



PREDIT 2002-2006

Ministère de l'Équipement, des Transports, de l'Aménagement du territoire, du Tourisme et de la Mer / DRAST

**SIMAU<sup>R</sup>IF :**

**Modèle dynamique de SIMul<sup>A</sup>tion de l'interaction  
Urbanisation-transports en Région Ile-de-France**

**Application à la Tangentielle nord**

**Rapport final de la 1<sup>ère</sup> année**

IAURIF

Université de Cergy Pontoise, THEMA

adpC



Septembre 2004

IAURIF 5.03.013  
DRAST n°03MT11

**Institut d'Aménagement et d'Urbanisme de la Région d'Ile-de-France (IAURIF)**

15, rue Falguière 75740 Paris cedex 15

Tél. 01.53.85.53.85 - Télécopie 01.53.85.76.02

E-mail : [dgcomm@iaurif.org](mailto:dgcomm@iaurif.org) –http : [www.iaurif.org](http://www.iaurif.org)

Directeur Général par intérim : François DUGENY

Directeur du Département Transports et Infrastructures : Alain MEYERE

**Université de Cergy-Pontoise**

**Laboratoire THEMA**

33 boulevard du Port 95011 Cergy-Pontoise Cedex

Tél. 01.34.25.63.37 - Télécopie 01.34.25.62.33

Directeur : Régis RENAULT

Cette recherche est menée par :

**Pr André DE PALMA**, Conseiller scientifique, THEMA

**Dany NGUYEN-LUONG**, IAURIF

**Kiarash MOTAMEDDI**, THEMA

**José MOYANO**, adpC

# Remerciements

Cette recherche a été soutenue par le PREDIT 3. Elle a également reçu une subvention de la DREIF (François Poupard et François Bertrand).

Nous tenons à remercier le Professeur Paul Waddell pour son aide concluante et précieuse tout au long de ces recherches. Nous avons aussi bénéficié de l'aide d'Eric Delattre, d'Omar Mekkaoui et de Nathalie Picard de l'Université de Cergy-Pontoise.

# PLAN

<b>REMERCIEMENTS.....</b>	<b>3</b>
<b>1 INTRODUCTION.....</b>	<b>6</b>
1.1 CONTEXTE.....	6
1.2 OBJECTIF.....	11
1.3 PHASAGE ET CALENDRIER .....	14
<b>2 APPLICATION À LA TANGENTIELLE NORD.....</b>	<b>16</b>
2.1 CONTEXTE.....	16
2.2 APPORT DE NOTRE RECHERCHE .....	18
<b>3 APERÇU DE LA LITTÉRATURE SUR LES MODÈLES D'URBANISATION... 19</b>	<b>19</b>
3.1 HISTORIQUE .....	19
3.2 ETAT DE L'ART .....	20
3.3 QUATRE EXEMPLES .....	21
3.3.1 MEPLAN.....	21
3.3.2 ISHTAR.....	22
3.3.3 DRAM/EMPAL.....	24
3.3.4 TRANUS.....	27
3.4 SYNTHÈSE .....	29
<b>4 MODÈLE D'URBANISATION URBANSIM.....</b>	<b>30</b>
4.1 INTRODUCTION.....	30
4.2 L'ARCHITECTURE D'URBANSIM.....	30
4.3 LES DIFFÉRENTS MODULES .....	32
4.3.1 <i>Le module d'accessibilité</i> .....	32
4.3.2 <i>Les modules de transition</i> .....	32
4.3.3 <i>Les modules de mobilité</i> .....	34
4.3.4 <i>Les modules de localisation</i> .....	34
4.3.5 <i>Le module de développement immobilier ou urbain</i> .....	36
4.3.6 <i>Le module de fixation du prix de terrain</i> .....	38
4.4 EXEMPLE : CHOIX DE LOCALISATION D'UN EMPLOI.....	39
4.5 LES DONNÉES D'ENTRÉE.....	39
4.6 LES PARAMÈTRES DES MODULES ET LEURS ESTIMATIONS.....	42
4.6.1 <i>Le calage des paramètres</i> .....	42
4.6.2 <i>Estimation</i> .....	43
4.6.3 <i>Contrôle et ajustement des estimations</i> .....	44
4.7 ETUDE DE CAS : EUGENE-SPRINGFIELD .....	44
4.8 SYNTHÈSE .....	48
<b>5 MODÈLE DE TRAFIC METROPOLIS .....</b>	<b>51</b>
5.1 INTRODUCTION.....	51
5.2 ELABORATION DES MATRICES ORIGINE - DESTINATION .....	52
5.2.1 <i>Introduction</i> .....	52
5.2.2 <i>Segmentation des motifs</i> .....	53
5.2.3 <i>Génération</i> .....	54
5.2.4 <i>Distribution</i> .....	57
5.2.5 <i>Choix modal</i> .....	62

5.2.6	<i>Choix de l'heure de départ</i> .....	63
5.2.7	<i>Affectation</i> .....	65
5.3	MODÈLE D'AFFECTATION EN TRANSPORTS EN COMMUN .....	66
5.4	SYNTHÈSE .....	67
<b>6</b>	<b>ARCHITECTURE DU MODÈLE INTÉGRÉ URBANSIM-METROPOLIS.....</b>	<b>68</b>
6.1	INTRODUCTION.....	68
6.2	STRUCTURE DU MODÈLE INTÉGRÉ .....	68
6.3	ECHANGES DE DONNÉES .....	69
6.4	DESCRIPTION TECHNIQUE .....	71
6.4.1	<i>Module d'ordonnement et gestion des simulations</i> .....	71
6.4.2	<i>Production des matrices Origine-Destination</i> .....	71
6.4.3	<i>Adaptateur des Logsommes</i> .....	71
6.5	DESCRIPTION DE L'INTERFACE GRAPHIQUE DU MODULE D'ORDONNANCEMENT ET DE GESTION .....	72
6.5.1	<i>Gestion des paramètres</i> .....	72
6.5.2	<i>Gestion avancée des paramètres</i> .....	73
6.5.3	<i>Gestion de l'exécution</i> .....	74
6.6	SYNTHÈSE .....	75
<b>7</b>	<b>CONSTRUCTION DES BASES DE DONNÉES ILE-DE-FRANCE .....</b>	<b>76</b>
7.1	INTRODUCTION.....	76
7.2	BASE DE DONNÉES EN ENTRÉE D'URBANSIM .....	77
7.2.1	<i>Structure de données</i> .....	77
7.2.2	<i>Table « gridcells »</i> .....	78
7.2.3	<i>Table « households »</i> .....	91
7.2.4	<i>Table « jobs »</i> .....	92
7.2.5	<i>Autres tables</i> .....	93
7.2.6	<i>Table des constantes d'UrbanSim</i> .....	103
7.3	BASE DE DONNÉES DE CALAGE .....	104
7.3.1	<i>Enquête Globale Transport (EGT)</i> .....	104
7.3.2	<i>Enquête Nationale Logement (ENL)</i> .....	105
7.3.3	<i>Base de données d'évolution du Mode d'Occupation du Sol (MOS)</i> .....	106
7.3.4	<i>Recensement général de la population (RGP)</i> .....	106
7.4	LES DONNÉES DE RÉSEAUX.....	107
7.5	SYNTHÈSE .....	108
<b>8</b>	<b>CONCLUSIONS.....</b>	<b>110</b>
	<b>BIBLIOGRAPHIE .....</b>	<b>111</b>
	<b>ANNEXE 1 : ETUDE DE CAS EUGENE –SPRINGFIELD : PRESENTATION DES DONNEES.....</b>	<b>114</b>
	<b>ANNEXE 2 : ETUDE DE CAS EUGENE-SPRINGFIELD : PARAMETRES DES MODULES.....</b>	<b>117</b>
	<b>ANNEXE 3 : MODELE DE DEMANDE DE L'IAURIF : PROCEDURE DE REDUCTION DES MOTIFS.....</b>	<b>131</b>
	<b>ANNEXE 4 : FONDEMENTS THEORIQUES DE METROPOLIS.....</b>	<b>133</b>
	<b>ANNEXE 5 : GLOSSAIRE .....</b>	<b>145</b>

# 1 Introduction

## 1.1 Contexte

L'interaction urbanisation-transports n'a encore jamais été modélisée en France<sup>1</sup>.

