utilisateur du niveau IRIS sur la plus grande fragilité statistique par rapport au niveau communal ou supra-communal. La diffusion des données à l'IRIS par l'INSEE est assortie d'indications sur la qualité des données. Selon ces indications, il est recommandé de procéder à des regroupements d'IRIS.

---

<sup>9</sup> Ilots Regroupés pour l'Information Statistique. Les communes de plus de 5000 habitants sont découpées en IRIS, répartis en 3 catégories : IRIS d'habitat, d'activités et divers. Le découpage en IRIS a été partiellement réactualisé entre les recensements de 1999 et de 2006.

### **3.5. Structure de données**

La structure de données sera **pré-établie** dans ULTISIM. Sa conception est donc fondamentale car elle conditionne toute la méthodologie d'application : recueil et mise en forme de données (quantité, qualité), estimation des modèles économétriques, chaînage des sous-modèles. Si l'on retient le principe de la microsimulation, on s'orientera vers une structure de données comparable à celle d'OPUS/UrbanSim mais avec un nombre d'attributs plus réduit et des attributs adaptés au cas des villes européennes, en somme un **commun dénominateur** à tous les cas mais avec une possibilité d'ajouter quelques variables spécifiques.

### **3.6. Modèle de demande de déplacements**

Depuis que le modèle à quatre étapes existe (cela fait plus de cinquante ans), on ne peut pas dire qu'il y ait eu des ruptures franches dans la modélisation classique (TRB, 2007). Le milieu de la recherche est assez actif depuis une dizaine d'années pour essayer de faire évoluer ce modèle : les modèles basés sur les activités (« activity based models »), les modèles dynamiques (par exemple METROPOLIS développé par le laboratoire THEMA de l'Université de Cergy-Pontoise) et les modèles désagrégés fonctionnant en micro-simulation (par exemple le modèle américain Transims) pour l'étape d'affectation. Mais quasiment partout, en pratique, on en est resté au modèle agrégé classique, amélioré seulement depuis quelques années par la rétroaction automatique entre les quatre étapes. C'est donc ce modèle classique qui sera retenu dans ULTISIM. A terme, **le modèle de demande de déplacements (génération, distribution et choix modal) sera complètement implémenté dans ULTISIM**, avec un paramétrage permettant au praticien de choisir des motifs de déplacements. Pour le cas de l'Île-de-France, il y aura une étape transitoire d'utilisation du modèle ANTONIN du STIF et d'une interface de connexion entre ULTISIM et ANTONIN.

### **3.7. Réseaux et affectation**

Une des grandes difficultés auxquelles on sera confronté est le développement d'une interface graphique conviviale de création-gestion des réseaux de transport multimodaux et de visualisation des résultats, ainsi que des algorithmes d'affectation. Les logiciels tels que Visem/Visum et TransCAD en ont fait leur argument principal de vente, la modélisation de la demande passant au second plan (c'est moins le cas pour Cube). Or il n'est pas dans notre



objectif d'investir en priorité dans ce type de développement. **Notre priorité doit d'abord être l'intégration parfaite d'un modèle de trafic et d'un modèle d'urbanisme** pour nous différencier des logiciels de trafic du commerce. Il faudra cependant envisager dans une phase ultérieure de développement l'intégration de ces fonctions indispensables à tout logiciel LUTI digne de ce nom, qui doit donc se suffire à lui-même.

Dans une première étape de la phase de développement, on se limitera à développer pour l'Île-de-France une interface entre ULTISIM et le modèle ANTONIN du STIF qui tourne sur Cube Voyager.

### ***3.8. Modes à prendre en compte***

En ce qui concerne la partie modélisation de trafic, il est habituel de se restreindre aux modes principaux : véhicules particuliers et transports collectifs. Nous ajouterons dans ULTISIM les modes **deux-roues** (motorisés et vélos) car il y a une demande croissante d'évaluation des politiques de circulations douces (vélos) dans les grandes villes européennes. En revanche, nous ne nous engageons pas dans l'intégration du mode marche à pied qui requiert une modélisation très spécifique et un découpage ultra-fin de l'aire d'étude. Enfin, nous envisageons de prendre en compte le trafic routier **poids-lourds**, soit d'une manière simple par pré-chargement des axes routiers à partir de comptages à collecter, soit en modélisant une matrice poids-lourds qui nécessite une expertise spécifique en logistique (voir par exemple le modèle FRETURB développé au LET à Lyon).

### ***3.9. Ergonomie du logiciel***

Cette question de l'ergonomie du logiciel est **fondatrice**. L'obstacle que rencontre le logiciel OPUS/UrbanSim pour s'implanter dans le milieu professionnel est son utilisation trop complexe qui nécessite des compétences pointues en langage Python. Même son interface GUI (Graphical User Interface) mise au point récemment, mais à notre avis trop tard (dix ans après le début du projet UrbanSim), n'a pas permis de renverser la vapeur et d'inciter plus de professionnels à se lancer dans une application. L'ingénieur-modélisateur veut rester dans son cœur de métier et consacrer sa réflexion à son étude d'évaluation socio-économique : collecte de données, estimation économétrique des sous-modèles, calage global, construction des scénarios, exécution, analyse des résultats, cartographie, calcul de

rentabilité, ce qui est déjà beaucoup. Compte tenu de la problématique actuelle des ressources humaines dans les bureaux d'étude et les services techniques de l'Etat et territoriaux, il n'a plus à ses côtés comme dans le passé des techniciens à sa disposition. Quant aux informaticiens aux compétences pointues, ils seraient bien utiles pour assister les modélisateurs mais ils préfèrent travailler dans des entreprises plus rémunératrices. Se retrouvant seul, l'ingénieur-modélisateur se doit d'être autonome et si possible polyvalent. Mais il y a des limites à cette polyvalence. Il est donc demandeur d'un logiciel qui lui facilite au maximum le processus de modélisation intégrée, l'abstrait de tout développement informatique complexe et l'affranchit d'assistance technique.

**La conception de l'interface homme-machine (IHM) sera donc une priorité** du projet ULTISIM. Nous nous orientons vers une méthode de type Agile c'est-à-dire de prototypage rapide et itérative (voir partie 4 sur la méthodologie de développement). Cette méthode Agile se veut plus pragmatique que les méthodes traditionnelles. Impliquant au maximum l'utilisateur final et permettant une grande réactivité à ses demandes, elle vise la satisfaction réelle du besoin de l'utilisateur final et non les termes d'un contrat de développement. Les étapes de programmation ne sont pas nécessairement séquentielles au niveau de la globalité du projet. La question de l'IHM est étroitement liée à la méthodologie de développement du logiciel qui sera développée dans la partie 4.

Une maquette très rudimentaire a été développée avec le logiciel Flash CS4. Elle permet de donner une première idée du produit final : un logiciel LUTI complètement intégré, où l'on peut trouver dans un même environnement un modèle de trafic et un modèle d'urbanisation à paramétrer. Elle illustre au travers de menus déroulants que le modèle de trafic et le modèle land use sont étroitement imbriqués et interdépendants dans le choix des variables, et ce dès le stade de la définition du zonage commun. Donc une interface utilisateur conviviale permettra de maximiser l'efficacité opérationnelle de l'utilisateur final.

Ce **prototypage rapide** correspond à une tendance lourde dans le monde du logiciel qui est le développement « agile », c'est-à-dire la livraison au plus vite de maquette informatique à présenter au client final pour le faire réagir, et non la livraison de papier. Le développement agile permet de maîtriser la complexité en faisant participer l'utilisateur final à la conception et au développement du logiciel.

L'avantage de Flash CS4 est qu'il permet d'ouvrir l'application sur un explorateur web et donc envisager à terme de proposer le logiciel ULTISIM dans un environnement de type

« cloud computing » (logiciel utilisable « on-line » sur internet et stockage des données à distance, comme le propose de manière précurseuse Citilabs avec leur module Mint).

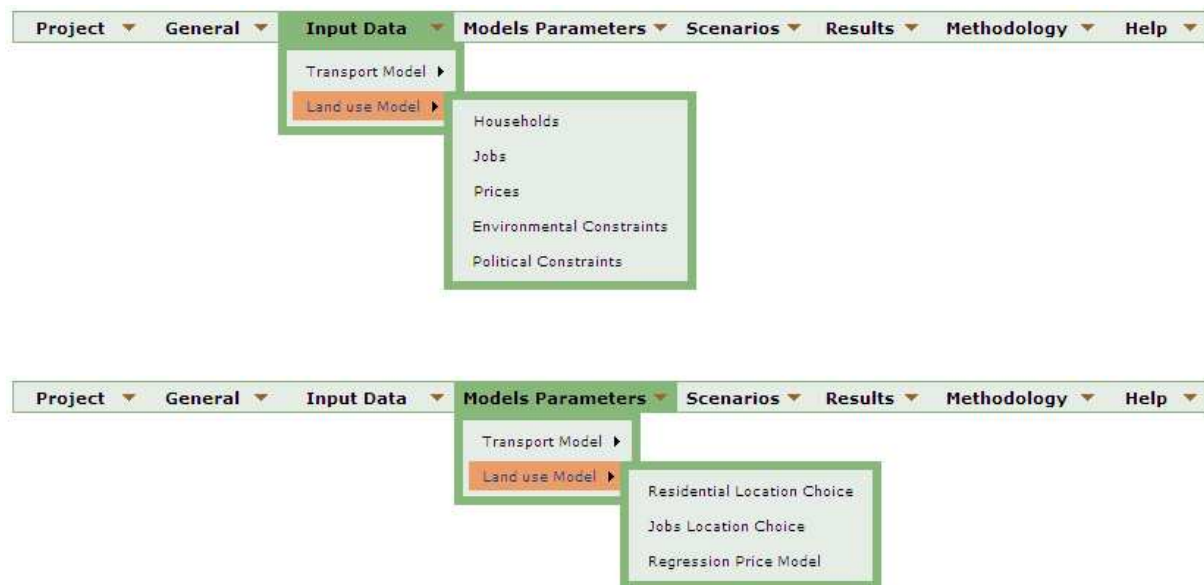


Figure 3 : Maquette de l'interface utilisateur d'ULTISIM

### 3.10. Positionnement d'ULTISIM

La modélisation intégrée transport-urbanisme constitue une avancée importante dans la modélisation en transports. Elle offre un nouveau paradigme et élargit vers de nouvelles disciplines la « vieille » modélisation de trafic à quatre étapes.

**Nous plaçons d'emblée ULTISIM sur le terrain du Logiciel et non du Modèle théorique.**

Le développement d'ULTISIM *ex-nihilo* nécessitera de longues années et des investissements importants, mais pas de la même ampleur que le produit OPUS/UrbanSim car nous savons quelles sont les erreurs à éviter et nous envisageons une méthodologie moderne de type « Agile » (voir partie 4) de plus en plus courante chez les éditeurs de logiciels. Notre objectif est bien de mettre au point un logiciel professionnel, prêt à être commercialisé. Il est clair que nous sortons du cadre de la recherche pure, mais restons dans celui de la « R&D » (recherche appliquée et développement). En se donnant cet objectif, nous orientons la conception et les fonctionnalités à développer afin qu'ULTISIM soit destiné au praticien, générique, adapté aux villes européennes, réellement intégré et pas trop exigeant en données.

La demande d'un tel logiciel n'arrive pas encore aujourd'hui à s'exprimer de manière formelle chez les maîtres d'ouvrage, même si on sent bien que l'argument principal aujourd'hui pour mettre en avant un grand projet d'infrastructure est l'impact sur le développement économique et non plus les trafics générés qui sont de toutes façons, ne l'oublions pas, un résultat des modèles LUTI (qui peut le plus peut le moins). Nous avons donc encore quelques années devant nous pour mettre au point ULTISIM mais il ne faut plus tarder.

La figure 4 montre le positionnement possible d'ULTISIM si le mode micro-simulation est retenu. On peut envisager de faire évoluer ULTISIM vers un modèle de micro-simulation à 100%, c'est-à-dire y compris la partie simulation de trafic. Mais ceci est une vision à long terme.

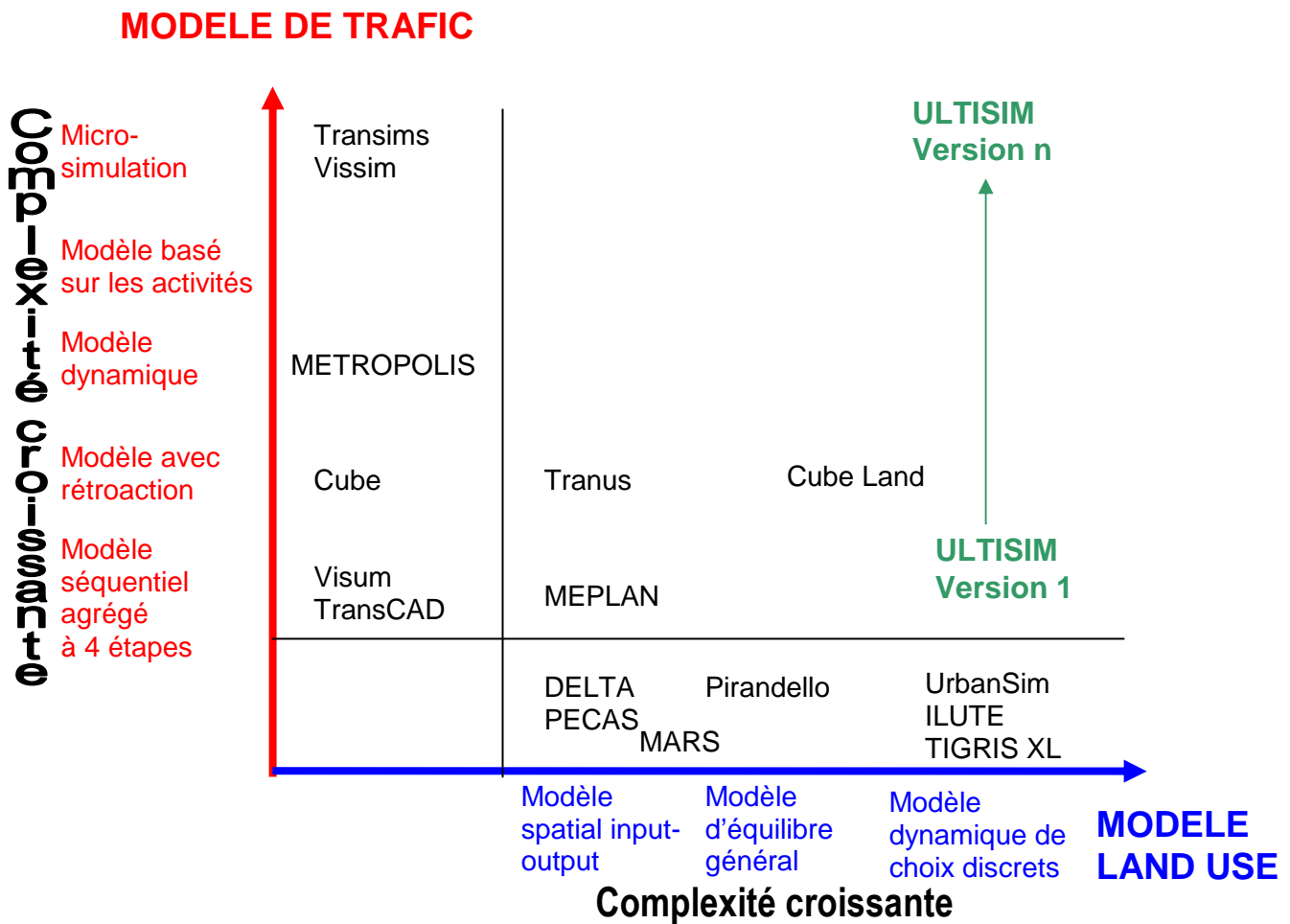


Figure 4 : positionnement d'ULTISIM comme logiciel LUTI

## 4. Méthodologie de développement

Le problème majeur dans le développement d'un logiciel LUTI se situe au niveau de **l'ingénierie du logiciel** (ou génie logiciel). Cette partie semblera peut-être à certains sans rapport avec le sujet de la recherche ULTISIM et pourtant, elle est **fondamentale**. A la limite, s'il n'y avait qu'un seul enseignement à retenir de ce rapport, ce serait même celui-ci. En général, les praticiens de la modélisation et les chercheurs n'ont pas cette formation en ingénierie du logiciel qui est un métier à part entière. C'est cette méconnaissance qui au final aboutit à des outils LUTI insatisfaisants, car non finis, bogués, inadaptés, non documentés ou inutilisables par des non spécialistes de l'informatique. Le développement de logiciels se déroule aujourd'hui dans un contexte d'innovation informatique accélérée et une complexité accrue difficilement gérable. A cela il convient d'ajouter une pression à laquelle les développeurs doivent faire face, notamment les réductions de budgets et de délais.