Cette prise en compte de l'interaction est essentielle. Pour décider de l'intérêt d'un projet de nouvelle infrastructure de transport, les acteurs de l'aménagement et des transports s'interrogent sur les effets à court, moyen et long terme de sa mise en service sur l'occupation du sol en terme de localisation résidentielle, d'activités économiques et d'emplois. Réciproquement, ils cherchent à mesurer les effets de ces impacts socio-économiques sur la mobilité et les déplacements.

Cette interaction suppose aussi une dimension « dynamique ». En effet, les impacts de la mise en service d'une grande infrastructure de transport, en plus des gains d'accessibilité, sont très complexes, durables et graduels sur l'urbanisation : l'environnement proche des nouvelles gares se densifie, de nouveaux équipements et commerces s'implantent, des entreprises se délocalisent tandis que des ménages tentent de se rapprocher de leur lieu de travail, changent de mode de transport, effectuent de nouveaux déplacements. On parle ainsi d'« effets structurants des transports », que ce soit de la densification, de l'étalement urbain ou du développement économique local. Les relocalisations sont aussi conséquence, bien sûr, des modifications du prix du marché du foncier. Ces changements prennent du temps, mais personne ne sait combien de temps.

Il existe donc un postulat bien ancré selon lequel la qualité des infrastructures de transport et le développement économique sont corrélés<sup>2</sup>. Mais jamais personne n'a réussi à établir une relation causale précise entre les deux. On trouve même un courant de pensée qui met en doute la réalité d'une causalité linéaire entre le développement d'une nouvelle infrastructure et des transformations spatiales, sociales ou économiques. Il nous semble pourtant évident (et les chiffres attestent ce point de vue) qu'une bonne accessibilité à un secteur donné et un niveau de service fiable en terme de vitesse de parcours favorisent la localisation des entreprises et des habitants dans ce secteur et que c'est une condition nécessaire mais non suffisante. Les effets structurants des transports sur l'économie et l'occupation des sols nécessitent des études supplémentaires, surtout pour mettre à jour des effets quantitatifs.

On ne sait pas bien isoler l'effet propre de l'infrastructure parmi tous les facteurs pouvant intervenir. Est-ce du ressort de la seule conjoncture, de la politique économique (incitations

---

<sup>1</sup> Il y a bien eu une tentative de modélisation dans l'agglomération de Lyon conduite par le CERTU avec le logiciel TRANUS en 1999. Il semble qu'elle n'ait pas donné entière satisfaction. Notons aussi l'existence de MOBISIM I, II et III, projet financé par le PREDIT. C'est un outil qui permet de tester des politiques de transports dans les aires urbaines par le biais de la dynamique des systèmes. Enfin, dans les années 60 à l'IAURP, il y a eu quelques recherches et développements mais 40 ans après il n'en reste rien à l'IAURIF.

<sup>2</sup> Le Président G. Pompidou déclarait lors d'un discours inaugural du dernier tronçon de l'autoroute Lille-Paris-Marseille à Savigny-lès-Beaune (A6) le 29 octobre 1970 : « Ce qu'il y a de plus important dans l'autoroute, [...] (c'est) son caractère lié à l'activité économique et le fait que l'autoroute n'est pas simplement chargée de faciliter la circulation, mais de répondre à des besoins, de les changer, en quelque sorte de les créer, en créant l'activité, en développant l'économie, en étant ce que l'on appelle un facteur d'entraînement ».

fiscales par exemple) ou de la dynamique locale via les communes et les promoteurs de zones d'activités ? L'analyse hédonique du prix du terrain fournit néanmoins certaines pistes. Les effets indirects ne sont pas à négliger non plus : les activités de logistique sont souvent les premières à bénéficier de la mise en service d'une infrastructure routière, le marché du foncier évolue dans les environs immédiats, ...

Ainsi, des interactions fortes et complexes existent entre urbanisation, demande de déplacement et offre de transport. Le tout évolue dans un ensemble de contraintes urbaines, environnementales, politiques et légales qui ne peuvent être ignorées. Or il n'y a à ce jour pas de prise en compte de ces interactions dans les modèles de trafic en France. Aujourd'hui, les sorties des modèles de trafic servent à alimenter les calculs de bilan socio-économique de type coûts-avantages. Mais tous les experts et les chercheurs reconnaissent que ce bilan ne suffit pas à appréhender l'ensemble des effets structurants d'une nouvelle infrastructure, à savoir l'économie urbaine. Par exemple, il existe une véritable lacune dans l'évaluation économique des effets de l'étalement urbain ou de la densification.

En principe, il faudrait modéliser dans une approche holistique ou systémique les interactions suivantes : 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 et les réseaux de transport sont alors modifiés à chaque horizon temporel. Ce processus itératif entre la modélisation de l'urbanisation et la modélisation de trafic des projets de transports permet de prendre en compte des décisions individuelles qui se situent à des échelles de temps très différentes :

- l'extrême court terme pour les choix d'heure de départ
- le très court terme pour les choix d'itinéraires ou de mode de transport
- le court-moyen terme pour les choix de destination
- le moyen-long terme pour les choix de localisation résidentielle ou de localisation d'activités.
- le long terme pour l'évolution des types urbains de l'occupation du sol.

La figure 1 montre la complexité des interactions entre urbanisation, transports, conjoncture économique, régulations par le politique, ...

Pour « mettre en équations » l'ensemble de ces interactions, on parlera de **modèle intégré urbanisation-transports** (ou urbanisme-transport). Cela consiste à interfacier deux types de modèle : **un modèle de trafic et un modèle d'urbanisation**. Comme la dimension du temps intervient aussi (implicite dans le schéma), on le qualifiera en plus de modèle **dynamique**. Il est évident que tout ne peut être pris en compte dans notre projet SIMAURIF et qu'il est raisonnable de penser que certaines relations sont trop complexes et chaotiques pour pouvoir être modélisées.

Pour le deuxième type de modèle, c'est le terme de modèle d'urbanisation qui a été retenu tout au long de notre projet. Les Anglo-Saxons parlent de « land-use model ». On trouve dans la littérature française plusieurs traductions : modèle d'occupation du sol, modèle d'usage du sol, modèle de localisation ou encore modèle urbain. On a préféré ne retenir aucune de ces quatre appellations car elles peuvent prêter à confusion avec d'autres expressions spécialisées :

- Mode d'Occupation du Sol (MOS) qui est l'inventaire numérisé de l'occupation du sol à l'IAURIF,

- Le terme « localisation » semble restrictif par rapport au terme « urbanisation » et de plus peut évoquer d'autres notions qui n'ont rien à voir tels que les systèmes de localisation GPS (Global Positioning System),
- « modèle urbain » est un terme consacré chez les urbanistes et recouvre des concepts de formes urbaines (ville monocentrique, polycentrique, étalée, dense, linéaire, ...).