En ingénierie du logiciel, on distingue aujourd'hui **deux écoles** : les adeptes des logiciels open source d'un côté, les adeptes des logiciels propriétaires de l'autre. Le choix entre ces deux mondes informatiques dépend de plusieurs facteurs : la question du coût, de l'innovation technologique (nouveaux langages de développement par exemple), des attentes de l'utilisateur final et de ses compétences et le niveau de maintenance du logiciel.

Passons d'abord en revue les quatre approches existantes en ingénierie du logiciel et leurs principales limites (Strohmeier et Buchs, 1996) puis nous discutons du choix entre l'open source et le logiciel propriétaire.

### 4.1. Outils de développement rapide

Ces outils sont basés sur le principe de la génération de code. Certains outils connus, catégorisés sous la notion de **RAD** (Rapid Application Development) tels que Windev ou 4D, procurent actuellement un bon moyen de livrer une application dans un délai relativement court.

Cependant, si ces applications atteignent un certain niveau de complexité, les outils se révèlent souvent incompatibles avec la sophistication des fonctionnalités demandées, dans la mesure où ces outils sont très efficaces pour répondre à leur fonction de base, mais incapables d'une quelconque valeur ajoutée dès lors que celle-ci n'a pas été prévue au départ.

Le principal problème pour les clients est alors l'incapacité à répondre aux besoins fonctionnels de l'application. C'est un risque peu acceptable pour les clients qui ont des besoins avancés.

## **4.2. Ateliers de Génie Logiciel**

Les Ateliers de Génie Logiciel (AGL) existent depuis longtemps et ont été largement vulgarisés pendant la vague client/serveur des années 90.

Les outils tels que Powerbuilder, Magic, NS-DK ont pris leur essor à cette époque car ils simplifiaient notamment les langages complexes tels que C++, langage standard à l'époque. Mais ces AGL ont perdu de leur popularité en raison de l'émergence d'un tout nouveau modèle de programmation avec le Web. La conséquence est que bon nombre de clients se sont à nouveau focalisés sur des langages tels que Java ou C#.

Une autre raison était la crainte de certains clients d'être verrouillés dans l'environnement d'un fournisseur, dans la mesure où ces outils fournissaient leur propre langage de programmation.

Finalement le principal problème pour les clients relevait de leur incapacité à se maintenir à jour des technologies, dans la mesure où la contrepartie de ces offres était la lenteur de leur évolution. A ce jour, les spécialistes en génie logiciel affirment qu'il n'existe pas d'AGL efficace tenable à long terme pour les clients.

## **4.3. L'off-shore**

Une autre approche consiste à réduire le coût des ressources via l'off-shore, c'est-à-dire la **sous-traitance** du développement informatique dans des pays à très bas coûts de main d'œuvre. Bien que cela semble économiquement alléchant avec des coûts allant de 1 à 10 suivant les pays, cette solution est loin de tenir ses promesses.

En fait, lorsque cela fonctionne, cela ne répond pas fondamentalement au défi mais diminue simplement le coût apparent des ressources de développement en faisant peser de nouveaux risques sur le projet, en particulier sur les délais. Au contraire, le coût total d'une application doit inclure les coûts de maintenance. Cette phase est bien souvent plus longue que la phase de développement initial. De plus, éloigner l'équipe de développement ajoute des difficultés pratiques (par exemple le décalage horaire lors des échanges téléphoniques) et empêche le client d'augmenter son savoir-faire.

Et ce qui peut paraître économique de prime abord se révèle problématique, avec des coûts de maintenance additionnels imprévus et de la perte d'agilité à long terme. Tout développeur

expérimenté sait que les possibilités d'évolution d'une application mal conçue sont faibles alors que toute évolution, même mineure, devient de plus en plus coûteuse (AFDEL, 2009).

#### **4.4. La méthode Agile**

Selon nous, les trois approches précédentes mettent l'accent de manière excessive sur les délais, le périmètre et le coût. Ces trois paramètres doivent être équilibrés de manière appropriée et c'est pourquoi nous privilégions une autre méthode pour relever le défi du développement d'ULTISIM en tant que logiciel LUTI professionnel. Cette méthode Agile s'oppose donc aux méthodologies classiques comprenant plusieurs tâches : l'analyse du besoin, l'élaboration des spécifications, la conceptualisation du mécanisme interne au logiciel, les techniques de programmation, le développement, la phase de test et finalement la maintenance, avec une équipe hiérarchisée et constituée de managers, sous-managers, chefs de projet, analystes, analyste-programmeurs, développeurs « off-shore », CDD, stagiaires. Au contraire, dans la **méthodologie Agile** (en France, elle s'appelle Scrum, qui est un terme emprunté au rugby et signifie mêlée), ce processus s'articule autour d'une petite équipe soudée et peu hiérarchisée, qui cherche à atteindre un but, comme c'est le cas au rugby pour avancer avec le ballon pendant une mêlée. Le principe de base de Scrum est de focaliser l'équipe de façon itérative sur un ensemble de fonctionnalités à réaliser, dans des itérations de durée fixe de une à quatre semaines, appelées sprints. Chaque sprint possède un but à atteindre, défini par le directeur de produit, à partir duquel sont choisies les fonctionnalités à implémenter dans ce sprint. Un sprint aboutit toujours sur la livraison d'un produit partiel fonctionnel. Pendant ce temps, le ScrumMaster a la charge de réduire au maximum les perturbations extérieures et de résoudre les problèmes non techniques de l'équipe.

Un principe fort en Scrum est la participation active de **l'utilisateur final** pour définir les priorités dans les fonctionnalités du logiciel et pour choisir celles qui seront réalisées dans chaque sprint. Il peut à tout moment compléter ou modifier la liste des fonctionnalités à réaliser, mais jamais celles qui sont en cours de réalisation pendant un sprint.

#### **4.5. Open source versus logiciel propriétaire**

Le mouvement « free and open-source software » existe depuis longtemps mais a pris son essor avec Internet à la fin des années 90. Les pionniers étaient les logiciels Linux, Apache, Ingres. Au début, installer Linux à la place de Windows, c'était un signe de rébellion contre

l'hégémonie de Microsoft. Aujourd'hui, le choix est plus **rationnel**. D'ailleurs, les éditeurs de logiciels propriétaires ont bien saisi le mouvement open-source et au lieu de le combattre, préfèrent aujourd'hui intégrer des composants open source dans leurs logiciels ou d'offrir aux produits open source des composants propriétaires. C'est le choix par exemple de la compagnie Sleepycat (rachetée en 2006 par Oracle) qui distribue d'une part gratuitement le logiciel de gestion de base de données Berkeley sous la licence Open Source, d'autre part sous une licence payante de logiciel propriétaire qui inclut de nombreuses fonctionnalités non disponibles dans la version gratuite (Fogel, 2010).

Mais revenons un peu en arrière. En 1998, les pionniers de **l'open source** se sont regroupés pour établir une charte des critères de définition d'une licence open source (Lerner et al, 2003) :

- Le code source d'un programme open source doit être gratuit ou livrable à un prix symbolique ;
- La redistribution du programme doit être autorisée et ce gratuitement ;
- Il en est de même pour la distribution du code modifié.

Ces critères étaient suffisamment larges pour englober à la fois la licence "General Public" (GPL : General Public License) et les licences qui permettent aux utilisateurs une plus grande liberté d'utilisation du code source. Ainsi, l'open source peut paraître à première vue un modèle séduisant pour les logiciels LUTI dont le développement nécessite des investissements conséquents. La possibilité d'accéder au code source directement peut aussi aider à comprendre comment les modèles fonctionnent et non plus comme dans une boîte noire (Patterson et al., 2007). De plus, l'utilisation de la licence Open Source est vue comme une condition préalable au **développement collaboratif**, comme l'explique le Professeur P. Waddell pour son projet OPUS/Urbansim (Waddel et al., 2005) : "L'open source permet un partage égal de l'information à tous les participants au projet, ce qui est crucial pour motiver les développeurs et encourager les utilisateurs à apporter leurs contributions. Au contraire, un projet de développement propriétaire d'un logiciel est soumis au code de la propriété intellectuelle et oblige les concepteurs à garder secret les informations pour préserver un avantage compétitif ». Certains détracteurs des logiciels propriétaires se revendiquent comme des résistants à l'hégémonie de Microsoft et de sa plateforme intégrée .NET.

En revanche, les **avantages du logiciel propriétaire** sont connus : facilité d'utilisation, stabilité, fiabilité, maintenance. Même s'il est gratuit au départ, l'open source nécessite en général l'aide d'un spécialiste en informatique pour installer et utiliser le logiciel (Mukherjee,



2010), ce qui a finalement un coût. Un logiciel open source ne coûte donc pas forcément moins cher qu'un logiciel propriétaire. C'est un enseignement fondamental qui va à l'encontre des idées reçues. L'éditeur du logiciel open source se paie sur le service apporté, la formation, la maintenance, etc. Par exemple, le Pr Waddell met au profit sa propre société, Urban Analytics, pour promouvoir son logiciel open source OPUS/Urbansim et vendre son expertise, ce qui est tout à fait normal aux Etats-Unis. Enfin, les éditeurs de logiciels propriétaires affirment que verrouiller le code source permet de le rendre plus fiable et plus sécurisé, contrairement au code de l'open source qui est modifiable par n'importe qui.

Pritchard et. Al en 2007 ont examiné différentes offres de logiciels libres dans le domaine de la modélisation en transport, avec l'objectif d'évaluer leur maturité en comparaison des logiciels propriétaires alternatifs (voir [www.sourceforge.net](http://www.sourceforge.net)). La revue a montré que sur des dizaines de projets open source recensés, 60% n'ont pas passé le stade de l'idée. Sur les 40% restants ayant abouti à la livraison d'un produit, un tiers a été abandonné. Pour les logiciels qui sont finalement sortis du lot, certains présentaient des fonctionnalités originales, introuvables sur les logiciels propriétaires comparables. Mais au final, leur diffusion dans la communauté des utilisateurs est restée marginale. Par exemple, le marché des logiciels de modélisation de trafic classique est toujours occupé à 90% par quatre logiciels propriétaires dans le monde. En résumé, l'open source permet aux développeurs une plus grande créativité mais ne répond pas nécessairement aux besoins fondamentaux des utilisateurs professionnels.

Prenons de nouveau l'exemple d'OPUS/Urbansim. Voici **quelques messages postés** sur la liste de diffusion d'Urbansim en avril 2011, qui illustrent parfaitement les problèmes rencontrés dans le monde de l'open source et qui montrent que l'on est très loin des attentes d'un logiciel land use prêt à utiliser et convivial à l'instar de ce que l'on trouve en modélisation de trafic chez des logiciels propriétaires du type Davisum ou Cube Voyager. Un logiciel développé en open source (par exemple en Java/Python/MySQL) par des universitaires comporte davantage de bogs qu'un logiciel développé par un éditeur informatique (par exemple en .NET). De notre expérience de l'utilisation d'Urbansim au cours du projet *SIMAU*RIF, nous évaluons à 50% le temps consacré au débogage des programmes d'Urbansim. Cette tâche est coûteuse et improductive. Le temps consacré au débogage d'un programme doit donc être minimum et ce temps est essentiel dans le choix de l'environnement informatique en modélisation LUTI.

**Question 1** : I am using the python script provided on the UrbanSim website to create zone database. I figured this would be a good way to get down and dirty into the details of creating the databases required for UrbanSim models. The database program

being used is PostgreSQL 8.2. I selected the proper database using the following uncommented portion of the script (see below).

```
# Select the database connection by (un)commenting and adapting the following options
#####
#metadata.bind = "sqlite:///Users/pwaddell/sqlite/sample_zone.db"
metadata.bind = "postgres://urbansim:urbansim1@localhost/zone_database"
#metadata.bind = "mysql://account:password@localhost/sample_zone"
#####
```

When I run the script this is the error I receive (see below). I check the path referenced in the script of the missing module (C:\Python26\lib\site-packages\sqlalchemy\dialects\postgresql\psycopg2.py) and was able to find the psycopg2.py. So I'm not really sure what the problem is or how to go about fixing it. I am new to it all, UrbanSim, Python, and PostgreSQL 8.2, so you will need to keep that in mind when responding. I'm sure there will be more questions to come. Thanks for the help!

```
IDLE 2.6.2   ==== No Subprocess ====
>>>
Traceback (most recent call last):
  File "C:\Documents and Settings\moreland_\Desktop\create_zone_database.py", line 6, in <module>
    metadata.bind = "postgres://urbansim:urbansim1@localhost/zone_database"
  File "C:\Python26\lib\site-packages\sqlalchemy\schema.py", line 1968, in _bind_to
self._bind = create_engine(bind)
  File "C:\Python26\lib\site-packages\sqlalchemy\engine\__init__.py", line 244, in create_engine
    return strategy.create(*args, **kwargs)
  File "C:\Python26\lib\site-packages\sqlalchemy\engine\strategies.py", line 60, in create
    dbapi = dialect_cls.dbapi(**dbapi_args)
  File "C:\Python26\lib\site-packages\sqlalchemy\dialects\postgresql\psycopg2.py", line 234, in dbapi
    psycopg = __import__('psycopg2')
ImportError: No module named psycopg2
>>>
```

**Answer 1 :** "The script needs to be edited. In the commented block there are three options for what kind of database to connect to. The one that is uncommented now is Postgres. This would be appropriate if you have a Postgres database server installed, and if you change the account name and password to match what you use for that. For testing purposes, I suggest you try the sqlite option (uncomment that line and comment out the postgres line), since sqlite is automatically installed. You can get a free sqlite database browser/editor to inspect the data. Later, you probably will want to upgrade to a more powerful (still free) database platform like Postgres or MySQL."

#### Question 2 :

Let me describe a couple of problems with the release 4.3.1 installed on Windows 7, 32 bit :

1. Export to mysql and import from mysql doesn't work:

Error importing opus\_gui.data\_manager.run.tools.opus\_data\_to\_sql

If I type in a Python prompt

```
>>> import MySQLdb
```

or

```
>>> import _mysql
```

there is an import error that there is no module named like this. I'm using Python 2.6, MySQL server 5.5 and MySQL workbench 5.2 CE. Following the installation recommendations, I reinstalled the package

<http://www.technicalbard.com/files/MySQL-python-1.2.2.win32-py2.6.exe>

but it didn't fix the problem. Note that I work with Windows 7, not with XP.

2. Variables specified as Python class (for whom Python scripts exist) are not recognized in 4.3.1, the message is Opus variable ... does not exist for dataset 'urbansim....' in year ... while in fact the data and the scripts exist; I was working with these variables in 4.3.0. It's possible to respecify some variables as expressions in a dialog box, but sometimes Python scripts are necessary.

**Answer 2-1** : the opus\_data\_to\_sql tool, I recall an extremely similar issue came up just over a year ago on this list. If I remember correctly, MySQLdb was not installing correctly, despite appearing as if it did. I solved the problem by locating another binary installer. Try the thread near the bottom of the page with labeled with (no subject):

<http://mailman-mail5.webfaction.com/pipermail/urbansim-users/2010-March/thread.html>

In particular, this post: <http://mailman-mail5.webfaction.com/pipermail/urbansim-users/2010-March/000131.html>

**Answer 2-2** : Regarding your second question, I am not aware of any python variable class breakage in 4.3.1. You may do these two checks: do you have a \_\_ini\_\_.py file in the variable folder? In a python command line shell, try the following: import urbansim.your\_dataset.your\_variable

**Answer 2-3** : My first problem is solved by reinstalling MySQL (indeed, it was not installed correctly). No problem with export and import MySQL<->OPUS. Concerning my second question, Python variable classes work correctly in 4.3.1. I was confused between urbansim and urbansim\_zone folders. Currently, I work with the latter with the zone version, but was trying to use the older path with the former, sorry.

De toute façon, l'argument avancé par les adeptes de l'open source selon lequel il faut libérer les codes sources des programmes se heurtera toujours à l'autre versant de la modélisation LUTI : **celui des données**. La mise à disposition gratuite des algorithmes dans le milieu Open Source facilite la reproductibilité des méthodes des chercheurs, mais pas la reproductibilité des résultats. En effet, dans le milieu de la recherche en transports, chacun ne dispose pas des mêmes données que son homologue à l'autre bout de l'Europe, et surtout pas de micro-données ou données désagrégées indispensables en modélisation LUTI. Même s'il y a un mouvement actuel vers l'accès gratuit à certaines bases de données communales (voir paragraphe 5.5.1 sur le mouvement open data), il n'est pas encore venu le jour où des bases de données précieuses tel que le MOS numérique de l'IAU seront distribuées gratuitement. Toute production de bases de données nécessite un investissement, une mise à jour, un stockage, qui ne sont pas gratuits. Mais la question est posée depuis toujours : jusqu'où va le droit de regard du producteur sur l'utilisation des bases de données financées en grande partie sur fonds publics ? Elle rejoint la question de la **propriété intellectuelle d'un logiciel** qui serait développé avec des subventions publiques. Le choix de l'open source signifierait que le code d'ULTISIM ne serait plus du tout contrôlable par les financeurs.