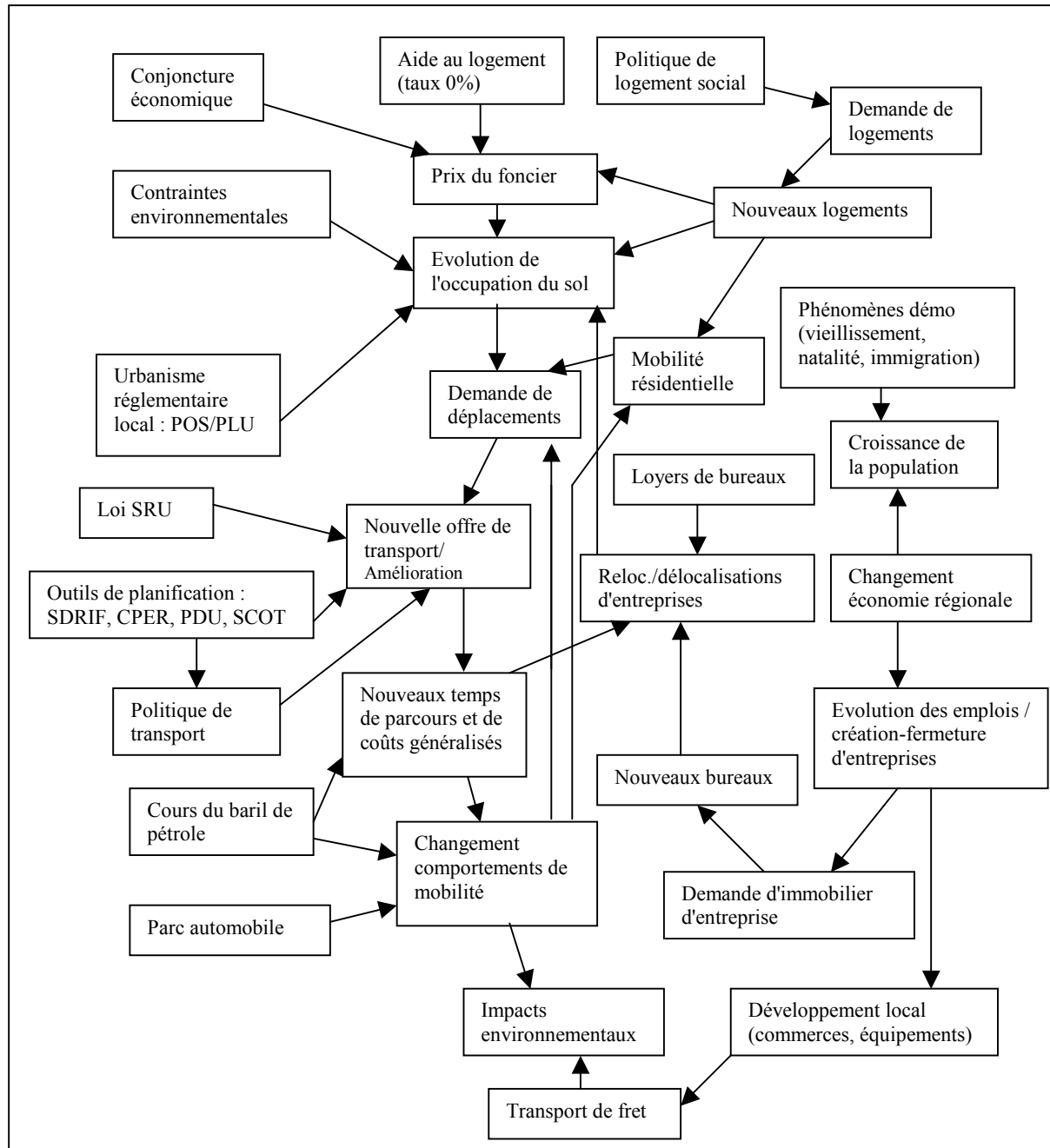


Figure1 : Schéma des interactions complexes du système urbanisation - transports

Il existe bien des modèles d'urbanisation. Cependant, ils ne sont pas utilisés actuellement en France alors qu'ils sont pratiqués couramment aux Etats-Unis et dans différents pays européens (Belgique, Pays-Bas, Norvège, Royaume-uni par exemple). La loi américaine sur l'air, ISTEA (1991), puis son extension TEA 21 en 1998 exigent en effet que les



problématiques d'occupation du sol et des transports soient complètement intégrées dans tout processus de planification régionale. De même, l'Union Européenne exige, dans de nombreuses études (voir partie 3.3.2 du projet ISHTAR, les analyses effectuées dans le projet MC-ICAM pour la DG TRENEN, etc. ) que la composante urbanisation soit prise en compte dans les études de trafic.

On peut donc s'étonner du retard pris en France dans la modélisation intégrée urbanisation – transports. A cela il peut y avoir plusieurs explications :

- l'absence de loi comme aux Etats-Unis obligeant à l'intégration urbanisation – transports dans un document de planification de type Contrat de Plan, Plan de déplacements urbains (PDU) ou Plan régional de la qualité de l'air (PRQA).
- l'absence de logiciel de prévision de l'urbanisation accessible, compréhensible, et suffisamment complet, destiné aux professionnels, alors qu'il existe plusieurs logiciels commerciaux de simulation de trafic (Davisum, Cube et Emme/2 pour les trois principaux utilisés en France). Pour les praticiens qui seraient un peu familiarisés avec les jeux vidéo ludiques, il serait souhaitable de trouver sur le marché un logiciel comparable au jeu SimCity mais à destinée professionnelle.
- les évidentes difficultés méthodologiques de la modélisation, ne serait-ce que parce qu'il est extrêmement difficile de dissocier les effets purement transports des effets exogènes sur l'urbanisation.
- le manque de polyvalence des modélisateurs de trafic, des urbanistes et des démographes. Les premiers n'ont en général qu'une approche mathématique, micro-économique et informatique de la simulation de trafic, les seconds n'ont en général qu'une approche qualitative des problèmes basée sur une longue expérience et pour qui les transports ne sont qu'une branche de l'urbanisme, les troisièmes raisonnent essentiellement en termes de soldes migratoire et naturel au niveau d'une commune (avec des paramètres tels que le taux de natalité, de décès, d'immigration, l'espérance de vie, ...) et les modèles démographiques de type âge-cohorte ne prennent pas en compte les variables d'accessibilité et de desserte par les transports. La barrière reste encore étanche entre ces différentes spécialités.

Ainsi, en France, le contexte est peu favorable au couplage des modèles d'urbanisation et de trafic. Par contre, aux Etats-Unis, les lois sur l'air ont fait de l'intégration urbanisation-transport un thème de recherche prolifique. Dans ce contexte, une équipe de recherche de l'Université de Washington aux Etats-Unis, dirigée par le Professeur Waddell, a commencé à développer en 1998 un logiciel d'urbanisation appelé **UrbanSim**. 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 présente de nombreux avantages pour notre projet :

- Il est l'aboutissement de nombreuses recherches commencées dès les années 60, d'abord en Amérique latine, puis aux Etats-Unis. Le concepteur d'UrbanSim, le Professeur Waddell, a su tirer des expériences de l'utilisation d'autres logiciels existants tels que Meplan, Tranus, MetroSim, 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 entreprises.

- Le logiciel a été conçu dès le départ pour fonctionner en interaction avec différents modèles externes classiques de prévision de trafic. De plus, en post-traitement, les résultats peuvent être exportés dans un système d'information géographique pour être visualisés, ce qui constitue un élément de communication et de diffusion incontournable aujourd'hui.
- 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 constamment mis à jour ([www.urbansim.org](http://www.urbansim.org)). Chaque mois, des nouveaux utilisateurs rejoignent cette communauté. On trouve sur le site de très nombreux articles, des rapports d'étude et de la documentation en ligne sur le logiciel.
- L'équipe du Professeur Waddell a opté pour la stratégie du logiciel dit « libre » (sous licence publique GNU) pour la diffusion d'UrbanSim. Celui-ci est développé entièrement en langage Java (langage « libre » lui aussi). Le code source est disponible de sorte que tout utilisateur peut l'utiliser, le modifier et le redistribuer. UrbanSim est gratuit et téléchargeable sur Internet, tout en offrant les caractéristiques d'un outil moderne, c'est-à-dire modulaire, ouvert et robuste.

Il est clair que c'est la disponibilité de ce logiciel qui a été l'élément déclencheur de notre recherche. Il s'avère que le Professeur Waddell a été mis au courant dès le début de notre recherche et de notre intérêt pour son logiciel. Il a accepté de nous apporter toute l'aide possible, tant au niveau de conseils méthodologiques qu'en adaptant, si nécessaire, le code informatique d'UrbanSim au cas de l'Ile-de-France. En effet, UrbanSim a été conçu dans le contexte du système et de l'économie urbaine américaine. Les variables expliquant tel phénomène aux Etats-Unis (par exemple le prix du terrain pour le choix résidentiel, le taux de motorisation pour le choix modal) ne sont peut-être pas discriminantes dans le cas de la région Ile-de-France. De même la variable d'accessibilité par le mode voiture particulière apparaît sur-représentée dans les différents modules d'UrbanSim, ce qui est compréhensible dans le cas américain où la voiture particulière est le mode prépondérant. Mais on ne trouve pas de variables décrivant l'accessibilité et la desserte par les transports en commun, sauf dans la formule de l'utilité. Il faudra remédier à cette lacune lors de l'adaptation au cas de la Région Ile-de-France.