En conclusion, le **choix du logiciel propriétaire** pour le développement d'ULTISIM s'impose pour des raisons de coût, d'efficacité, de marché commercial, de longévité.

## 5. Données nécessaires

### 5.1. La question des données

Le point commun aux différentes catégories de modèles intégrés concerne la question des données nécessaires à leur développement. Les modèles LUTI sont très **exigeants en données**. C'est une question cruciale qui est en général très mal décrite dans les rapports de recherche et dans la littérature scientifique qui privilégient la présentation des équations puis les résultats. Il est vrai que l'identification des bases de données, la collecte, le contrôle de l'intégrité, la complétude et la mise en forme sont des tâches fastidieuses, ingrates et peu reconnues dans le milieu de la recherche. Or l'adage anglo-saxon « **Garbage in, Garbage out** » est particulièrement évocateur lorsque l'on sait que la modélisation LUTI nécessite un volume de données gigantesque, aussi bien en entrée pour le faire tourner qu'en phase de développement pour estimer économétriquement les différents sous-modèles. Ce thème des données comporte de **nombreuses questions** aussi bien théoriques que pratiques :

- Quelles sont les sources de données disponibles sur l'aire d'étude et sur une longue période passée ?
- Qui sont les producteurs de données ?
- Quel est le coût de production et de collecte des données ?
- Comment faire en cas de données manquantes ?
- Quel est le niveau d'incertitude des données ?
- Quelle est l'échelle pertinente (micro/méso/macro) ?
- Quels sont les outils informatiques pour traiter, vérifier l'intégrité et mettre en forme les données une fois la collecte réalisée ?
- Comment pérenniser les bases de données ?

Il faut donc revaloriser la tâche sur les données qui devient stratégique, reconnaître la valeur et l'intérêt du savoir et de l'expertise sur les bases de données, et parler désormais **d'ingénierie des données**. La situation d'aujourd'hui est très différente d'il y a une quinzaine d'années à cause des progrès de l'informatique, tant au niveau de la puissance des micro-processeurs (qui se mesure en 2011 en centaines de milliers de milliards d'instructions par seconde) que des logiciels tels que les systèmes d'information géographique (SIG). Hier encore, la limite de puissance des ordinateurs de bureau imposait une contrainte dans la collecte des données et leur traitement. Aujourd'hui, cette contrainte n'existe plus et on peut concevoir un projet de modélisation LUTI dynamique désagrégée et de microsimulation sans se soucier des temps de calculs. Ainsi, il est tout à fait envisageable

aujourd'hui d'opter pour un découpage d'ULTISIM sur la région d'Île-de-France (12 000 km<sup>2</sup>) de quelques milliers de zones (par exemple le zonage IRIS) alors qu'il y a à peine 15 ans, parler de quelques centaines de zones pour un modèle de trafic classique était considéré comme très audacieux.

La question des données nécessaires dépend aussi largement du **contexte de l'aire d'étude**. Chaque étude de cas est un cas particulier, et nous recommandons d'identifier clairement les spécificités de chaque aire d'étude avant de se lancer dans la modélisation. Par exemple, en Île-de-France, une hypothèse fondamentale dans l'utilisation des modèles d'équilibre général est l'équilibre entre les prix de l'immobilier résidentiel, les localisations et les coûts de transport, conformément à la théorie monocentrique. Mais en Île-de-France, le marché immobilier de l'immobilier d'entreprise fonctionne-t-il comme le marché résidentiel? Les études de l'IAU ont pu montrer que le marché de l'immobilier d'entreprise présente des particularités : une suroffre de bureaux alors que la demande n'est pas satisfaite (est-ce dû à l'obsolescence d'une partie des bureaux tertiaires actuels ?), un marché très dépendant de la conjoncture économique (mais jusqu'à quel degré ?)<sup>10</sup>, des prix affichés fluctuants sur le court terme (qu'en est-il des prix finals négociés en fonction de la durée des baux ?), une faible attractivité des zones d'emplois tertiaires autour des deux aéroports internationaux (Roissy et Orly) contrairement aux autres pays (est-ce principalement dû à la faiblesse des dessertes en TC ?). Un critère qui devrait aussi prendre de plus en plus d'importance à l'avenir concerne le Grenelle 2 sur les contraintes énergétiques des bâtiments.

Une autre réflexion sur la localisation des ménages n'a pas encore été traitée par les modélisateurs : l'influence des flux migratoires étrangers sur la dynamique du marché du logement. Quelles données sont nécessaires pour sa prise en compte dans un modèle de choix de localisation résidentielle ? C'est ce genre d'analyse spécifique à chaque aire d'étude qui doit être préalable à toute collecte de données.

---

<sup>10</sup> Une étude récente publiée en juin 2011 du cabinet Deloitte recense toutes les opérations neuves et les restructurations de bureaux de plus de 1000 m<sup>2</sup>, en chantier et livrées « en blanc », c'est-à-dire sans utilisateur ou locataire au moment du permis, sur Paris et la petite couronne. Elle montre que la chute de production de bureaux neufs se poursuit depuis 2008, et que le total des chantiers en cours, moins de 780000 m<sup>2</sup>, est en effet à son plus bas niveau depuis huit ans. Dans ce contexte de panne générale, les Hauts-de-Seine cumulent 80% des chantiers de bureaux engagés sur Paris et la petite couronne. La capitale concentre à elle seule cinq des neuf opérations recensées par l'étude, mais avec un seul chantier nouveau lancé entre octobre 2010 et mars 2011. L'est de Paris est quasiment à l'arrêt : seulement deux nouveaux ont démarré dans le Val-de-Marne. C'est donc dans les Hauts-de-Seine que la construction de bureaux a été la plus dynamique, avec cinq chantiers ouverts, représentant 82% des surfaces mises en chantier sur Paris et la première couronne. Le quartier de La Défense reste assez actif, avec trois tours en construction.

## 5.2. *Éléments de réponse*

Le projet *SIMAU*RIF entre 2003 et 2008 a permis de réaliser un travail approfondi et pionnier sur ces questions en Île-de-France et nous renvoyons le lecteur aux quatre rapports d'étude (Nguyen-Luong et de Palma, 2002 à 2008, B) et au rapport sur l'impact du tramway T3 sur les prix de l'immobilier (Boucq, Nguyen-Luong et Papon, 2011). Rappelons que la construction des bases de données du projet *SIMAU*RIF a nécessité deux années de préparation avec :

- d'une part le **SIGR**. Toutes les données basées sur le référentiel spatial MOS (Mode d'Occupation du Sol) ont été projetées sur le maillage en 50 000 cellules et concernent :
  - les logements et localisation des ménages
  - les emplois et leurs localisations
  - les activités (bureaux, entrepôts, commerces) et leurs localisations
  - les valeurs foncières et immobilières
  - la présence d'aménités et leurs proximités aux données désagrégées
  - les données de réseaux (routier, transport en commun)
  - les contraintes environnementales (zones inondables, etc)
  - les contraintes de politiques urbaines (SDRIF, PLU, etc)
  - les matrices origine-destination de temps de parcours
  - les projets de transports et leur calendrier de réalisation
  
- d'autre part le logiciel **SAS** pour les données désagrégées : les RGP de 1990 et 1999, les Enquêtes Globales Transport de 1991-1992 et de 2001-2002, l'ERE 1997 et 2001, les données de transactions de la base BIENS des Notaires qu'il faut acheter. Par exemple, un géocodage automatique de chacun des 5 millions de ménages et des 4,5 millions d'emplois dans le maillage en 50 000 cellules a été effectué.

D'une manière générale, il faut privilégier l'utilisation de données désagrégées et géolocalisées quand elles sont disponibles car cela améliore significativement la qualité des paramètres estimés.

Remarquons qu'il existe une autre catégorie de données assez peu prises en compte. Ce sont celles qui concernent la **règlementation fiscale**. La fiscalité étant variable en fonction de la localisation, il serait intéressant de construire une base de données décrivant les niveaux de taxe d'habitation, taxe foncière, cotisation foncière des entreprises, contribution économique territoriale (CET) et allègements d'impôts sur les zones spécifiques (par exemple les zones franches urbaines ZFU et les zones de redynamisation urbaines ZRU

offrent des exonérations d'impôts aux entreprises, ce qui est un facteur d'attractivité pour certaines activités).

En résumé, les réponses aux questions du paragraphe 5.1 dépendent de quatre facteurs.

En premier lieu, elles dépendent de la **catégorie de modèle intégré**. Entre un modèle relativement simple d'équilibre général et un modèle dynamique de choix discrets fonctionnant en micro-simulation, les efforts pour construire les bases de données ne sont pas du tout comparables. Les modèles de choix discrets nécessitent de recueillir les données désagrégées et de les localiser à un niveau spatial suffisamment fin. L'expérience montre que la construction d'une base de données d'estimation d'un modèle de choix discret requiert 90% du temps contre 10% pour la partie économétrique. Or cette étape préalable est rarement décrite de manière détaillée et transparente dans les rapports de recherche.

En second lieu, elles dépendent des **résultats attendus**. Par exemple, si on envisage que le modèle intégré comporte un sous-modèle de prix, il y a nécessairement des données de transactions à acquérir avec un certain nombre d'attributs dits « intrinsèques » aux biens (caractéristiques du logement). Si on prend la base de données des Notaires, quelles sont les variables disponibles et quelles sont les autres variables à construire pour prendre en compte les caractéristiques « extrinsèques » d'un bien immobilier (par exemple les distances de chaque logement aux aménités) ? Ce qui est important de retenir, c'est que d'une part les données doivent servir à rendre les modèles applicables **en prédiction** et non à construire des modèles **explicatifs**, d'autre part il faut de la cohérence entre les données utilisées dans les différents sous-modèles. Donc il faudra se soucier d'une certaine parcimonie dans le choix des variables et ne pas oublier l'objectif d'intégration de tous les sous-modèles si l'on vise un objectif d'opérationnalité en mode routine.

Ensuite, la collecte et le traitement des données dépendent de la  **finesse**  du découpage géographique du modèle. Si le découpage est agrégé (communes ou regroupement de communes), la question des données est plus simple à traiter que si le découpage est infracommunal comme c'était le cas dans le projet *SIMAU RIF* (découpage de l'Île-de-France en 50 000 cellules carrées).

Enfin, une fois les choix précédents effectués, il y a encore la question de l'identification des bases de données existantes adéquates. C'est là que les problèmes commencent réellement. Tout modélisateur a déjà rencontré **le problème de l'accès aux données qui sont dispersées et du coût d'acquisition**. Faute de moyens, il arrive souvent qu'un

modélisateur recueille des données incomplètes, biaisées et plus ou moins « bricolées » mais il ne veut pas savoir quels sont ces problèmes car il a déjà suffisamment de soucis avec son travail de modélisation, il fait donc « avec » et puis il sait que de toutes façons, dans le milieu de la recherche, seuls les résultats qu'il produira seront présentés et valorisés.

Une méthode pour traiter le problème des données manquantes est la méthode de fusion des données qui permet de lier des bases de données individuelles de différentes sources. Il existe ainsi des procédures statistiques permettant de faire hériter les attributs d'une base aux observations d'une autre base (Morency, 2008). Par exemple, il est possible d'exploiter conjointement les données provenant des enquêtes ménages déplacements et des recensements.

### **5.3. Comparaison du cas Île-de-France avec deux cas étrangers**

#### **5.3.1. Pratiques de modélisation**

Quelques mots sur les pratiques de modélisation LUTI menées dans l'aire urbaine de Turin en Italie et la Randstad aux Pays-Bas. A Turin, le laboratoire Laq-TIP de l'Ecole Polytechnique de Turin est en train de développer un modèle intégré en utilisant le logiciel OPUS/Urbansim (Caneparo *et al*, 2009). De plus, le Laq-TIP a développé un Système d'analyse multi-dimensionnelle et de modélisation générative de la dynamique urbaine et régionale. Concrètement, le laboratoire a connecté son système de réalité virtuelle à OPUS/Urbansim pour permettre la visualisation en 3D des résultats issus d'OPUS/UrbanSim.

La question de la visualisation des résultats sortis d'un modèle LUTI est importante pour communiquer et diffuser les résultats, mais reste néanmoins secondaire par rapport aux données d'entrée. Il est clair aujourd'hui que la 3D sera la prochaine étape des post-traitements. Elle nécessitera la généralisation de nouvelles méthodes d'acquisition, de traitement des données tridimensionnelles (photogrammétrie, modèle numérique de surface, réalité « augmentée ») et de développement de logiciels pour non-spécialistes.

De leur côté, les Pays-Bas sont assez avancés dans la recherche et le développement des modèles intégrés. Le modèle intégré utilisé est le modèle TIGRIS XL développé par le bureau Significance (voir paragraphe 2.3.7.). L'Université de Delft a développé un autre modèle de micro-simulation à l'état de prototype, appelé SFM (Spatial Firm-demographic



Micro-simulation), qui permet d'estimer les effets d'agglomération. Il a été appliqué à la région du sud des Pays-Bas (De Bok, 2009).

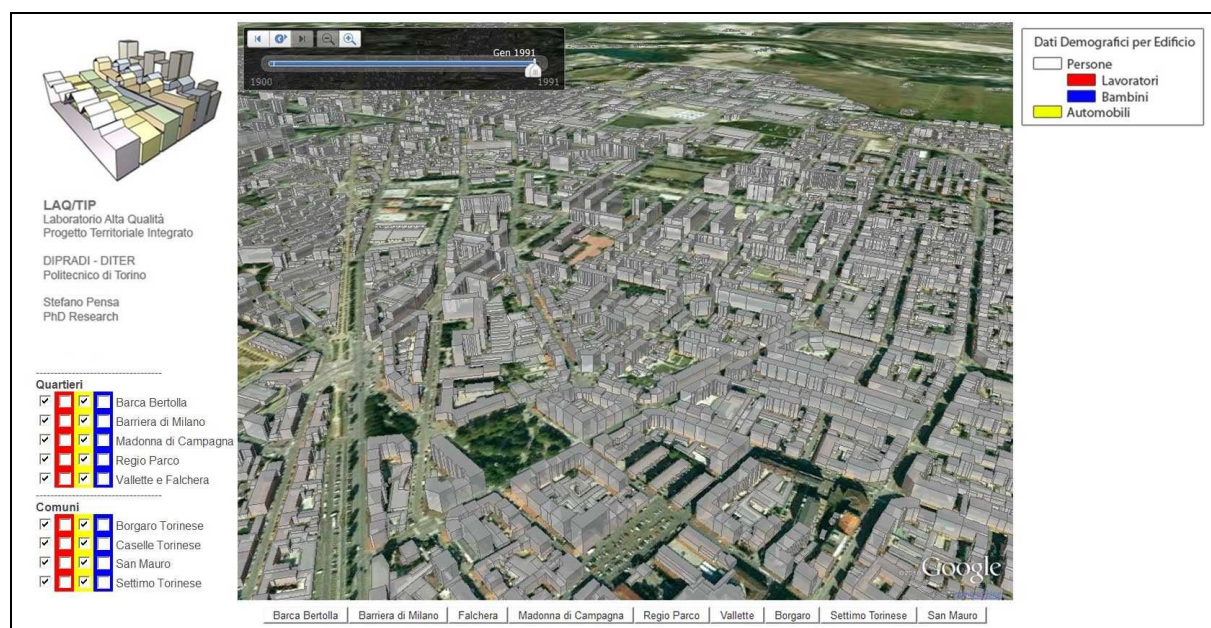


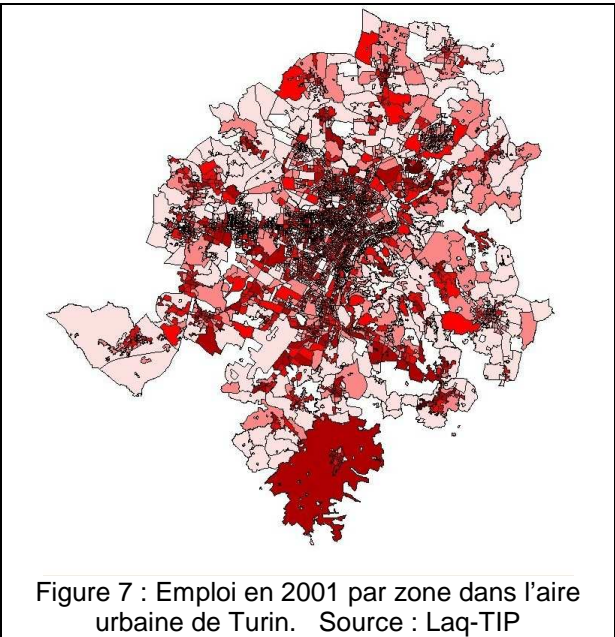
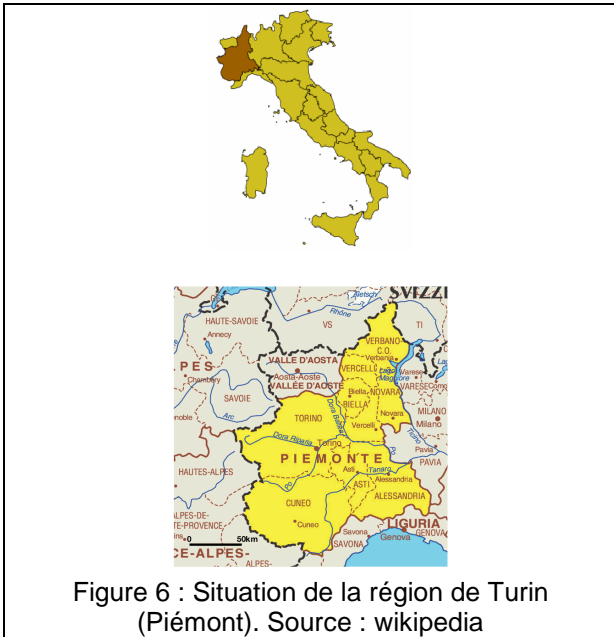
Figure 5 : Extrait d'une sortie 3D exportée depuis le système du Laq-TIP vers Google Earth

### 5.3.2. Chiffres clés des aires d'étude

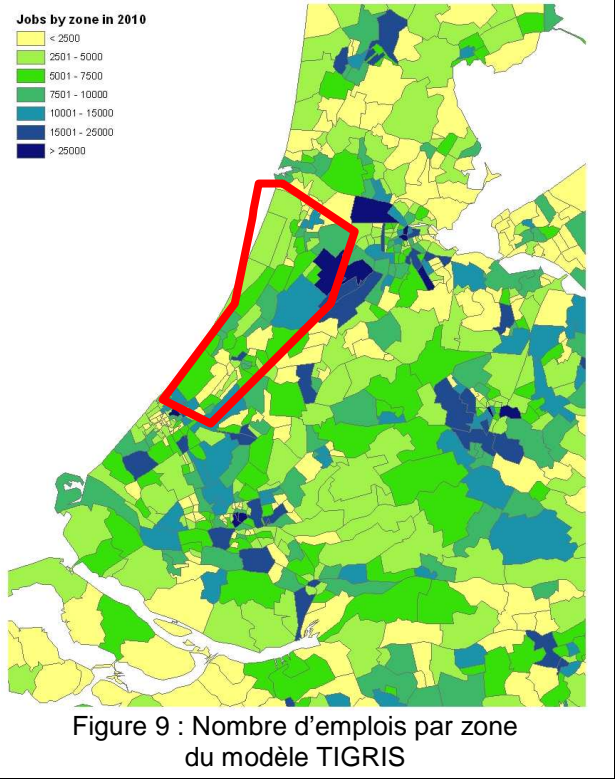
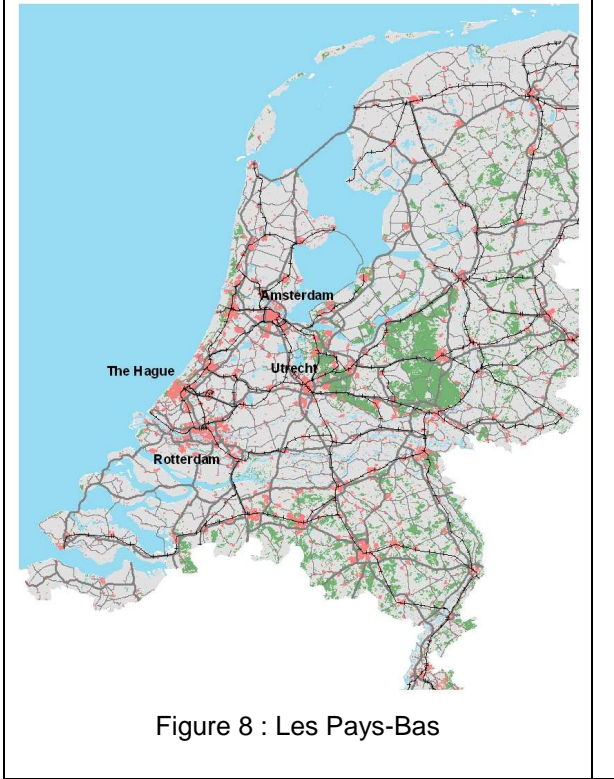
	IAU îdF	Laq-TIP	Significance
<b>Aire d'étude</b>	Île-de-France	Aire urbaine de Turin	Randstad
<b>Modèle</b>	SIMAU <sup>RIF</sup> basé sur UrbanSim	Application d'UrbanSim	TIGRIS XL
<b>Population</b>	11,5 millions	2,2 millions	7,2 millions
<b>Emplois</b>	5 millions	0,7 million	3,8 millions
<b>Surface (km<sup>2</sup>)</b>	12 000	6 821	5 820

Tableau 1

**Turin** est la capitale du Piémont qui est un centre de production important de l'Italie septentrionale. Les activités économiques sont très nombreuses : automobiles et sous-traitance, construction ferroviaire, assurances, banques, transports routiers, industrie agroalimentaire et dérivés, ce qui en fait un des poumons économiques de l'Italie. L'aire urbaine de Turin comprend la ville de Turin et 34 communes. Elle compte près de 2,2 millions d'habitants et 0,7 millions d'emplois sur 6821 km<sup>2</sup>.



La **Randstad** est le nom de la région où sont situées les quatre principales villes hollandaises : Amsterdam (aéroport de Schiphol), Rotterdam (le plus grand port du monde), La Haye (siège du gouvernement) et Utrecht (noeud ferroviaire national).



C'est l'une des plus importantes régions économiques européennes, où se concentrent 46,2 % de la population néerlandaise soit 7,2 millions d'habitants sur une population totale de 16,7 millions, et 3,8 millions d'emplois sur 5420 km<sup>2</sup>.

### 5.3.3. Tableau synthétique de comparaison

	<b>IAU îdF</b>	<b>Laq-TIP</b>	<b>Significane</b>
<b>Modèle de trafic</b>	Modèle propre de l'IAU sous Davisum en 606 zones Modèle ANTONIN du STIF	Modèle de l'Agence de Mobilité de l'Aire urbaine en 267 zones	Modèle du Centre de recherche en Transport sous Cube (TRC), en 1327 zones sur tout le pays
<b>Autres zonages</b>	- Communes (1300) - IRIS (5100) - Cellules carrées de 500 m (50 000)	Communes (35)	- Communes (443) sur tout le pays - Cellules carrées de 100 m
<b>Enquête ménages déplacements</b>	EGT 2001 (source IAU)	Household Travel survey (source Agence de Mobilité de l'Aire urbaine)	National Travel Survey
<b>Population</b>	RGP 1990 et 1999 (source INSEE)	Recensement national de la Population 1991 et 2001 (source ISTAT / C.S.I. Piemonte National Statistical Institute)	- BD nationale à la commune (source : CBS) - Prévisions (CBS et Agence nationale de l'évaluation environnementale NEAA)
<b>Emplois</b>	- RGP 1990 et 1999 - Base Astree (source IAU)	Recensement national des activités 1991 et 2001 (source ISTAT)	- BD diverses (données nationales LISA, CBS, enquête emplois EBB, Eurostat) - Prévisions (Bureau d'analyses économiques CPB)
<b>Immobilier</b>	Base BIENS à la commune (Source Chambre des Notaires)	OICT (Observatoire de l'immobilier de Turin)	NR
<b>Foncier</b>	NR	FIMAA (Association nationale des professionnels de l'immobilier et du foncier)	NR
<b>Occupation du sol</b>	MOS 1990 et 1999 en 83 postes (source IAU) RGP pour le nombre de logements SIGR pour les commerces (source IAU) SIGR pour les contraintes environnementales	Land use de Turin 1990 et 2000 (source : Ville de Turin) Corine Land Cover	- BD nationale en 38 postes - Données CBS pour le nombre de logements - BD sur le commerce (source IBIS et BAK) - BD sur les contraintes environnementales (source : Ecological network, Natura2000)

<b>Modèle de localisation des ménages</b>	EGT 2001 (source IAU) qui comprend la donnée de revenu	- Registry offices - RNP 1991 et 2001 (source ISTAT) - Revenu (National Statistical Institute)	- Données CBS - BD sur le logement (source : WBO, tous les 4 ans, Ministère du Logement) - Revenu (source : Experian)
<b>Modèle de localisation des emplois</b>	- ERE 1997 et 2001, CLAP 2006 (source IAU) - RGP et fichier DGI (Direction générale des Impôts)	- Registre des entreprises (Source : Chambre de commerce) - Recensement général des activités 1991 et 2001 (source ISTAT)	NR
<b>Modèle de prix de l'immobilier</b>	Base BIENS géolocalisée (Source Chambre des Notaires) OLAP (Observatoire des Loyers de l'Agglomération Parisienne)	NR	SO (car prix endogènes) Pas de données individuelles
<b>Modèle de développement urbain (version Urbansim 2)</b>	EVOLUMOS (source : IAU)	NR	SO
<b>Modèle de localisation de projets (version UrbanSim 4)</b>	Données géolocalisées de ZAC et de ZAE (source IAU), Base de données Projets urbains et immo_construction (source IAU)	NR	SO

Tableau 2

NR : non renseigné en juin 2011  
SO : Sans objet

En Île-de-France (projet SIMAURIF) et à Lyon (projet SimBad), le modèle de localisation des emplois a fait l'objet de recherches moins avancées que le modèle de localisation des ménages qui pose davantage de problèmes. Une des bases de données très importante pour estimer ce modèle est l'**Enquête Régionale Emplois (ERE)** géolocalisée, devenue CLAP en 2006. C'est une enquête réalisée par l'INSEE tous les 5 ans en moyenne depuis 20 ans, auprès des entreprises, tandis que le RGP recense les lieux de travail des individus. CLAP ne recense pas les emplois non-salariés (indépendant, libéraux), mais il a l'avantage d'être géolocalisé. Il peut y avoir des effets de déclaration au siège social même si l'enquête concerne les établissements et non les sièges. Pour les grandes entreprises (RATP, SNCF, ADP, Air France, etc), cet effet a été corrigé par l'IAU mais pas pour les PME de services à la personne ou aux entreprises (par exemple, les entreprises de nettoyage).

Cette base de données permet de calibrer le modèle de choix de localisation des emplois à une échelle infracommunale.

#### **5.4. Données européennes**

**Une des orientations données à ULTISIM est sa capacité à être générique.** Pour cela, nous avons voulu savoir s'il existait des bases de données européennes standard, accessibles facilement, gratuites si possible, permettant en mode « RAD » (Rapid Application Development) de développer une application en quelques mois une fois que l'on disposerait du logiciel ULTISIM.

D'abord, comme base d'occupation du sol, on accède facilement à la base **Corine Land Cover** qui couvre l'Europe entière pour les années 1990, 2000 et 2006. Il y a 25 postes d'occupation du sol mais ils sont très spécifiques au domaine de l'environnement. Par exemple, la base distingue les sites d'extraction de minéraux, les champs de riz, les forêts de conifères, les forêts mixtes, etc. A contrario, le MOS (Mode d'Occupation du Sol) de l'IAU IAU comporte 83 postes différents et il y a un bon équilibre entre les postes décrivant le type d'habitat, les activités, les transports, les écoles, les équipements, les commerces et les espaces naturels. On peut obtenir en ligne des extraits de la base Corine Land cover en format shape. Le système de coordonnées n'est pas compatible directement avec le système Lambert 93 utilisé en France. En conclusion, la base Corine Land Cover ne sera pas suffisante pour répondre aux besoins d'ULTISIM. Par ailleurs, la Commission européenne a essayé de mettre en place un SIG à l'échelle infra-régional appelé SIRE mais ce projet en est resté à l'état de maquette.

Ensuite, en ce qui concerne les données démographiques et socio-économiques, il existe trois bases européennes (Eurostat, 2008) :

- **Eurostat NUTS** (Nomenclature des Unités Territoriales Statistiques): les données sont disponibles seulement de manière agrégée au niveau régional, et pour certaines régions européennes, à un niveau équivalent au département français ;  
[http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts\\_nomenclature/introduction](http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction)
- **Eurostat micro-données PCM** (Panel Communautaire de Ménages) : cette base comprend des données sur des unités statistiques individuelles : revenus, situation financière au sens large, vie professionnelle, conditions de logement, relations sociales,

santé et biographie. Le PCM est une enquête par panel (un échantillon de ménages et de personnes ont été interrogés chaque année) qui commence à dater. Elle a eu lieu pendant huit ans, de 1994 à 2001. Le coût est quasiment gratuit pour les laboratoires de recherche (250 €) ;

- **Urban Audit database** : c'est une base de données qui date de 2003-2004 sur 58 grandes villes (Paris et Londres n'en font pas partie) mise en place par la Commission européenne en 1999. Elle comporte pour chaque ville environ 480 variables. Aucune extraction n'est possible depuis internet. Un réseau de « correspondant données » dans chaque ville avait été mis en place mais ce système a périclité. La base n'a pas été mise à jour depuis quelques années.

**En résumé, il ne faut pas trop attendre du côté des bases de données européennes.**

Elles ne sont pas assez détaillées spatialement, pas assez représentatives pour une aire d'étude donnée, pas adaptées à la micro-simulation, ne sont pas mises à jour et ne sont pas accessibles facilement par internet. La seule solution pour construire les bases de données et d'estimation pour une application d'ULTISIM est le recueil de données existantes localement. Il faudra donc penser à ne pas verrouiller la structure de données d'ULTISIM afin de permettre la prise en compte des données au cas par cas.

## **5.5. Nouvelles sources de données**

### **5.5.1. Mouvement « Opendata »**

Le développement des technologies de l'information (Internet, Smartphone, 3G+, etc.) a conduit à un développement tous azimuts de l'offre d'information sur les transports et les déplacements. Les prestataires (sociétés de services, exploitants, internautes, etc) sont confrontés à l'accès aux données de base. Depuis un an, des grandes villes ont décidé d'ouvrir leurs bases de données gratuitement et de les mettre en ligne. Après Rennes et Paris, voici que Bordeaux, Brest, Marseille, Montpellier et Nantes vont bientôt entrer dans **l'ère de l'« open data »**. L'objectif commun est d'optimiser l'utilisation de leurs services, en particulier les transports, et inventer de nouveaux usages. Les prestataires qui attendaient cette « libération » des données depuis longtemps peuvent désormais développer les applications interactives. Les applications les plus attendues concernent l'aide à la mobilité : recherche d'une place de stationnement, calcul d'itinéraires pour des déplacements intermodaux, etc. Pour les modèles LUTI, ce sont les microdonnées urbaines qui sont les

plus intéressantes : lieu et nombre de places de stationnement, adresse des équipements scolaires, culturels, de loisirs, etc. En fait, ce sont les données comparables à celles qui sont stockées dans le SIGR de l'IAU pour la région Île-de-France, dont la richesse des données constitue sans doute une exception et donc n'est pas reproductible.

A l'avant-garde du mouvement open data en France, l'**agglomération de Rennes** a lancé un concours de services innovants et recueilli 43 applications Web et smartphone en projet ou déjà en oeuvre. Les lauréats doivent s'engager à offrir un service gratuit durant un an. Après, le service pourra être payant, ce qui risque de relancer le débat sur la marchandisation des données publiques. Une des applications à Rennes a été réalisée par un citoyen féru d'informatique. Elle propose des itinéraires à la demande aux personnes handicapées ([www.handimap.org](http://www.handimap.org)). Le système, couplé à un moteur de recherche, empile et compile plusieurs couches de données : arrêts de bus accessibles, places de stationnement réservées aux personnes à mobilité réduite, surfaces « podotactiles » sur les trottoirs, etc. Il suffit à l'utilisateur géolocalisé de taper sa destination sur son smartphone pour voir s'afficher la route à suivre sur Google Maps.

Le mouvement n'a pas échappé au Gouvernement français, qui a lancé en février 2011 une mission appelée **Etalab**, chargée de mettre en place un portail Web permettant aux citoyens de consulter en toute transparence l'ensemble des informations publiques de l'Etat et de l'administration. Rappelons que la loi Cada de 1978 garantit l'accès aux documents administratifs, et une directive européenne de 2005 renforce cette obligation, fort opportunément servie par la révolution numérique.