Dès octobre 2000, l'IAURIF a téléchargé la première version bêta d'UrbanSim (0.9) puis l'a testée sur l'étude de cas proposée sur le site. Il y avait concrètement un produit tangible, bien réel, avec un code source disponible, un interpréteur Java et une interface utilisateur de type Windows (visiblement développée au départ sous Linux) appelée « dataprep ». Cette réalité nous a convaincu de garder par la suite une veille technologique sur les développements futurs d'UrbanSim et de nous mettre en contact avec le Pr Paul Waddell.

Invité par l'IAURIF et l'Université de Cergy-Pontoise, le Pr Paul Waddell a fait une présentation d'UrbanSim lors de la 5<sup>ème</sup> Journée Transports organisée par l'Université de Cergy-Pontoise le 24 octobre 2002 (voir <http://dynamic.psh.u-cergy.fr/news>). L'application d'UrbanSim sur une aire aussi grande que la région d'Ile-de-France serait pour lui une première et une référence. Séduit par le « challenge » que constitue notre projet de recherche, il nous a fait l'honneur de rejoindre notre équipe dans le cadre d'une année sabbatique subventionnée partiellement par la National Science Foundation (NSF). Depuis le 1<sup>er</sup>

septembre 2004, Paul Waddell travaille à l'IAURIF à temps partiel en tant que chercheur associé. Sa participation au projet constitue un encouragement fort.

Il s'agit maintenant de rattraper le retard français en ce domaine. On constate à l'heure actuelle un frémissement de la recherche française, comme en témoignent :

- tout d'abord notre projet, SIMAURIF (SimulAtion de l'Urbanisation de la Région Ile-de-France), dont les prémices datent de 2001 à l'IAURIF.
- le projet français SIMBAD (Simuler les MoBilités pour une Agglomération Durable), initié par le LET et l'Agence d'Urbanisme de Lyon et qui devrait démarrer au dernier trimestre 2004. Il vise à développer un modèle liant les transports et l'urbanisation sur l'aire urbaine de Lyon. Une équipe de 8 personnes a été constituée.
- la création récente du laboratoire LVMT (Laboratoire Ville, Mobilité et Transports) en juin 2003, issu d'une coopération entre l'ENPC et l'INRETS. La mission de ce laboratoire est de mieux structurer et renforcer la recherche dans le domaine stratégique de l'interaction entre la planification régionale et les systèmes de transports. Un partenariat avec la RATP est en cours de création. Le Laboratoire comprend 23 chercheurs dont 8 thésards.
- le projet de la DREIF qui consiste à mettre au point un modèle de projection de l'occupation du sol en 250 zones sur la Région Ile-de-France et à le coupler au modèle de trafic MODUS 2.

On peut envisager des collaborations, au moins des échanges d'expériences, entre les cinq organismes cités : IAURIF, THEMA, LET, LVMT et DREIF.

Citons aussi deux projets actuellement en cours en Europe :

- le projet européen, PROPOLIS. Son but est de développer un modèle intégré urbanisation – transport. Cet outil doit permettre de tester des politiques de transport dans le but de définir des stratégies de développement durable et de montrer leurs effets dans sept villes européennes (voir <http://www.ltcon.fi/propolis>). PROPOLIS fait partie du projet global européen LUTR (Land Use and Transportation Policies for the City of Tomorrow).
- Le projet de l'Ecole Polytechnique Fédérale de Zurich (EPFZ) qui vise à développer un modèle intégré utilisant VISUM et UrbanSim appliqué à la Région de Zurich.

## 1.2 Objectif

L'objectif du projet SIMAURIF consiste à élaborer une nouvelle méthodologie permettant de traiter l'interaction au cours de temps entre urbanisation et transport. Il s'agit de concevoir un processus itératif entre la modélisation globale du système urbain et la modélisation de trafic sur des projets de transports, chaque étape produisant des sorties qui seront des entrées pour l'étape suivante. Pour cela, on développera une interface entre le logiciel **UrbanSim** et le logiciel de prévision de trafic **METROPOLIS**, adapté à l'Ile-de-France, développé par le laboratoire THEMA de l'Université de Cergy-Pontoise. Les deux modèles seront d'abord calibrés séparément, puis calibrés dans leur ensemble. En ce qui concerne la demande de déplacements (génération, distribution et choix modal), c'est le modèle de l'IAURIF qui sera porté sur METROPOLIS. Le bouclage entre prévision de l'occupation du sol et prévision de trafic se fera automatiquement selon un pas prédéterminé.

Nous appliquerons le produit intégré à un cas concret : le projet de transport ferré Tangentielle nord entre Sartrouville et Noisy-le-Sec en Ile-de-France. Il s'agit de réaliser une modélisation

interactive urbanisation – transport sur le territoire desservi par cette grande infrastructure ferrée, en nous basant sur un calage prenant en compte toutes les données disponibles pour des simulations sur une période à définir (de 20 à 40 ans).

Une grande partie de notre travail consistera à caler d'abord globalement le couple UrbanSim-METROPOLIS sur l'ensemble de l'Ile-de-France puis de manière très fine sur le territoire desservi par la Tangentielle nord. Nous prendrons comme année initiale 1990 et le ferons tourner d'un pas annuel jusqu'à 1999, année de calage ou année cible. L'expérience d'utilisateurs américains et l'expérience un peu malheureuse de TRANUS à Lyon montrent que cette phase de calage est la plus ardue, la plus longue et la plus délicate. Il ne faudra cependant pas négliger la première phase qui est la phase de recueil, de centralisation et de mise en forme des données. C'est une étape fastidieuse mais cruciale, et dans la plupart des projets de recherche en ce domaine, elle est un peu trop vite expédiée sous le prétexte un peu facile que les données sont incomplètes, hétérogènes voire inexistantes. Nous reconnaissons pour notre part que l'IAURIF dispose de données sur la Région Ile-de-France extrêmement riches, parfois même insoupçonnées, et que si l'on ne parvient pas à constituer des bases de données d'entrée et de calage digne de ce nom, on n'y parviendra nulle part ailleurs. Il s'agit maintenant de les exploiter au mieux et à bon escient, en particulier grâce à l'outil système d'information géographique. On part du principe que les méthodes de simplification habituelles en ce domaine ne peuvent être que l'ultime recours.

L'application de notre méthodologie permettra de confronter les résultats issus de la modélisation interactive et ceux issus d'une approche qualitative plus classique utilisée par les urbanistes pour réfléchir sur l'interaction entre la Tangentielle nord et l'urbanisation. Il est toujours intéressant de relire des articles écrits par des urbanistes pour mieux comprendre leur perception, leur vision de l'urbanisation de la Région d'Ile-de-France. Dans les deux encadrés suivants, on trouvera deux extraits des Cahiers de l'IAURIF (n°102, novembre 1992) :

### **Un milieu urbain vivant, qui se transforme sans cesse**

#### Une alternance d'éruptions et de sédimentations urbaines

Depuis le milieu du siècle, tous les 20 à 30 ans, une poussée massive d'urbanisation autour de la capitale a installé l'un des tissus urbains composant la banlieue d'aujourd'hui.

D'abord, de 1880 à 1910, la mise en place de l'appareil de production industriel et de son accompagnement d'habitat ouvrier ; de 1920 à 1935, la nappe d'extension des grands lotissements pavillonnaires ; puis de 1950 à 1980, le semis des ensembles d'habitat collectif de toutes tailles, aux formes urbaines et aux clientèles variées.

Entre ces vagues successives, des périodes de croissance urbaine ralenties ou arrêtées sont celles où se réparent progressivement les dégâts et les erreurs de la période précédente, où s'effectue une lente maturation qualitative qui transforme plus lentement l'espace urbain. C'est ainsi que des milliers d'hectares de lotissements, créés dans les années 1920 pour la plupart, sans aucun équipement, même de voirie de desserte, ont été progressivement équipés et valorisés au cours des trente années suivantes, au frais des "mal lotis", des communes et de l'Etat.