Le mouvement opendata se répand aussi aux Pays-Bas. Voir [www.opendatanederland.nl](http://www.opendatanederland.nl). Ce site ne fonctionne pas encore très bien. Il semble qu'une masse importante d'informations très détaillées y soit stockée, mais de manière encore fragmentaire, et certaines livrées dans un format brut.

De même, l'Ecole Polytechnique de Turin participe à plusieurs projets relevant du mouvement Open Data. Voir <http://nexa.polito.it/category/topic/open-data> et le site de la Région piémontaise, <http://www.dati.piemonte.it>, qui mentionne le portail français de la mission Etalab [www.data.gouv.fr](http://www.data.gouv.fr).

Un point crucial concernant ce mouvement "open data" est l'indispensable respect des **formats standards** pour que les données puissent être exploitables par les prestataires et les éditeurs de logiciels développant des applications liées aux transports. Par exemple, il existe un format européen normalisé d'échange de données appelé **Trident**. Ce format a été

utilisé par la RATP pour son système d'information sur les situations perturbées SIPRE, et également pour les échanges d'informations au niveau francilien entre les systèmes de la RATP et ceux de l'AMIVIF, pour les horaires des transports en commun.

### 5.5.2. Projet Belgrand-GEBD

Le projet Belgrand est un projet de Grand Equipement de Base de données (GEBD) pour la ville et la mobilité. Mis en place début 2011, il est porté par le Pôle de recherche et d'enseignement supérieur (Pres) Université Paris-Est, coordonné par l'IFSTTAR, et réunit le lab'Urba, l'IGN, l'ENSG et l'Ecole des Ponts ParisTech. L'objectif est de créer un **portail Web collaboratif** à l'horizon 2015 qui connectera entre eux les outils et ressources utiles pour exploiter et mutualiser diverses bases de données existantes (logiciels, référentiels, catalogues, métadonnées ...) ainsi que les nouvelles formes de données issues de capteurs GPS/GSM. Dans une première phase financée par le MEEDDM, le projet Belgrand se présente comme un réseau de laboratoires (le réseau GEBD) et un lieu de rencontres et d'échanges sur les bases de données pour la ville durable.

Pour en savoir plus : <http://belgrand-gebd.ifsttar.fr>

### 5.5.3. Projet CASD (Centre d'Accès Sécurisé Distant)

Le projet CASD est très similaire au projet Belgrand-GEBD. Il a été sélectionné dans le cadre du premier appel à projets en 2010 du programme Investissements d'avenir sur les Equipements d'excellence (Equipex) du Ministère de la Recherche. Il est porté par le Groupe des Ecoles Nationales d'Economie et Statistique (GENES) / INSEE, le Gis Quetelet (CNRS, EHESS, Universités, Sciences Po), l'ENS Cachan, l'Ecole Polytechnique et HEC. Le projet vise à créer un équipement et une infrastructure hautement sécurisés permettant aux chercheurs d'accéder à des **bases de données individuelles détaillées**. Par exemple, les chercheurs se plaignent de la difficulté d'accéder à des bases de données désagrégées sur le logement (RGP, Enquêtes Ménages Déplacements, Enquêtes Nationales Logement) qui sont indispensables pour mener des études sur les trajectoires résidentielles et de modélisation.

Le projet a reçu une subvention du Ministère de la Recherche de 4 millions d'Euros. Il a donc réussi à se « vendre » comme projet donnant à la recherche française des outils qui jusqu'alors lui faisaient défaut au regard de ses concurrents européens et d'Amérique du nord (Canada, Etats-Unis), en profitant de l'engouement actuel pour le « cloud computing »



(concept informatique qui consiste à déporter sur des serveurs distants des traitements informatiques traditionnellement localisés sur des serveurs locaux ou sur le poste client de l'utilisateur). Notons que le géant Apple vient d'annoncer en juin 2011 la création de son service iCloud.

Que pensez des projets Belgrand et CASD ? On peut se réjouir de voir émerger des projets qui montrent une prise de conscience dans la communauté des chercheurs de **l'importance stratégique des données** en modélisation urbaine et des transports, mainte fois soulignée par les experts de l'IAU. Il faut arrêter de parler d'« achat » ou de « recueil » de données. On est entré dans l'ère de **l'ingénierie des données**.

Cependant, il faut être très pragmatique et veiller à ce que ces deux projets « montagne » n'accouchent pas de souris. Prenons le cas de l'IAU : il est le centre de ressources de données pour l'ensemble des acteurs en Île-de-France et il dispose de tous les outils modernes de traitement de données spatiales. En tant que producteur- fournisseur-utilisateur, l'IAU a une expertise sur les données reconnue dans le milieu de la recherche. Or le SIGR de l'IAU est l'aboutissement de 25 ans de travaux en interne. Peut-on imaginer raisonnablement que le projet Belgrand parvienne en 5 ans à créer un SIGR bis ? Peut-on croire un instant que l'IAU mette à disposition gratuitement l'accès au SIGR via le centre Belgrand et à l'Enquête Globale Transports via le CASD, que l'INSEE en fasse de même avec ses BD désagrégées ? Les co-financeurs de l'EGT se réservent traditionnellement depuis la première EGT en 1969 une exclusivité d'exploitation pendant au moins trois ans suivant la livraison des fichiers, avant de fournir les fichiers bruts aux chercheurs dans le cadre d'études conventionnées. Par ailleurs, l'argument avancé d'inviolabilité du CASD n'est pas nécessairement un atout. D'une part, la valeur du contenu n'est pas comparable à celle de la Banque de France par exemple et donc le niveau de sécurité d'accès à des données urbaines ne sera peut-être pas aussi exigeant. D'autre part, l'histoire montre que les algorithmes de cryptographie comporteront toujours des failles. Pour preuve le piratage en avril 2011 des centres d'hébergement de données de deux grandes entreprises, Sony et Amazon<sup>11</sup>, aux enjeux financiers énormes.

---

<sup>11</sup> Sony a confirmé en mai 2011 que le système PlayStation Network (PSN) reliant les consoles de jeux de salon PS3 à Internet a été piraté. Des données personnelles dont probablement des coordonnées bancaires ont été dérobées. De plus, des données personnelles appartenant à 24,6 millions d'abonnés au SOE, Sony Online Entertainment, auraient été volées. Apparemment, le pirate informatique qui a « hacké » le PSN aurait utilisé les services d'Amazon, Elastic Compute Cloud, un service de cloud computing. Amazon est une des plus grandes entreprises proposant des services de Cloud Computing.

#### 5.5.4. Données issues de capteurs GPS et GSM

En plus des données traditionnelles (recensement général, enquêtes déplacements, comptages automatiques routiers, enquêtes aux gares, dans les bus, etc), il existe des méthodes innovantes de collecte de données individuelles utilisant des **capteurs GPS, GSM** et autres technologies de localisation à courte distance (Wifi, Bluetooth, etc). Ces nouvelles approches permettent de suivre à la trace tous les modes de transport (suivi de tous véhicules et de piétons) et par conséquent de compléter, voire de remplacer, des méthodes plus classiques de comptages et d'enquêtes ménages déplacements qui sont indispensables à la modélisation de trafic et à la modélisation LUTI. Des développements sont encore à mener dans l'interopérabilité des signaux de localisation (par exemple GPS/Wifi). Ces données occupent de plus en plus de volume qu'il faudra bien stocker. Les américains parlent de « big data ».

##### 5.5.4.1. Les capteurs GPS

Les enquêtes de suivi des individus par GPS commencent à se mettre en place, en complément des enquêtes traditionnelles en face-à-face. Certaines enquêtes ont même été conduites intégralement par GPS et sans enquêteur (Cincinatti 2010 – Etats Unis et Sydney 2011 – Australie). Le coût des traceurs GPS passifs a fortement diminué au cours des 3 dernières années et la technologie des microprocesseurs GPS est de plus en plus performante. Il y a juste encore une dernière limitation au niveau de l'autonomie des batteries, mais qui devrait être levée dans les prochaines années. Une recherche Predit GO 3 menée par l'IAU est en cours sur l'élaboration d'une nouvelle méthodologie d'enquête ménages déplacements intégralement réalisée à l'aide d'un GPS, d'un Système d'information géographique et d'un système expert. Les premiers résultats de la recherche montrent que les enquêtes par GPS peuvent tout à fait se passer d'enquêteurs à domicile et donc **remplacer** les enquêtes ménages déplacements traditionnelles si elles sont **complétées** par des enquêtes légères conduites par téléphone et/ou par internet et si des progrès sont réalisés au niveau des traceurs (hybridation signaux GPS/Wifi, intégration d'accéléromètre miniaturisé). Elles ont l'avantage d'apporter des informations nouvelles sur les itinéraires et la multimodalité en semaine, améliorent la précision des attributs fondamentaux que sont les durées et les distances de déplacements et sont moins coûteuses, ce qui permettrait une plus grande fréquence de réalisation (tous les 3 ans contre tous les 9 ans aujourd'hui). L'utilisation de fichiers issus du suivi GPS ou de personnes pose des questions de respect ou protection de la vie privée mais ce n'est pas nouveau.

Voir [www.iau-idf.fr/eqtpargps](http://www.iau-idf.fr/eqtpargps)

#### 5.5.4.2. Les capteurs GSM

Le GSM (Global System for Mobile Communications) est la norme numérique pour la téléphonie mobile. Cette norme utilise la gamme de fréquences des 900 MHz. Un dérivé du GSM, le DCS, Digital Communication System, utilise la gamme des 1800 MHz.

Les deux technologies principales de géolocalisation utilisant le GSM sont :

- **Cell ID** (Identification de cellule) : dès que le portable est dans une zone couverte par le réseau, il se connecte à une antenne relais GSM. C'est à partir de l'identification de cette antenne que l'on peut localiser le portable. Cette localisation est très rapide (moins de 5 secondes) mais elle est peu précise car elle dépend du nombre d'antennes relais et de leur rapprochement. En zone urbaine, la précision varie entre 100 et 700 mètres et en zone rurale, cela peut aller jusqu'à 10 kilomètres, ce qui est largement insuffisant pour les applications envisagées. De plus, cette localisation n'est pas possible dans toute l'Europe, par exemple en Italie.

- La **triangulation** : un croisement des données est effectué à partir de trois antennes relais utilisées. Cette technique nécessite l'installation préalable d'un programme sur la carte SIM du portable. La localisation s'effectue en environ cinq secondes et est plus précise que le Cell-ID : de l'ordre de 150 mètres en zone urbaine et 5 kilomètres en zone rurale. Il existe plusieurs méthodes de triangulation : Advanced Forward Link Trilateration (AFLT), Angle of Arrival (AOA), Enhanced Forward Link Trilateration (EFLT), Enhanced Observed Time Difference (EOTD), Time Difference of Arrival (TDOA).

La figure 10 compare les différentes technologies de géolocalisation selon leur précision spatiale.

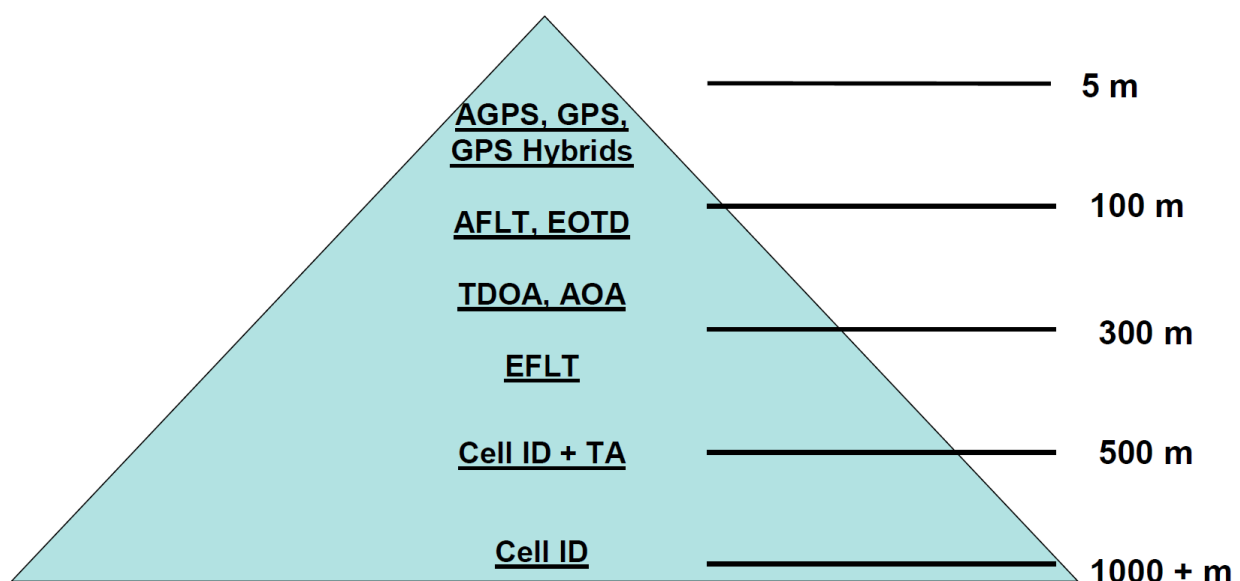


Figure 10 : précision de localisation selon différentes technologies.

En raison des limites imposées par la CNIL (Loi informatique et libertés) en France, les opérateurs de téléphonie mobile sont réticents à ouvrir leurs bases de données individuelles, même anonymisées. Il est sans doute plus facile d'obtenir en amont la **collaboration active** des individus que d'obtenir leur permission d'utiliser des données après coup. Prenons l'exemple du service « Childlocate » en Grande-Bretagne de localisation téléphonique proposé par une société privée qui a conclu un accord avec plusieurs opérateurs (Vodafone, Orange, T-Mobile). L'individu intéressé par ce service paie pour pouvoir localiser en temps réel sur une carte ses enfants équipés d'un téléphone portable.

Voir : <http://www.childlocate.co.uk>

De plus, il nous semble aujourd'hui indispensable de poser la question de la **rétribution** des enquêtés. Aux Etats-Unis, l'enquêté est payé. En France, il n'est pas coutume de le faire, surtout pas en cash. Il faut bien admettre qu'il est de plus en plus difficile pour un enquêteur à domicile de se voir ouvrir la porte pour interroger un ménage pendant une heure trente en moyenne. Les modalités de cette rétribution sont à étudier : remboursement d'un mois de forfait téléphonique, remboursement d'un mois de carte Navigo, etc

### **5.5.5. Mouvement « crowdsourcing »**

Le terme « crowdsourcing » est un néologisme sémantiquement calqué sur l'outsourcing (externalisation). La traduction littérale de crowdsourcing est « approvisionnement par la foule, ou par un grand nombre de personnes », mais l'expression ne reflète pas vraiment le sens anglo-saxon du terme : « externalisation distribuée à grande échelle » semble une traduction plus précise. Le crowdsourcing consiste à regrouper les connaissances et les compétences de personnes, le plus souvent des internautes, pour résoudre un problème et ensuite partager librement la réponse avec tout le monde. Google et Wikipédia sont les plus gros utilisateurs de crowdsourcing (source : Wikipedia).

Le travail peut être collaboratif ou au contraire s'effectuer purement en parallèle. Dans une approche économique, il peut s'agir de remplir une tâche au moindre coût, mais des approches plus collaboratives, sociales ou altruistes existent, faisant appel à des réseaux spécialisés ou au grand public. Certaines démarches de sciences participatives et sciences citoyennes l'utilisent, pour acquérir un plus grand nombre de données, à des échelles géographiques qui seraient autrement inaccessibles à des chercheurs insuffisamment nombreux ou ne pouvant faire preuve d'ubiquité.

En quoi le crowdsourcing peut-il être une nouvelle source de données en modélisation des transports ? Donnons deux exemples :

- le suivi de déplacements d'individus par GPS et l'enquête complémentaire légère auto-administrée sur internet ;
- la confrontation des résultats de modélisation : les citoyens qui ont une bonne connaissance du contexte local pourraient accéder par internet à des représentations visuelles en 3D des résultats de modélisation intégrée et interagir avec le système. L'Ecole Polytechnique de Turin travaille actuellement sur un projet de plateforme internet de simulation urbaine à l'échelle régionale appelé **NextCity** destinée aux décideurs mais aussi aux citoyens.

## 6. Spécifications fonctionnelles d'ULTISIM

Ce chapitre décrit l'**architecture** d'ULTISIM. Il tire les leçons des travaux de recherche en cours ou passés en Europe et aux Etats-Unis. Après une discussion sur le degré de formalisme des spécifications fonctionnelles, nous examinons successivement la structure modulaire, la prise en compte des contraintes de politiques urbaines et environnementales, l'approche dynamique, la mesure de l'accessibilité, la question de l'échelle géographique, le modèle de développement urbain et enfin l'interaction entre ULTISIM et ANTONIN.

### 6.1. Degré de formalisme

Quel degré de formalisme doit-on utiliser pour les spécifications d'un logiciel ? C'est une question importante, mais la réponse risque d'être décevante : il n'y a pas de recette miracle ou de formalisme universel qui marche à tous les coups dans le domaine du génie logiciel. Les experts informaticiens disent qu'ils ont vu autant de projets échouer que réussir avec des spécifications très détaillées et très formalisées (AFDEL, 2009). Ceci dit, peut-on se passer de spécifications et compter seulement sur une commande exprimée oralement ? Probablement non. Cette méthode artisanale peut marcher pour des petits projets mais pas pour un projet comme ULTISIM où l'on va sans doute évoluer en marchant.

Nous proposons un **processus simple de documentation évolutif des spécifications**. Le formalisme sera ajusté au cours du projet, ce qui permettra une mise à jour facile du document de spécifications (on en revient à l'image de la charpente d'une maison). C'est la raison pour laquelle il est préférable de mettre en place un processus souple de documentation qui se focalise sur les éléments clés à développer, et pas plus.

**Trois éléments** sont nécessaires :

- Description textuelle de ce que fait le logiciel. Si on n'est pas capable d'écrire littéralement ce qu'un programme doit faire, on sera encore moins capable de le développer ;
- Une représentation visuelle de l'interface utilisateur final, en présentant des copies d'écran de l'interface désirée ;
- Une liste la plus complète possible de règles et de comportements, incluant même ce qui arriverait dans les situations limites. Par exemple, la nécessité d'un zonage commun

entre le modèle de trafic et le modèle land use, la cohérence entre les motifs de déplacements et les modes d'occupation du sol à prédire.

## 6.2. Structure modulaire

Une structure modulaire permet une plus grande **flexibilité** en autorisant un calage par sous-modèle et en offrant une possibilité d'intégration de nouveaux sous-modèles dans le futur.

La structure générale d'ULTISIM est composée de **huit sous-modèles**, quatre pour le modèle de trafic, quatre pour le modèle land use.

Modèle de trafic	Modèle land use
Génération	Choix de localisation résidentielle des ménages
Distribution	Choix de localisation résidentielle des entreprises ou des emplois
Choix modal	Choix de localisation de projets urbains (bureaux, activités, commerces)
Affectation	Prix du foncier et de l'immobilier

Tableau 3 : les 8 sous-modèles d'ULTISIM

Le modèle de trafic se présente comme un ensemble de quatre étapes exécutées de manière **séquentielle**, tandis que le modèle land use se présente comme un ensemble de quatre étapes fonctionnant de manière **parallèle**. En théorie, les quatre étapes du modèle land use sont imbriquées les unes dans les autres, les sorties de l'un étant des entrées de l'autre. Ainsi, le prix est une variable explicative des trois modèles de localisation. La localisation des projets urbains urbains (immeubles d'habitation, immeubles de bureaux, entrepôts, locaux commerciaux) permet de créer de la disponibilité en logements et en emplois nouveaux et d'introduire une contrainte de capacité par zone. Ce modèle de localisation des projets urbains permet donc au système de ne pas affecter indéfiniment sur une même zone très attractive.

La figure 11 schématise l'architecture globale d'ULTISIM.

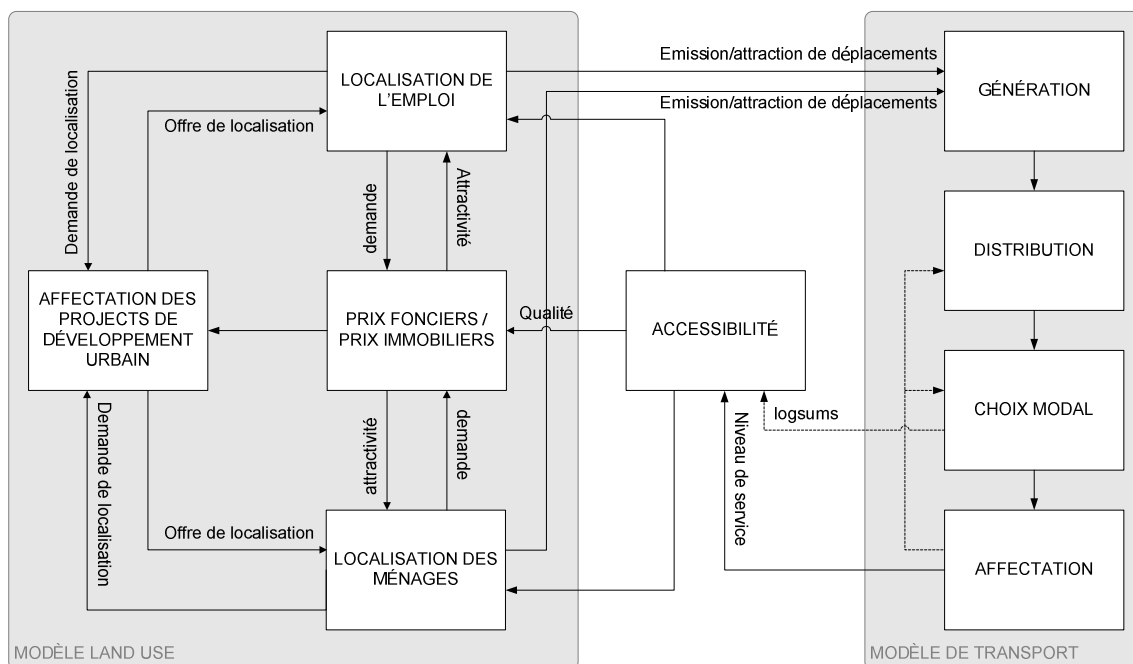


Figure 11: Architecture globale d'ULTISIM

Le résultat de ces modèles de localisation est la répartition par zone de la population et des emplois. Ces deux vecteurs sont utilisés comme émission et attraction par zone dans le modèle de génération.

Les prix du foncier et de l'immobilier sont des **facteurs déterminants** dans les choix de localisation des ménages, des entreprises et des projets urbains. Il y a des effets réciproques entre les prix et l'attractivité territoriale qui peuvent se modéliser par un modèle d'équilibre général soit par une cascade de modèles dynamiques s'échangeant les entrées-sorties par pas de temps. Suite à la mise en service d'une nouvelle infrastructure de transport, les populations et les emplois chercheront à bénéficier d'une part de l'accessibilité nouvelle créé par le projet entraînant des gains d'utilité collective et de croissance, d'autre part à bénéficier des économies d'agglomération (concentration d'activités dans une même zone). En revanche, cette demande renforcera la pression foncière sur les zones desservies par le projet due à l'attractivité résidentielle et l'attractivité économique locale autour des stations ou le long de l'axe routier.

Le **deuxième facteur** essentiel à prendre en compte est la disponibilité d'offre (pénurie ou sur-offre), en regard de la demande de localisation.



Le modèle d'urbanisme est complètement intégré au modèle de trafic. L'interaction entre les deux est réalisée par l'impact des mesures d'accessibilités (logsum) sur les localisations. En retour, ces nouvelles localisations (population et emplois par zone) seront injectées dans le modèle de génération. Dans le cas du modèle dynamique ULTISIM, cette rétroaction se fera par un bouclage par pas de temps annuel ou pluriannuel.

### **6.3. Prise en compte des politiques urbaines et environnementales**

La partie land use du modèle ULTISIM est conçue pour être **flexible** à différents niveaux de gouvernance ou de développement urbain. D'abord, l'interventionnisme de la puissance publique peut être très variable selon les régions européennes, depuis un libéralisme assumé (Grande-Bretagne) à une réglementation stricte sur l'occupation du sol et sur la constructibilité des zones (France). Citons quelques contraintes environnementales françaises : les zones de protections autour des sites et monuments protégés, les zones inondables, les Zones Naturelles d'Intérêt Ecologique Faunistique et Floristique (ZNIEFF), les zones Seveso, etc.

Par conséquent, il faut prévoir dans ULTISIM la prise en compte de ces **contraintes** au niveau du modèle de développement urbain. Ce modèle simule les changements d'occupation du sol et l'implantation de nouveaux projets urbains. Il prend en compte l'occupation du sol actuelle, la demande du marché, le niveau de régulation du marché et les schémas d'urbanisme territoriaux.

En résumé, les facteurs suivants devraient être pris en compte dans le contexte des villes européennes, davantage soumises aux politiques de régulation que les villes américaines :

- Objectif de construction de logements (par exemple, +60 000 par an en Île-de-France entre 2006 et 2030) dans le SDRIF 2007 ;
- Plans locaux d'urbanisme (PLU) en France qui fournit pour chaque commune les plans d'urbanisme pour le résidentiel, l'activité, le mixte et le non-résidentiel ;
- Les contraintes environnementales ;
- Les contraintes de politiques urbaines (schémas directeurs territoriaux, schéma directeur régional) qui définissent dans les grandes lignes de la destination des sols à un horizon futur.

## **6.4. Simulation dynamique**

La **dynamique temporelle** sera une des caractéristiques d'ULTISIM. Le logiciel sera distinguera ainsi des deux seuls **modèles d'équilibre général** de type commercial que sont Tranus et Cube Land. ULTISIM est conçu comme un modèle de croissance incrémental, comprenant des modèles partiels qui ne sont pas nécessairement en équilibre. L'approche incrémentale considère l'état à l'année n comme le point de départ pour simuler les localisations à l'année n. C'est fondamentalement différent des modèles d'équilibre général où le calcul de l'équilibre est exécuté en une fois (« one single shot »). Mais cela suppose de procéder à un calage sur une **période passée suffisamment longue** (au moins 10 ans) pour observer des évolutions de l'occupation du sol, et non sur une année de référence comme en modélisation de trafic classique ou en modélisation intégrée par approche de l'équilibre.

Nous optons pour l'approche par la simulation dynamique et non par l'équilibre général pour deux raisons. D'abord il est peu probable que l'équilibre général soit atteint en une seule année, en raison du temps de réponse très long des modifications de l'occupation du sol. Le deuxième argument est de s'interroger sur l'existence réelle d'un équilibre général entre prix, localisation et coûts des transports tel que le formalise le modèle monocentrique (Alonso, 1960). Si l'équilibre existait réellement en Île-de-France, il n'y aurait pas de sur-offre actuelle de bureaux alors que la demande des entreprises de relocalisation n'est pas satisfaite. Pour l'Île-de-France, il faut en plus tenir compte de l'inertie du système aménageur-promoteur-client.

La partie land use d'ULTISIM utilise un pas de temps annuel pour distribuer les populations et les emplois par zone, tandis que la partie modèle de trafic peut tourner à un pas de temps de trois à cinq ans. En effet, compte tenu du rythme actuel d'investissement en infrastructure de transport, on peut se permettre de décrire l'offre nouvelle avec un pas de temps pluriannuel (cinq ans voire plus).

## **6.5. Accessibilité**

Le concept d'accessibilité remonte aux années 1950, avec Hansen (1959) qui définit l'accessibilité comme « le potentiel d'opportunités pour des interactions » (Reggiani et Bucci, 2007). Makrí et Folkesson (1999) fournissent une revue de littérature sur les mesures d'accessibilité, dont les fondements théoriques sont soit les modèles d'interaction spatiale,

dont la logique initiale s'apparente à la loi de gravitation universelle de Newton, soit les modèles économiques, issus de la théorie des choix discrets, et qui supposent un comportement rationnel des agents.

La mesure d'accessibilité est basée sur les modèles économiques des choix discrets, issus de la théorie de l'utilité. Les agents économiques maximisent leur utilité à se déplacer, dont la fonction contient les variables représentant les attributs de chaque destination possible (attractivité et fonction de résistance du transport, mais aussi parfois caractéristiques socio-économiques de l'agent). La vraisemblance d'un agent choisissant une destination  $j$  dépend de l'utilité retirée de ce choix comparativement à l'utilité de tous les choix possibles, et l'indicateur d'accessibilité correspond alors au dénominateur du modèle logit multinomial issu du programme de maximisation de l'utilité à se déplacer, connu également sous le nom de **logsum**. Du point de vue économique, le logsum peut être décrit comme le surplus des usagers, c'est-à-dire le bénéfice net qu'ils retirent de leurs choix de destination.

L'accessibilité d'une zone est exprimée comme un niveau de service moyen vers l'ensemble des autres zones. Il est possible de représenter de manière désagrégée ces variables d'accessibilité. La qualité d'accessibilité d'une zone de résidence (origine) est classiquement définie comme une somme pondérée des effectifs d'emplois localisés dans les diverses zones (de destination), pondérés respectivement par un coefficient atténuateur en fonction du coût de transport origine-destination par mode. On parlera en fait **d'accessibilité globale**, pour la distinguer de **l'accessibilité locale** (proximité de chaque zone aux noeuds ferroviaires et routiers les plus proches, nombre de gares/stations dans chaque zone).

On peut envisager pour ULTISIM de désagréger non seulement par origine, mais encore par CSP pour permettre d'évaluer les effets redistributifs, et même aussi selon la localisation de l'emploi détenu, pour la partie des ménages qui choisissent leur domicile sans remettre en cause l'emploi qu'ils occupent.

Aussi, l'accessibilité aux emplois d'un individu situé dans la zone  $i$  peut s'écrire

$$A_i = \sum_j Q_j \exp\left(\frac{-C_{ij}}{\lambda}\right),$$

avec  $Q_j$  le nombre d'emplois dans la zone  $j$ ,  $C_{ij}$  le coût généralisé de transport entre  $i$  et  $j$ , et  $\lambda$  un paramètre. Le logsum est le logarithme de  $A_i$ .

Notons qu'il existe des indicateurs d'accessibilité moins complexes, qui seront prévus dans la structure de données d'ULTISIM : les temps d'accès par mode depuis chaque zone vers le

centre de Paris, vers les principaux lieux attracteurs de déplacements (aéroports, La Défense, etc).

## 6.6. Le découpage géographique

Quatre niveaux de découpage existent pour l'Île-de-France :

- Microscopique : niveau de la parcelle ou de la maille carrée de 150 mètres à 500 mètres de côté (figure 12) ;
- Mésoscopique : infracommunal, quartier, IRIS (figure 13) ;
- Communal (1300) ou regroupement de communes (TAZ) (figure 14) ;
- Macroscopique ou stratégique : une centaine de zones (figure 15).

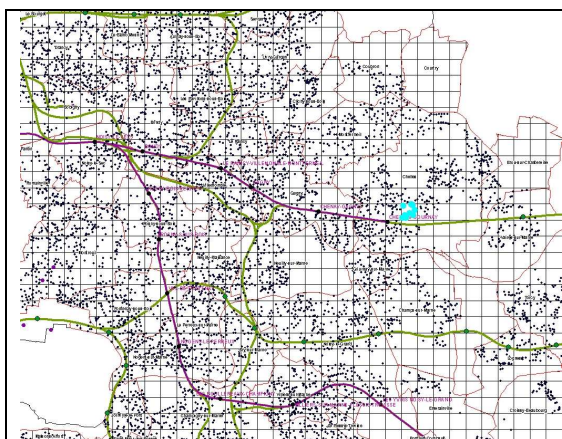


Figure 12 : Maillage carré de 500 m de côté (découpage de SIMAURIF)



Figure 13 : Extrait du découpage à l'IRIS de l'INSEE

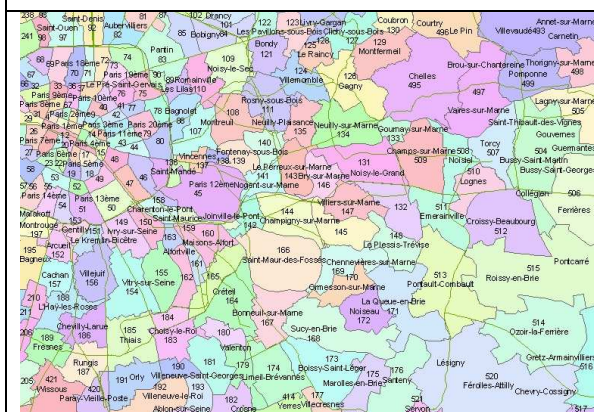


Figure 14 : Extrait du découpage en 606 zones (TAZ) du modèle global de l'IAU



Figure 15 : Découpage stratégique de l'IAU en 117 secteurs

Nos réflexions nous ont amenés à opter pour le **découpage mésoscopique à l'IRIS** (5100 zones) pour l'Île-de-France. Ce découpage serait commun au modèle de trafic et au modèle land use.