Enfin, durant l'une et l'autre de ces phases d'évolution, éruption ou sédimentation, un phénomène permanent mais d'amplitude variable, se poursuit : celui de la mutation des espaces urbains existants, qui prend des formes très variables depuis la réhabilitation/extension des pavillons jusqu'aux rénovations "chirurgicales" de grande ampleur.

Mais il faut souligner que ce phénomène de mutation urbaine ne cesse jamais, même dans les périodes de stagnation urbaine les plus marquées et qu'il faut donc en tenir compte dans toutes les prévisions et les projets d'aménagement de l'agglomération.

### Des changements de population et d'activités qui transforment l'espace

Les évolutions démographiques, économiques, sociales, comportementales sont, en effet, les déterminants permanents des transformations de l'espace urbain.

Même en l'absence de toute croissance démographique, à effectif constant, la diminution de la taille moyenne des ménages - elle même liée à des évolutions de comportements familiaux (la recherche de plus de confort : surface, équipement du logement) - entraîne une demande de logements neufs ou de transformations de logements anciens qui sont autant de ferments de transformations des espaces urbains. Toujours à effectif constant, les évolutions locales de population par classe d'âge et composition sociale entraînent également des transformations de l'espace urbain, privatif et collectif.

De même les activités, même lorsqu'il n'y a pas création d'entreprise ni d'emplois nouveaux, sont conduites du fait des évolutions technologiques, à transformer ou à étendre l'espace qu'elles occupent déjà ; les équipements et services publics et privés connaissent les mêmes processus.

Au total, il existe ainsi un courant continu de transformation de l'espace urbain, plus ou moins intense suivant les périodes.

Mais une dynamique de transformation urbaine inéluctable peut constituer un levier très efficace dans les politiques de valorisation de la banlieue.

### Une valorisation qui s'appuie sur le temps et l'action locale

Le temps est l'un des matériaux de base de la transformation et de la valorisation de la ville. Quel qu'en soit le processus, de l'action ponctuelle au grand projet urbain, l'unité de temps est de l'ordre de la dizaine d'années, et il faut bien sûr plusieurs décennies - voire un siècle - pour pouvoir obtenir une transformation globale significative d'une agglomération existante. C'est dans cette perspective temporelle qu'il faut inscrire les projets d'action des collectivités.

**Par Jean-Pierre LECOIN\* (architecte-urbaniste)**

### **A propos d'Orbitale (projet de rocade ferrée en banlieue proche de Paris)**

Orbitale est le support d'un véritable projet urbain de valorisation, de transformation de la couronne : il a été conçu, tracé avec ce principal souci.

Il participe diversement aux mutations urbaines selon les lieux qu'il dessert et peut engendrer différentes transformations :

- il dessert en plusieurs points les zones "stratégiques" qui seront l'objet de profonds changements et de grands projets urbains,
- il conforte ou crée de nouvelles centralités pouvant entraîner des transformations, soit ponctuelles, soit beaucoup plus importantes de type ZAC etc., la réalisation des stations pouvant être traitée en véritables projets urbains,
- il rend plus accessible la majorité des lieux importants de l'agglomération : emploi, équipement, centres-villes,
- il participe au développement de l'économie en offrant de nouvelles vocations à des zones auparavant "hors circuit",
- il rompt l'isolement de certains quartiers jusqu'ici oubliés.

La ville n'est pas figée, elle ne se développe pas seulement par extension, mais aussi par transformation sur elle-même.

Réaliser un transport en commun en zone déjà urbanisée favorisant, dirigeant ces mutations, c'est l'ambition du projet ORBITALE, résolument tourné vers l'avenir.

**Par Jean-Louis PAGES (architecte-urbaniste)**

Au cours de la mise en place de cette application, il ne nous a pas échappé que cette modélisation interactive urbanisation-transport constituait une étape intermédiaire permettant de mettre au point un outil prototype ou une maquette. Dans la phase de création de cet outil intégré et dynamique, nous optons pour une approche basée sur une description axiomatisée des comportements des usagers (approche économique néoclassique de choix de localisation et de mode de transport) et des lois de congestion paramétriques pour le mode voiture particulière. Ces choix peuvent apparaître réducteurs mais constituent, nous semble-t-il, un compromis satisfaisant à l'heure actuelle.

L'outil ainsi créé permettra de tirer des recommandations sur les actions de planification stratégique qu'il conviendrait de mener pour maximiser les retombées de l'investissement de la Tangentielle nord. Ces actions peuvent se référer à la planification spatiale, à la programmation des investissements ou à la mobilisation des différents acteurs concernés.

Quoi qu'il en soit, un des intérêts de notre projet de recherche est de proposer une approche multidisciplinaire permettant d'établir un dialogue entre modélisateurs, ingénieurs des transports, urbanistes et démographes. A l'issue de ce projet, nous devrions être mieux en mesure de déterminer les voies de développement futures nécessaires à la mise en place d'un modèle français d'interaction entre transport et urbanisation.

### ***1.3 Phasage et calendrier***

La première année a été consacrée aux développements méthodologiques et au recueil de données, la deuxième année sera consacrée au calage et à l'application de l'outil au cadre de la Tangentielle nord au travers de multiples scénarios. Les parties nécessitant un développement informatique spécifique sont sous-traitées à la société adpC.

Le projet a été planifié sur deux ans. Mais au terme de la première année de recherche, on se rend compte que la recherche devra sans doute être prolongée au delà des deux années si l'on veut tirer profit au mieux des résultats et valoriser le produit final.

Les différents thèmes de la recherche sont les suivants. Ils ne s'enchaînent pas de manière séquentielle mais sont menés en parallèle :

#### **Année 1**

- Revue de la littérature sur les modèles d'urbanisation
- Recueil des données d'entrée de l'Ile-de-France pour UrbanSim
- Constitution de bases de données de calage à partir de différentes enquêtes
- Etude théorique des différents modules et de la structure de données d'UrbanSim
- Prise en main d'UrbanSim sur une étude de cas américaine
- Mise en place de l'architecture de l'outil intégré SIMAURIF
- Programmation du modèle de demande de l'IAURIF dans METROPOLIS
- Conception de l'interface entre les deux modèles

#### **Année 2**

- Programmation de l'interface
- Calage des différents modules d'UrbanSim
- Estimations économétriques et calage de METROPOLIS sur la Région
- Mise en place de la convergence de l'outil intégré

- Application au cas de la Tangentielle nord, élaboration des scénarios
- Etude de sensibilité des variables
- Présentation des résultats
- Recommandations

Suite à un rapport intermédiaire diffusé en juin 2004, le présent rapport fait le bilan de la première année. Il comporte 6 chapitres :

- 1) Intérêt de l'application à la Tangentielle nord
- 2) Revue de la littérature sur les modèles d'urbanisation
- 3) Présentation du modèle UrbanSim
- 4) Présentation du modèle METROPOLIS et du modèle de demande de l'IAURIF
- 5) Architecture théorique du modèle intégré
- 6) Recueil de données d'entrée et de calage pour UrbanSim

Des annexes techniques complètent le rapport.





offrir une nouvelle offre de transport qui donnerait de nouvelles possibilités de déplacement grâce à son maillage avec des lignes radiales existantes : RER A, B, C, D, E, les lignes banlieue de Paris-St-Lazare, Paris-Nord et Paris-Est, le tramway T1. Il devrait donc améliorer fortement la desserte voyageur des territoires traversés, créant ainsi une dynamique de développement économique.

Le **territoire** directement concerné par ce projet couvre 17 communes (600 000 habitants). Le territoire élargi est constitué d'une quarantaine de communes (1,3 millions d'habitants).

Ces territoires apparaissent aujourd'hui comme un espace urbain socialement défavorisé et peu dynamique au plan économique. Ils sont plutôt cloisonnés et faiblement structurés, fortement dépendants de pôles régionaux qui lui sont extérieurs. Ils disposent pourtant d'un potentiel de régénération important qui pourrait être renforcé par la Tangentielle nord. Il existe en effet :

- des leviers de développement que la Tangentielle peut mettre en réseau contribuant ainsi à une meilleure structuration des territoires,
- un potentiel économique et foncier que la Tangentielle peut rendre plus accessible, permettant une revitalisation économique,
- un patrimoine urbain et naturel que la Tangentielle peut aider à renouveler et à revaloriser.