## **6.7. Le modèle de développement urbain**

### **6.7.1. Objectif**

Ce modèle est un des trois modèles de choix discrets, avec le modèle de localisation des ménages et le modèle de localisation des emplois. Il est à ce jour le modèle **le moins bien traité** dans la communauté des modélisateurs et représente à l'heure actuelle le point faible des modèles intégrés dynamiques. C'est en particulier la collecte des données qui est l'obstacle majeur. Or ce sous-modèle est essentiel au fonctionnement de l'ensemble. Il est le pendant, en quelque sorte, de la contrainte de capacité dans le modèle d'affectation routière. Si ce modèle est négligé, cela revient à localiser les ménages et les emplois sans limite de disponibilité foncière et immobilière. Rappelons que dans Urbansim 2, ce modèle était un modèle de transition d'état de type logit, alors que dans Urbansim 4, il est devenu un modèle de choix de localisation de projets urbains.

L'objectif est l'affectation au sol des projets de développement urbain (projets d'immobilier résidentiel et d'entreprises, de centres commerciaux, d'entrepôts).

### **6.7.2. Spécifications fonctionnelles**

La conception de ce modèle est illustrée sur la figure 16. En théorie, les variables intervenant dans la production de logements sont les schémas directeurs régionaux et locaux, le taux de reconstruction sur la ville, les emprises foncières disponibles et les contraintes de la destination des sols (résidentiel, activités, mixte, non urbanisables, etc) données par les Plans locaux d'urbanisme des communes (PLU).

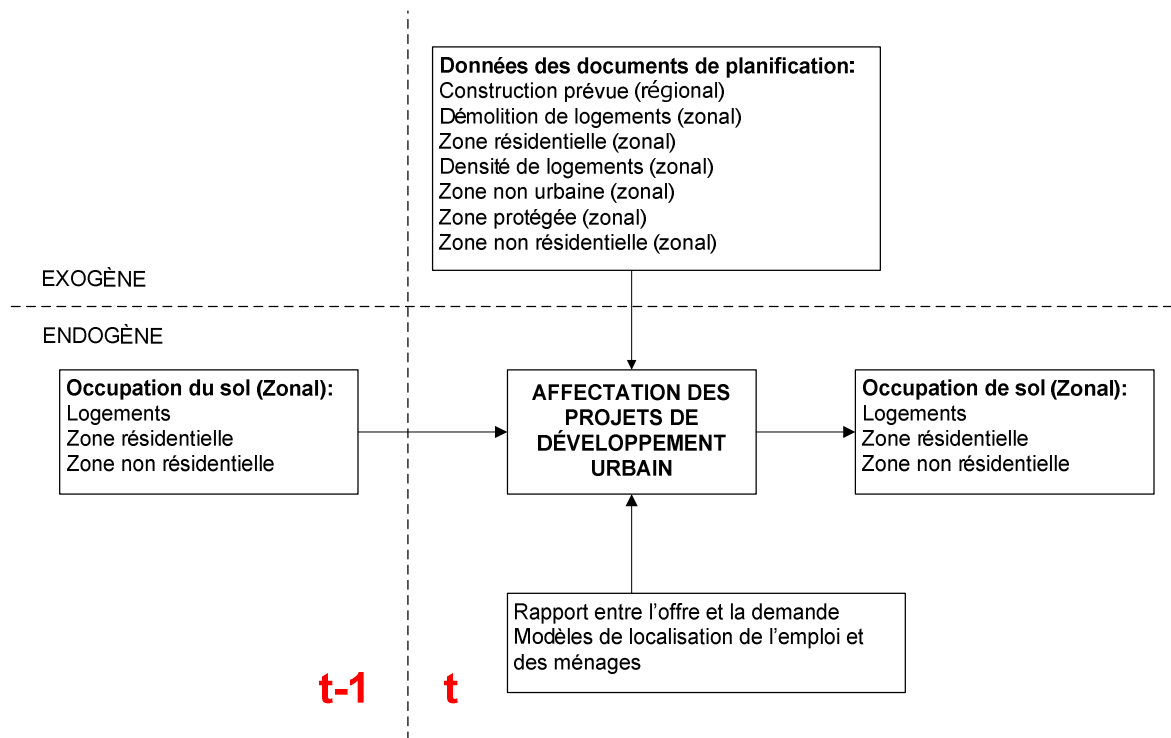


Figure 16 : Spécifications fonctionnelles du modèle de développement urbain

### 6.7.3. Base de données d'historiques

L'IAU dispose de la couche « immo\_construction » mise à jour en juillet 2010 dans le SIGR.

Elle comporte 3194 points et 7 attributs :

- Code INSEE
- Surface
- Type d'opération (0 = reconversion en autre usage que bureaux, 1 = création nouvelle, 2 = restructuration, 3 = reconstruction)
- Année de livraison de l'opération
- Adresse postale
- Nom de l'opération
- Codification du type d'immobilier (1 = bureau, 2 = entrepôt, 3 = activité)

Cette couche permet de construire une base de données désagrégées d'historiques et de calibrer le modèle de choix de localisation de projets urbains du point de vue du promoteur. Une variable explicative est évidemment le prix du terrain (foncier) mais les données sont éparses et non exhaustives. Le recueil de cette donnée se fait donc au cas par cas.

## **6.8. Interface entre *ULTISIM* et *ANTONIN***

### **6.8.1. Introduction**

Pour le cas de l'Île-de-France, l'objectif de cette description est de préparer la méthodologie permettant de connecter *ULTISIM* au modèle *ANTONIN* du STIF, dans une première phase, sachant que la partie demande de déplacements sera implémentée en propre dans *ULTISIM* lors d'une seconde phase.

**ANTONIN** est un système de prévision des déplacements de personnes sur l'ensemble de la région Île-de-France. C'est un modèle multimodal car l'usage de chaque mode de transport ne peut être analysé que dans le cadre d'une problématique globale qui intègre l'ensemble des modes en présence. Il fonctionne sur le logiciel Cube Voyager.

Le modèle permet en prenant en entrée un réseau de TC, un réseau routier et des données zonales sur un découpage donné d'obtenir des matrices de déplacements et une affectation sur chacun des arcs du réseau, et ce pour chaque mode et différentes périodes de la journée.

On distingue plusieurs grandes étapes :

- La génération des boucles de déplacements ;
- Le calcul du niveau de service, c'est-à-dire du temps et du coût pour chaque déplacement entre deux zones, pour chaque mode et chaque période ;
- Le choix pour chacune des boucles de la destination et du mode utilisé ;
- L'affectation de ces déplacements : pour chaque déplacement, l'itinéraire emprunté par l'individu est déterminé.

### **6.8.2. Description des données de réseaux du modèle *ANTONIN***

#### *6.8.2.1. Le découpage zonal*

Pour pouvoir réaliser des prévisions à des échelles géographiques allant du niveau global de la région d'Île de France au niveau local, l'Île-de-France a été découpée en **1400 zones** selon un zonage pertinent vis à vis de l'offre et de la contexture des réseaux de transports collectifs. Paris, par exemple, regroupe presque 300 zones. A mesure qu'on s'éloigne de Paris, les zones deviennent de plus en plus larges. En Grande Couronne, elles regroupent souvent plusieurs communes.

De nombreuses variables sont renseignées pour les zones du découpage de l'Île de France retenu pour le modèle :

- des données de population (population totale, répartition par âge, sexe, taille des ménages) ;
- des données d'emploi (emploi total, surfaces commerciales) ;
- le nombre de places d'études au lieu d'étude ;
- des données sur la tarification du stationnement.

#### *6.8.2.2. Le réseau de voirie*

Le réseau de voirie utilisé dans le modèle comporte 12 000 nœuds et 31 000 arcs environ. Chaque arc du réseau de voirie comprend plusieurs variables descriptives dont notamment :

- une classe de vitesse : il y a 11 classes de vitesse correspondant aux différents types de voirie (de la voie autoroutière à la voie de desserte locale) ;
- une classe de capacité : il y a 33 classes de capacité renvoyant à des courbes débit-vitesse différentes (on distingue les voies à sens unique des voies à double sens) ;
- un indicateur précisant où se situe l'arc (à Paris, en zone dense hors Paris, en zone rurale).

Tous les nœuds sont codés dans le réseau de voirie, qu'il s'agisse de nœuds routiers (les carrefours), de centroïdes de zones ou d'arrêts de transports collectifs.

#### *6.8.2.3. Les réseaux de transports collectifs*

Les réseaux de transports collectifs sont codés à l'heure de pointe du matin et à l'heure creuse du matin. Ils utilisent comme support de codification géographique le réseau de voirie du modèle, et comprennent une description exhaustive de l'offre ferrée (RER, trains de banlieue, métro et tramway) et de l'offre bus RATP. L'offre bus des réseaux OPTILE (bus d'entreprises privées desservant principalement la grande couronne) est décrite pour une sélection de lignes, retenues pour leur trafic important ou parce qu'elles assurent un rabattement vers une gare du réseau ferré ayant plus de 500 entrants par jour.

Le réseau de transports collectifs est ainsi décrit par environ 1200 missions. Les lignes de bus empruntent les arcs du réseau de voirie. Les lignes de TCSP ou du réseau ferré utilisent des arcs spécifiques.



Des données sur les temps de correspondance aux stations de métro et aux principaux pôles d'échanges ont été collectées lors d'une enquête spécifique.

La tarification simplifiée du voyage est reconstituée sur chaque origine – destination pour une carte orange mensuelle ou un billet plein tarif.

Pour chaque nœud correspondant aux stations ou gares du réseau ferré, on précise si l'accès en voiture est possible ou pas en fonction de l'existence d'un parc relais ou de la disponibilité réelle de places de stationnement autour des gares.

### 6.8.3. Description des modèles et des étapes de calcul d'ANTONIN

Il est basé sur une série de modèles désagrégés de choix discret dont l'enchaînement correspond aux étapes déterminantes qui vont conduire à la réalisation d'un déplacement. Ces modèles sont capables de prendre en compte des variables socio-économiques et démographiques qui ont une influence majeure sur les décisions qui conduisent à se déplacer et sur les caractéristiques de ces déplacements. En particulier, le revenu et l'activité interviennent dans le choix de motorisation. L'âge, le sexe et la catégorie socioprofessionnelle conditionnent le nombre et la nature des déplacements, ....

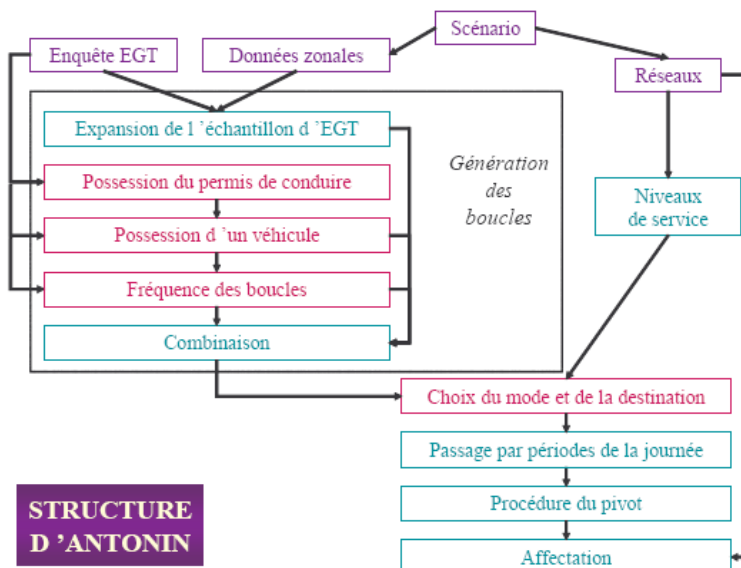


Figure 17 : Structure d'ANTONIN

#### 6.8.3.1. Architecture de modélisation

ANTONIN comporte plusieurs modèles **désagrégés** de choix discret de type logit. Ces modèles visent à expliquer le comportement des individus face à des choix entre un nombre fini d'alternatives. Cette logique découle de la théorie micro-économique de l'utilité : chaque individu accorde une utilité plus ou moins grande aux différentes alternatives possibles, l'alternative qu'il choisit est celle qui maximise son utilité.

L'enchaînement des différents modèles correspond aux étapes déterminantes qui vont conduire à effectuer un déplacement. En particulier, des choix de long terme conditionnent les choix effectués à court terme. Ce constat a conduit à utiliser l'enchaînement suivant de modèles estimés à partir des résultats de l'Enquête Globale de Transports de 2001 :

#### 6.8.3.2. Modèles appliqués aux individus

- modèle désagrégé de possession du permis de conduire associé à un modèle de cohorte du nombre total de permis de conduire en Île-de-France ;
- modèle désagrégé de possession d'une voiture au sein des ménages ;
- modèle désagrégé de possession d'un abonnement de transports collectifs ;
- modèle désagrégé de fréquence des boucles de déplacements par motifs.

#### 6.8.3.3. Modèle appliqué aux boucles de déplacement

Le modèle suivant est de type modèle désagrégé de choix **conjoint** du mode et de la destination. Le fait d'élaborer un modèle de choix du mode et de la destination conjoint permet de prendre en compte les interactions importantes existant entre le choix d'une destination parce qu'elle procure une certaine utilité et la facilité ou la difficulté à s'y rendre suivant les différents modes de transport.

L'approche retenue dans ANTONIN est la modélisation non pas de déplacements mais de **boucles de déplacements**. C'est une des innovations principales. Les boucles sont une simplification de la chaîne complète de déplacements : ce sont des allers-retours ayant pour origine soit le domicile, soit le lieu de travail avec un motif principal identifié à partir d'une hiérarchisation des différents motifs. Une approche par boucle de déplacement est plus cohérente qu'une approche par déplacement : le mode de déplacement est le même pour l'ensemble de la boucle. Il y a 10 modèles différents, un pour chaque motif de boucle de déplacements (domicile – travail, domicile – études, domicile – affaires professionnelles, domicile – achats, ...).

On notera par ailleurs que les modèles sont appliqués à des boucles de déplacements sur la journée. Les **niveaux de service** utilisés dans ces modèles sont donc des moyennes du niveau de service journalier obtenues en pondérant les niveaux de service de pointe du matin pointe du soir (qui est considérée comme égale à la transposée de l'heure de pointe du matin) et heures creuses sur la base de la répartition des allers-retours dans ces différentes périodes que l'on peut obtenir dans l'EGT. Pour la route, on se sert à ce stade du déroulement d'ANTONIN de vitesses réalistes fixées par type de voie, grandes zone géographiques et périodes de la journée.

L'analyse des déplacements en Île-de-France montre que, si globalement le choix modal est en faveur de la voiture particulière, l'utilisation des transports publics est plus forte que dans les autres agglomérations françaises. Cette situation s'explique principalement par la qualité et la variété de l'offre de transports collectifs. Ainsi, comme c'est le cas pour l'usage de la voiture particulière, le voyageur dispose généralement de plusieurs chemins alternatifs en transports collectifs pour un couple origine - destination donné. Ces chemins alternatifs peuvent donner lieu à des chaînes modales complexes sur le réseau de transports collectifs faisant appel à plusieurs modes différents. Le mode de rabattement sur le réseau (à pied ou en voiture particulière) vient ajouter un niveau de complexité supplémentaire.

Il importe donc de disposer d'une description suffisamment précise des modes de transports collectifs disponibles et de l'offre proposée qui permette la modélisation de cette situation complexe.

Pour ce faire, lors de la mise en œuvre des 10 modèles désagrégés de choix conjoint du mode et de la destination, les alternatives du choix modal ne se limitent pas à la voiture ou aux transports collectifs (VP / TC), mais ce sont **13 modes de déplacement** qui sont utilisés (10 modes TC -combinant le réseau ferré lourd, le métro et le bus- 2 modes VP –passager et conducteur- et 1 mode « doux »). Pour les transports collectifs, les différents enchaînements de sous-modes sont combinés avec le mode d'accès à pied ou en VP. Toutefois, pour les chaînes modales complexes, l'accès à pied ou en VP n'a pas été identifié séparément.

#### **6.8.4. Passage de matrices de boucles journalières à des matrices de déplacement**

Les résultats doivent permettre des analyses sur les déplacements journaliers ainsi que sur l'heure de pointe dimensionnante qui est pour les transports collectifs franciliens la pointe du

matin. Pour ce faire, les matrices de boucles de déplacement journalières sont converties en matrices de déplacements allers et retours. Les allers et les retours sont répartis par période horaires selon des coefficients issus de l'EGT pour **quatre grands modes** (voiture conducteur, voiture passager, modes doux et transports collectifs) et **trois grands motifs** (domicile – travail, domicile - autres, déplacements non basés au domicile).