Le projet constitue donc une occasion exceptionnelle pour enclencher un processus de régénération dans les territoires urbains qu'il dessert.

**Du point de vue de l'exploitation**, il a été envisagé dans une première variante un matériel roulant de type RER circulant sur des voies communes aux trains de fret et de voyageurs et une fréquence minimale de 15 minutes compte tenu de la densité actuelle du trafic fret. Selon les prévisions de trafic présentées dans le Schéma de principe, il y aurait un trafic journalier de 100 000 voyageurs. Le coût total des travaux est évalué à environ 880 millions d'Euros (matériel roulant et travaux d'infrastructure).

Une nouvelle variante est à l'étude depuis 2003 : le tram-train à gabarit réduit (2,65 m de large sur 80 m de long, capacité de 500 personnes contre 2000 pour le type RER). Elle suppose la création de deux voies dédiées aux voyageurs parallèles aux deux voies actuelles exploitées pour le fret. La fréquence serait portée à 5 minutes. La vitesse commerciale est encore en débat (40 à 50 km/h) mais devrait être comparable à celle de matériel RER.

Cette variante aurait l'avantage d'une desserte intensifiée, plus fiable et adaptée aux besoins. Le nouveau matériel roulant serait constitué de rames modernes de type urbain, courtes et rapides, aptes à franchir des fortes rampes et des courbes serrées.

L'avantage de ce système est qu'il permettrait de répondre progressivement à la demande. Les gares seraient conçues pour permettre des évolutions ultérieures en fonction des développements urbains décidés par les collectivités. Si cette solution est retenue, il faudra cependant procéder à des expropriations. Le trafic prévisionnel a été estimé à 130 000 voyageurs, pour un coût total d'investissement de 760 millions d'Euros (acquisitions foncières et protections phoniques comprises), donc comparable à la première variante.

L'avancement actuel des études et la concertation conduisent à repousser à 2012 la date de mise en service initialement prévue à 2009 dans l'ancien Schéma de principe.

## ***2.2 Apport de notre recherche***

La recherche théorique sur l'interaction urbanisation-transport restera sans avenir si des efforts importants de modélisation ne sont pas engagés. On peut s'interroger en effet sur l'existence ou non d'une relation entre urbanisation et transports, mais cela mènerait à des discussions sans fin sur les relations directes, indirectes, de concomitance ou de congruence. Nous avons préféré laisser de côté ce débat pour modéliser sans état d'âme. Dans le cadre de cette recherche, nous réaliserons donc une modélisation interactive de l'interaction sur le territoire desservi par la Tangentielle nord entre Sartrouville et Noisy-le-Sec.

Il est intéressant de s'attaquer à un territoire sur lequel plusieurs études de trafic ont déjà été réalisées depuis 10 ans, avec des résultats souvent contradictoires, mais avec au moins un point commun : la non-prise en compte des effets des transports sur l'urbanisation. Nous nous donnons donc comme objectif de contribuer à la réflexion sur l'avenir de ce territoire et sur l'impact à terme de la Tangentielle nord. Deux séries de scénarios seront élaborées autour des variantes train et tram-train.

Parallèlement, les urbanistes travaillent beaucoup avec des outils juridiques (OPAH, PADD, PLU et loi SRU : voir glossaire) et des opérations ponctuelles de type ZAC du ressort des communes. De plus, il faut souligner le rôle important des aménageurs et promoteurs immobiliers. Il est clair que tous ces éléments peuvent avoir un impact sur l'urbanisme local. Mais dans le cadre de notre recherche, ils ne peuvent être pris en compte qu'au travers d'une approche globale de l'urbanisation. Lorsque cela sera possible, on intégrera les données de la base de données numérisées des POS/PLU appelée BD-POS, gérée par la DREIF.

Les résultats de nos simulations seront confrontés à la vision des urbanistes experts de l'IAURIF et permettront alors de tirer des recommandations sur les actions de planification stratégique qu'il conviendrait de mener pour optimiser au mieux les retombées de l'investissement de la Tangentielle nord.

## 3 Aperçu de la littérature sur les modèles d'urbanisation

### 3.1 Historique

Les modèles d'urbanisation visent à étudier et à prévoir l'affectation du sol entre les différentes activités (résidences, activités économiques et équipements). Ces modèles permettent au planificateur de mieux prévoir les effets des politiques d'urbanisation ou de transports, et de tirer des enseignements quant à l'optimisation des investissements. L'analyse théorique des phénomènes de localisation des agents économiques (entreprises, ménages) est ancienne (von Thünen [41], Christaller [11], Losch [31]). En plus de ces travaux en économie spatiale, les premiers travaux sur les valeurs foncières datent du 18<sup>ème</sup> et du 19<sup>ème</sup> siècle. Les modèles basés sur la théorie de la croissance urbaine se basent sur ces travaux et ainsi prennent le relais des théories qualitatives. Ces modèles théoriques tentent d'expliquer le comportement des ménages et des activités économiques en cherchant un compromis entre la proximité (ou pénibilité de trajet), la disponibilité de terrain et son prix. Parmi ces travaux, celui d'Alonso (Alonso [2]) est probablement plus connu. Certains modèles dans cette catégorie ont aussi tenté de trouver les optimums économiques ou sociaux. Dans le même axe de recherche, nous pouvons citer également les travaux récents de passage des villes monocentriques aux villes ayant plusieurs centres (Arnott [3]).

Par ailleurs, les modèles dits descriptifs tentent de reproduire la réalité en utilisant des analyses statistiques simples telles que les méthodes de régression multiple. Nous recensons parmi ces derniers le modèle de l'Institut d'aménagement et d'urbanisme de la région parisienne (IAURP) élaboré dans les années 60. Il tente de retracer l'évolution du développement de la banlieue parisienne (Merlin [35]). La capacité opérationnelle de ces modèles est à souligner malgré le fait qu'ils ne reposent pas toujours sur des bases théoriques solides. L'outil mathématique de base était la recherche opérationnelle (méthodes d'optimisation), très à la mode dans les années 60.

Rappelons aussi brièvement le modèle de Lowry ([32]) qui a marqué son temps au milieu des années 60. C'est un modèle statique reposant sur des hypothèses simplistes : le choix de localisation résidentielle des ménages est fonction uniquement de l'accessibilité aux lieux de travail et le choix de localisation des entreprises est déterminé par l'accessibilité de la clientèle potentielle.

A ces modèles s'ajoutent les modèles d'équilibre partiels qui se focalisent sur les domaines géographiques et sectoriels particuliers.

A la suite de ces efforts, les modèles comme ceux de l'université de Caroline du Nord (Donnelly et al. [20]) et de Montréal (Laroche [29]) ont réussi à obtenir de meilleurs résultats en ajoutant les notions de modélisation stochastique et les méthodes telles que les simulations de Monte-Carlo.

Au milieu des années 70, l'élaboration des modèles de choix discrets (modèle logit emboîté) a fait avancer le développement des modèles d'urbanisation (Lerman [30]). La nouvelle génération des modèles basés sur ces méthodes rassemble les bases de la théorie économique, les estimations économétriques, la notion de l'incertitude, les phénomènes stochastiques et les aspects dynamiques (pour plus d'informations, nous renvoyons le lecteur à Fujita [24]).

Dans la prochaine section, nous discutons de l'état de l'art en matière de logiciels de modélisation.

### 3.2 Etat de l'art

Les modèles de choix discret constituent aujourd'hui, sans conteste, le cadre mathématique le plus utilisé pour décrire la demande dans les modèles urbains. Ce type de modèle se réfère à la théorie de la maximisation de l'utilité aléatoire élaborée initialement par Daniel McFadden, ce qui lui a valu le Prix Nobel d'économie en 2001. Ces modèles désagrégés sont appliqués pour l'analyse et la prévision du comportement des acteurs. Ils impliquent la définition des lois comportementales et de fonctions d'utilités pour les acteurs qui ont à faire un choix parmi plusieurs alternatives qui sont mutuellement exclusives.

La figure ci-dessous montre les tendances dans le développement des modèles d'urbanisation : prise en compte de l'hétérogénéité des agents et des nombreux paramètres des déplacements (mode, chaîne, tranche horaire), modélisation d'évolution des prix de l'immobilier, prise en compte du cycle de vie des ménages, accessibilités, équilibre partiel etc. C'est dans ce contexte que les modèles de choix discrets ont été appelés à jouer un rôle important. Le diagramme suivant est présenté par Miller [34]. Il a été conçu à la suite des nouvelles réglementations portant sur les études des projets de transports aux Etats-Unis.

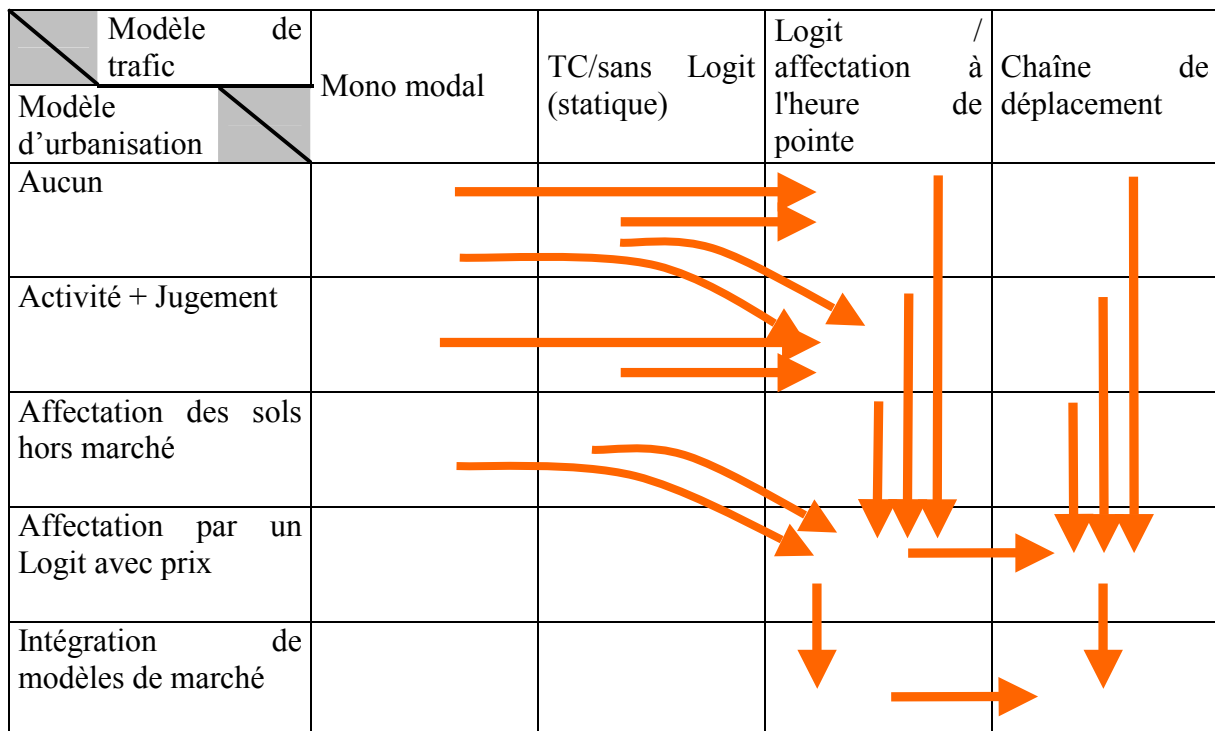


Figure 2 : Diagramme de l'évolution des modèles d'urbanisation et de trafic (source Miller et al [34])

Le logiciel UrbanSim a été conçu dans l'esprit de ce diagramme. Il cherche à satisfaire les exigences de modélisation relatives à l'occupation de sol et à éviter les lacunes des modèles précédents tels que DRAM/EMPAL et TRANUS.

Afin de donner un aperçu de l'état de l'art, nous présentons dans la section suivante, sous forme synthétique, quatre modèles d'urbanisation : MEPLAN, ISHTAR, DRAM/EMPAL et TRANUS. Cette description nous servira de référence pour la description du logiciel UrbanSim, décrit dans la partie 4.

### 3.3 Quatre exemples

Les deux modèles DRAM et EMPAL étaient considérés, au cours des années 90, comme des outils de références aux Etats-Unis. Ils ont été utilisés dans plus d'une vingtaine d'agences de planification régionale (MPOs) aux Etats-Unis. MEPLAN et TRANUS sont deux autres modèles utilisés le monde. En France, TRANUS a été choisi comme le modèle d'urbanisation par l'Agglomération de Lyon (Du Crest [21]). Enfin ISHTAR est un modèle global qui est en cours d'élaboration par un consortium européen. Ce modèle comporte une partie urbaine implantée dans un environnement de type système d'information géographique (SIG).

#### 3.3.1 MEPLAN

MEPLAN est un modèle développé depuis 1978 par l'équipe du Professeur Martial Echenic en Grande Bretagne (<http://www.meap.co.uk/meap/ME&P.htm>). Il est basé sur les principes d'équilibre économique. Il a été conçu sous une forme personnalisable et flexible adaptée à différents projets. Le tableau 1 présente les caractéristiques essentielles de ce modèle (voir Abraham [1]).

Nom	MEPLAN
Origine	Grande-Bretagne
Portée	Études régionales, urbaines
Structure du modèle	Entrées - sorties spatiales, Incrémental
Choix de localisation des ménages	Oui
Choix de localisation des emplois	Oui
Classification des emplois	Agrégé, défini par utilisateur
Développement de l'immobilier	Implicitement en état d'équilibre
Classification de l'immobilier	Agrégé, défini par utilisateur
L'unité de mesure de l'immobilier	Acres, Unités, espace au sol
Prix de l'immobilier	Oui, équilibre
Élément géographique	Zones définies par l'utilisateur
Élément de choix	Agrégé sur la zone
Représentation des pouvoirs publics	Entrées exogènes des politiques tels que le zonage, la modification de réseau de transport ou la tarification
Base temporelle	Équilibre statique, Incréments assez petits de 1 à 5 ans
Modèle de trafic	Choix modal VP/TC : modèle explicite en module séparé
Transport de fret	Possible
Structure modulaire	Non
Applications importantes	Londres, Sud est d'Angleterre, Cambridge (GB), Tokyo (Japon), Sacramento (USA), Modèle national de Suède

Tableau 1 - les caractéristiques de MEPLAN

MEPLAN est basé sur trois principes : l'équilibre des importations et des exportations (Input – Output) des zones, la théorie du maximum d'utilité appliquée aux choix des acteurs et la concurrence spatiale conformément à la théorie d'Alonso. Le choix de localisation des activités primaires se fait dans un premier temps. Les ménages et les secteurs de service se localisent ensuite en fonction des forces de la concurrence et de l'accessibilité. La Figure 3 présente les deux marchés de l'immobilier et des transports pour lesquels MEPLAN cherche à trouver l'équilibre. Le prix de l'immobilier et le coût des transports issus de l'équilibre entre l'offre et la demande ainsi que l'interaction spatiale constituent les éléments de choix de localisation.