#### **6.8.5. L'affectation des déplacements sur les réseaux**

Les matrices de déplacements VP conducteur et l'ensemble des sous-matrices TC sont affectées sur les réseaux de transport de manière classique.

ANTONIN dispose d'un module d'affectation pour les voitures qui prend en compte les courbes débit-vitesse du logiciel DAVIS développé par l'INRETS (aujourd'hui IFSTTAR).

Pour les transports collectifs, l'affectation est réalisée en fonction des temps perçus pour chacune des alternatives possibles. Pour un couple origine – destination et un mode donnés, plusieurs chemins (ceux qui ont un temps généralisé assez « proche » du temps perçu du meilleur chemin) sont retenus.

La probabilité de choisir un chemin parmi l'ensemble des chemins retenus est fonction du temps perçu associé à ce chemin.

#### **6.8.6. Les résultats produits par le modèle ANTONIN**

Le modèle produit différentes catégories de résultats dans des fichiers de type matriciel (format \*.trp ou \*.skm par exemple, formats reconnus par Cube et utilisables par des scripts depuis le module MATRIX) ou sous la forme de réseaux enrichi des trafics (format \*.net).

##### *6.8.6.1. Réseaux chargés à l'heure de pointe*

Le modèle produit des réseaux chargés pour les transports collectifs et pour les réseaux routiers aux heures de pointe du soir, du matin et pour la période heures creuses d'un jour moyen de semaine.

Pour les transports collectifs, lorsque plusieurs missions se superposent sur un même tronçon, les charges ainsi que les montants / descendants correspondent à la somme des différentes missions.

Ce réseau est la représentation sous Cube d'un fichier dbf, qui contient, mission par mission, rigoureusement les mêmes informations que le réseau chargé, mais sur les 2 heures de pointe.

#### *6.8.6.2. Trafic*

Le modèle produit 13 matrices Origine/Destination (format .trp). Elles correspondent aux 10 combinaisons de modes TC, aux 2 modes VP (conducteur et passager) et aux modes actifs. Chacune de ces matrices comprend de 3 à 9 onglets, dont notamment le 1er qui correspond aux résultats sur la journée, le 2ème sur les 2 heures de pointe du matin et le 3ème sur l'heure creuse.

#### *6.8.6.3. Temps de parcours*

Ces informations sont produites par Origine/Destination et par chaîne modale pour chaque période de la journée (Heure de pointe / heure creuse). Il s'agit donc de 13 matrices O/D (format .skm) qui correspondent aux 10 combinaisons de modes TC, aux 2 modes VP (conducteur et passager) et aux modes « doux ».

Chacune de ces matrices comprend 8 onglets, qui représentent pour chaque O / D outre le temps total, le temps d'accès entre la zone et la chaîne modale, le temps dans les transports, le temps passé dans le métro, le temps passé dans un bus, le temps passé dans un train, le temps d'attente, le temps passé dans les correspondances, le nombre de correspondances, le temps de rabattement et le tarif.

### **6.8.7. Echanges entrées-sorties entre ULTISIM et ANTONIN**

L'accessibilité est un facteur de choix de localisation et détermine donc l'avantage d'une zone par rapport à une. ULTISIM utilise l'accessibilité issue d'ANTONIN comme variable dans les modèles de localisation et le modèle de prix.

Les données zonales sont produites par ULTISIM et les données de temps de parcours par ANTONIN. Il faut donc prévoir un module d'**agrégation** des données zonales à l'IRIS (5100 zones) dans le découpage d'ANTONIN en 1400 zones et un module de **calcul des logsum** en 1400 zones dont les zones IRIS hériteront.

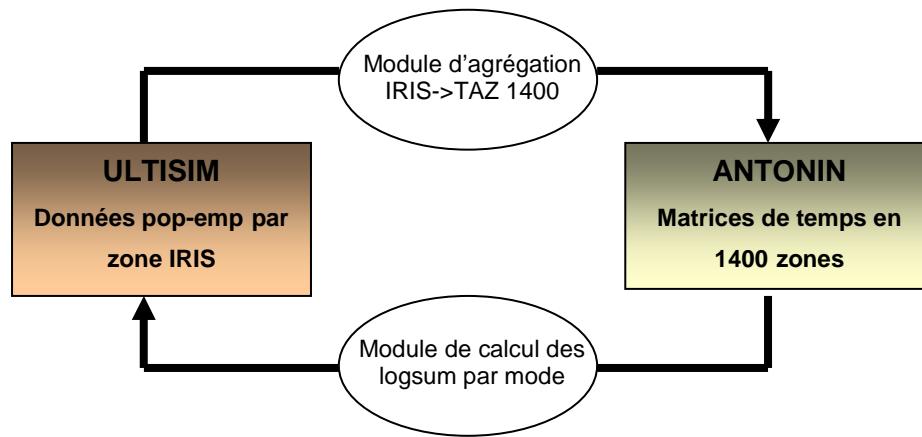


Figure 18 : Echanges entrées-sorties

ANTONIN produit des calculs intermédiaires de logsums au niveau de l'individu, du ménage et pour différents motifs. Le Logsum doit tenir compte de la structure particulière d'ANTONIN (combinaisons de modes et boucles). Il est donc défini par une somme pondérée de logsums où interviennent les ménages, les individus, la zone origine et les boucles :

$$L_{h,o} = \sum_{m,p} f_{m,p,h,o} L_{m,p,o}$$

avec :

- $L_{h,o}$  : logsum pour un ménage de type h en zone origine o
- $L_{m,p,o}$  : logsum pour le motif m, l'individu de type p en zone origine o
- $f_{m,p,h,o}$  nombre moyen de boucles pour le motif m, l'individu de type p, le ménage de type h et la zone origine o.

Des logsum similaires sont calculées pour le modèle de localisation des emplois en utilisant comme opportunités d'autres données que les emplois totaux, par exemple les emplois tertiaires et les commerces. C'est au stade du développement « agile » d'ULTISIM que ces formules seront précisées.

## 7. Conclusions, orientations, recommandations

Au terme de cette première phase du projet ULTISIM, nous faisons les **douze** conclusions, orientations ou recommandations suivantes pour la suite. Nous estimons qu'elles sont valables pour tout autre projet de développement d'un logiciel LUTI.

1. Les outils de modélisation intégrés actuels issus du milieu universitaire ne permettent pas de répondre en **mode opérationnel** à une commande d'un maître d'ouvrage. Pour permettre la **pratique au quotidien** d'un modèle intégré transport-urbanisation, il faut développer un logiciel professionnel, de type commercial. C'est l'objet d'ULTISIM.
2. Le **blocage actuel à une pratique quotidienne** de la modélisation LUTI en France n'est pas dû à des lacunes **théoriques** (les modèles dynamiques et les modèles d'équilibre sont largement documentés) mais à l'absence d'un logiciel **ergonomique**, prêt à utiliser et adapté au contexte des villes européennes, comme il en existe en modélisation de trafic classique.
3. Le logiciel ULTISIM doit être un modèle **réellement intégré** et non un outil connectant un modèle land use existant et un modèle de trafic existant. C'est donc *ex nihilo* en concevant une structure globale que cette intégration deviendra possible.
4. La structure de données d'ULTISIM doit prévoir cette réelle intégration, par une **cohérence des données** en entrée entre les deux ensembles de modèles et par la définition d'un **zonage unique**.
5. Les applications d'ULTISIM visant les villes européennes, la structure de données devra prévoir la prise en compte de **données typiquement « européennes »** (par exemple, prise en compte des modes Transport en commun et vélo, des politiques urbaines et environnementales).
6. Les **différents sous-modèles** doivent être **adaptés** aux villes européennes. Ainsi, le modèle de prix sera un modèle de prix de l'immobilier et non du foncier.
7. Il n'existe pas de base de données standard dans chaque ville européenne. Les données étant principalement recueillies localement, il faudra laisser une certaine **souplesse** dans la définition des variables tout en prévoyant un **commun dénominateur** à la structure de données.

8. ULTISIM sera un modèle **dynamique et désagrégé de choix discrets** pour la partie land use et non un modèle d'équilibre général. Dans une première étape, la partie modèle de trafic sera un **modèle classique à quatre étapes propre** à ULTISIM, basé sur des motifs de déplacements choisis en cohérence avec les données du modèle de développement urbain et sur les modes mécanisés (VP, TC, 2 Roues), sans oublier les marchandises. La priorité doit être accordée au développement de la partie modèle d'urbanisation et à l'intégration d'un modèle de trafic simple. Dans une seconde étape, on peut envisager de mettre au point un modèle de demande plus sophistiqué de type désagrégé.
9. Dans une **étape de transition**, le modèle de trafic sera le modèle **ANTONIN** du STIF qui sera couplé à la partie land use du modèle ULTISIM.
10. ULTISIM sera orienté **utilisateur final**, destiné aux praticiens. La méthodologie de développement sera « **agile** » et centrée sur l'élaboration d'une interface utilisateur qui sera peaufinée au fur et à mesure, en étroite concertation avec les utilisateurs finaux. L'effort doit être porté sur **l'ingénierie du logiciel**. Le choix du logiciel propriétaire apparaît le mieux adapté.
11. L'absence d'un logiciel comme ULTISIM obligera les praticiens à continuer à faire du **bricolage** coûteux et improductif à partir d'outils existants.
12. Le développement d'ULTISIM nécessitera des **investissements relativement importants**, en particulier pour travailler étroitement avec une société de services et d'ingénierie informatique (SSII). Sans aller jusqu'à une exploitation commerciale intensive du produit, on peut justifier les moyens financiers nécessaires par des besoins croissants en études d'évaluation socio-économique des bénéfices élargis. A ce stade, il paraît indispensable de **mobiliser et de mutualiser des moyens financiers** pour enclencher cette phase de développement, peut-être en cherchant du côté des financements européens.



## Bibliographie

1. AFDEL (Association Française des Editeurs de Logiciels) – La conduite de l'innovation logicielle : le livre blanc de la Commission R&D (2009).
2. Ahas R., Mark U., Aasa A., Kalle H. – Mapping human behavior in Tallinn. in A. Litjens et al, ed., Tallinn. Delft: TU Delft (2005).
3. Alonso W.– A theory of the Urban Land Market. Papers and proceedings of the Regional Science Association. Vol. 6, No1, 1960, pp.149-157 (1960).
4. Anas A., Yu L. – A regional economy, land use and transportation model (RELU-TRAN) : Formulation, algorithm design and testing (2007).
5. Boucq E., Nguyen-Luong D., Papon F. – Impact du tramway T3 sur les prix de l'immobilier résidentiel. Application de SIMAURIF. Predit 4 (2011)
6. Bramley G., Simmonds D., Dobson A. – Using Land Use Transport Interaction (LUTI) Models to assess the sustainability of strategic planning and infrastructure proposals : the potential scope and illustrative applications. 7th International Workshop: Towards methodological innovation in planning evaluation. Edinburgh, 10-12 December 2008.
7. Cambridge Systematics Inc. – Deriving Economic Development Benefits of Transit Projects from Integrated Land Use Transportation Models. For Federal Transit Administration, Office of Planning and Environment (2009)
8. Caneparo L., Guerra F., Montuori A. – Uncertainty and Feature Selection of Integrated Land Use-Transportation Microsimulation Models (2009)
9. Charlot S., Lafourcade M. – Économie géographique et infrastructures de transport, dans *Les Effets d'agglomération dans la métropolisation de l'économie*, P.-P. Combes et I. Thomas (eds.), CIFOP, Bruxelles (2000).
10. Citilabs – Cube Land - Reference Guide (2010).
11. De Bok M., Geurs K., Zondag B. – Application of a LUTI model for the assessment of land use plans and public transport investments. Proceedings of the 12th WCTR in Lisbon (2010).
12. De Bok M. – Estimation and validation of a microscopic model for spatial economic effects of transport infrastructure. Elsevier, Transportation Research Part A 44-59 (2009).
13. Eagle N., Pentland S. – Eigenbehaviors: Identifying Structure in Routine. Behavior Ecology Sociobiology 63, 1057-1066 (2009).
14. Eurostat Commission européenne. – Statistiques urbaines et régionales. Guide de référence (2008).

15. Fogel K. – Producing Open Source Software - How to Run a Successful Free Software Project. Available at: <http://producingoss.com/en/index.html> . Chap. 9. (2010).
16. Gambs S., Killijian M.-O., Núñez del Prado Cortez M. – Show me how you move and I will tell you who you are. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS (SPRINGL '10). ACM, New York, NY, USA, 34-41 (2010).
17. Geurs K.T., Van Wee B. – Land-use/transport Interaction Models as Tools for Sustainability Impact Assessment of Transport Investments: Review and Research Perspectives EJTIR, 4, no. 3, pp.333-355 (2004).
18. González M.C., Hidalgo A.C., Barabási A.-L. – Understanding Individual Human Mobility Patterns. Nature 453, 779-782 (2008).
19. Hansen W.G., 1959. “How accessibility shapes land use”. Journal of the American Institute of Planners 25:73-76.
20. Hunt J.D., Abraham J.E. – Design and Application of the PECAS Land Use Modelling System. In Proceedings of the 8th Computers in Urban Planning and Urban Management Conference, Sendai, Japan (2003).
21. Lerner J. et Tirole J. – The Scope of Open Source Licensing , IDEI Working Paper, n. 219 (2003).
22. Makri M. et Folkesson C., 1999. “Accessibility measures for analyses of land-use and travelling with geographical information systems”, Working paper, Department of Technology and Society, Lund Institute of Technology, Sweden.
23. Martínez F.J. – MUSSA : A Land Use Model for Santiago City. Transportation Research Record 1552: Transportation Planning and Land Use at State, Regional and Local Levels, 126-134 (1996).
24. Morency C. – Contributions à la modélisation totalement désagrégée des interactions entre mobilité urbaine et dynamiques spatiales. Thèse de l'Université de Montréal (2005).
25. Mukherjee D. – Open Source vs Proprietary Software - The never ending Battle. Available at : <http://brajeshwar.com/2010/open-source-vs-proprietary-software-%E2%80%93-the-never-ending-battle/> (2010).
26. Nguyen-Luong D., De Palma A., Picard N., Motamedi K., Ouaras H. – Quatre rapports SIMAURIF, disponibles sur le site Internet de l'IAU [www.iau-idf.fr](http://www.iau-idf.fr) dans la rubrique Nos etudes/Transport et mobilité/Simulation de trafic (2004, 2005, 2007 et 2008).

27. Patterson Z. et Bierlaire M. – An UrbanSim Model of Brussels within a Short Timeline. Proceedings of the 7th Swiss Transport Research Conference. September 2007, Monte Verità, Switzerland (2007).
28. Pfaffenbichler P. – The strategic, dynamic and integrated urban land use and transport model MARS (Metropolitan Activity Relocation Simulator)—Development, testing and application, Institute for Transport Planning and Traffic Engineering, Beiträge zu einer ökologisch und sozial verträglichen Verkehrsplanung Nr. 1/2003, Vienna University of Technology, Vienna (2003)
29. Pfaffenbichler P., Emberger G., Shepherd S. – The Integrated Dynamic Land Use and Transport Model MARS, Springer Science + Business Media, LLC (2007).
30. Pritchard D.R. et Buliung R.N. – Transportation analysis and research using free/open Source software." Report to Project Open Source | Open Access Student Experience Program Available at : <http://open.utoronto.ca/> (2007).
31. Simmonds, D.C., Feldman O. – Advances in Integrated Urban/Regional Land-Use/Transport Modelling Using the DELTA Package, World Conference on Transport Research, Berkeley, Calif. (2007).
32. Strohmeier A., Buchs D. – Génie logiciel : principes, méthodes et techniques. Lausanne, Presses polytechniques et universitaires romandes (1996).
33. Transportation Research Board, Special Report 288 – Metropolitan Travel Forecasting: Current Practice and Future Direction. Transportation Research Board of the National Academies, Washington, D.C. (2007).
34. Waddell P. – UrbanSim: Modeling urban development for land use, transportation and environmental planning, Journal of the American Planning Association, 68, 297-314 (2002).
35. Waddell P, Ševčíková H., Socha D., Miller E., Nagel K. – Opus: An Open Platform for Urban Simulation. Presented at the Computers in Urban Planning and Urban Management Conference, London, U.K, June (2005).
36. Waddell P., Ulfarsson G.F., Franklin J. P. & Lobb J. – Incorporating land use in metropolitan transportation planning, Transportation Research Part A: Policy and Practice, Volume 41, Issue 5, pp. 382-410 (2007).
37. Wegener M. – Overview of land-use transport models, Spiekermann & Wegener Urban Regional Research (2003).
38. Zhong M., Hunt J. D., Abraham J.E. – Design and Development of a Statewide Land Use Transport Model for Alberta. Journal of Transportation systems engineering and information technology. Volume 7, Issue 1 (2007).