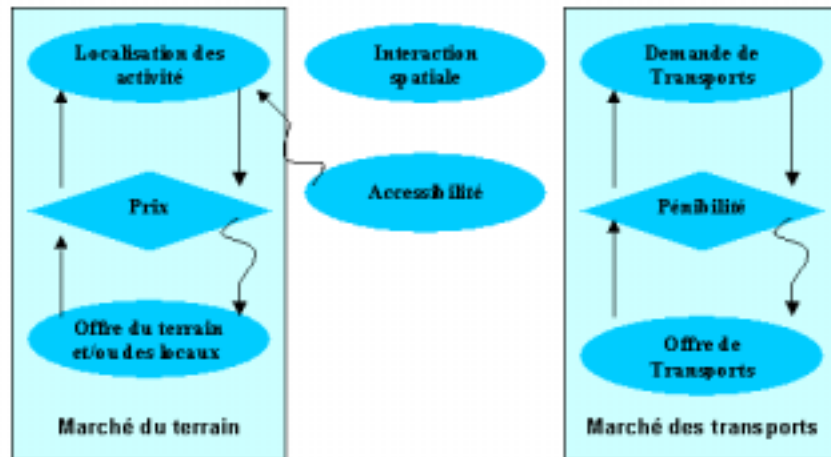


Figure 3 : l'architecture de MEPLAN (adapté de source [1])

### 3.3.2 ISHTAR

ISHTAR est un projet européen, en cours de développement (<http://www.ishtar-fp5-eu.com>). Il vise à modéliser toute la chaîne des causes à effets, depuis les différentes politiques urbaines et économiques jusqu'à l'efficacité du système de transports. L'intervention des différents modèles se fait autour d'une base de données gérée par un SIG.

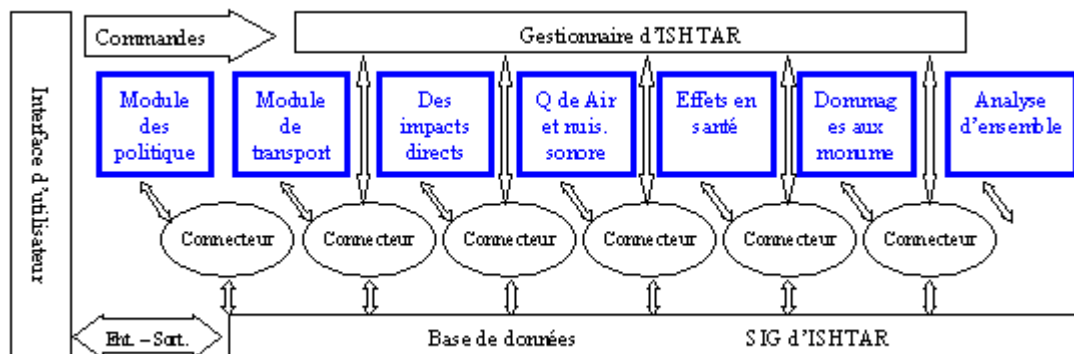
Nom	ISHTAR
Origine	Un consortium européen piloté par l'Agence nationale italienne pour les nouvelles technologies, l'énergie et l'environnement
Portée	Études urbaines
Structure du modèle	Basée sur un Système d'Information Géographique
Processus de modélisation	Les tendances démographiques et économiques au niveau de ville -> évolution de la population et de l'emploi par cellule (localisation des ménages et évolution exogène de l'emploi) -> nouvelle demande de transports -> nouvelles mesures de l'accessibilité et de la pollution
Choix de localisation des ménages	Oui
Choix de localisation des emplois	Non
Classification des emplois	Agrégé
Développement de l'immobilier	Par tendance (exogène)
Classification de l'immobilier	Agrégé

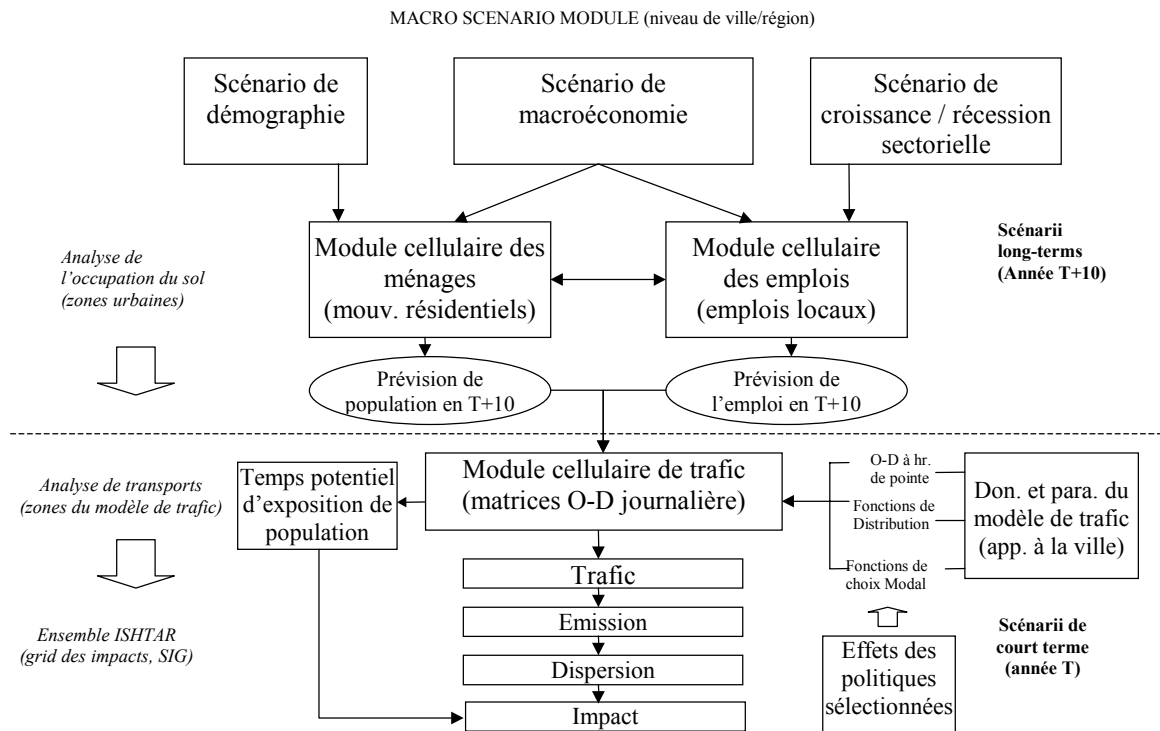
L'unité de mesure de l'immobilier	m2
Prix de l'immobilier	Exogène
Élément géographique	Cellules définies par utilisateur
Représentation des pouvoirs publics	Entrées exogènes des politiques comme plans urbains, zonage, modification de réseau de transport ou tarification
Base temporelle	Extrapolation vers un nouvel horizon temporel
Modèle de trafic	Un modèle de simulation intégré du trafic sur 24 heures. Logiciels VISUM et METROPOLIS.
Transport de fret	Non
Structure modulaire	Non
Applications importantes	Etudes de cas prévues pour Athènes, Bologne, Bruxelles, Graz, Grenoble, Paris et Rome.

Tableau 2: Caractéristiques d'ISHTAR

Ce modèle s'intéresse aux impacts environnementaux (émissions de polluants et de bruit), à leur monétarisation et aux impacts sur la santé (maladies et décès provoqués).

L'architecture d'ISHTAR est présentée sur la Figure 4. Un système d'information géographique joue le rôle de pivot. Les données y sont essentiellement enregistrées et les calculs sont faits dans cet environnement.





### 3.3.3 DRAM/EMPAL

Les modules DRAM (Disagregated Residential Allocation Model) et EMPAL (Employment Allocation model) tournent séparément. DRAM prédit la localisation future des ménages et la probabilité d'occurrence des déplacements domicile – travail entre chaque paire de zones. EMPAL, de son côté, calcule la localisation future des emplois de différents secteurs avec une fonction d'attractivité par zone basée sur une mesure d'accessibilité à la population étendue dans les zones plus ou moins proches. Dans les premières versions du logiciel, ces calculs ont été basés sur les modèles gravitaires puis ils ont évolué vers les modèles de choix discret Logit et Logit emboîté (Southworth [40]).

Les deux modèles DRAM et EMPAL sont intégrés dans un nouveau modèle ITLUP (Integrated Transportation – Land Use Package) ou ultérieurement, METROPILUS qui comprend également le module "Calib". Ce dernier est chargé d'effectuer les estimations des paramètres du modèle par la méthode de maximum de vraisemblance. Ces modèles sont développés depuis le début des années 70. En 1995, ils ont été utilisés par la plupart des organismes chargés de planification urbaine aux Etats-Unis (Putman [38]). Une interface avec le modèle classique à quatre étapes TRANSPLAN a été développée. Le logiciel intégré a été appliqué à la région de Los Angeles en 1994.



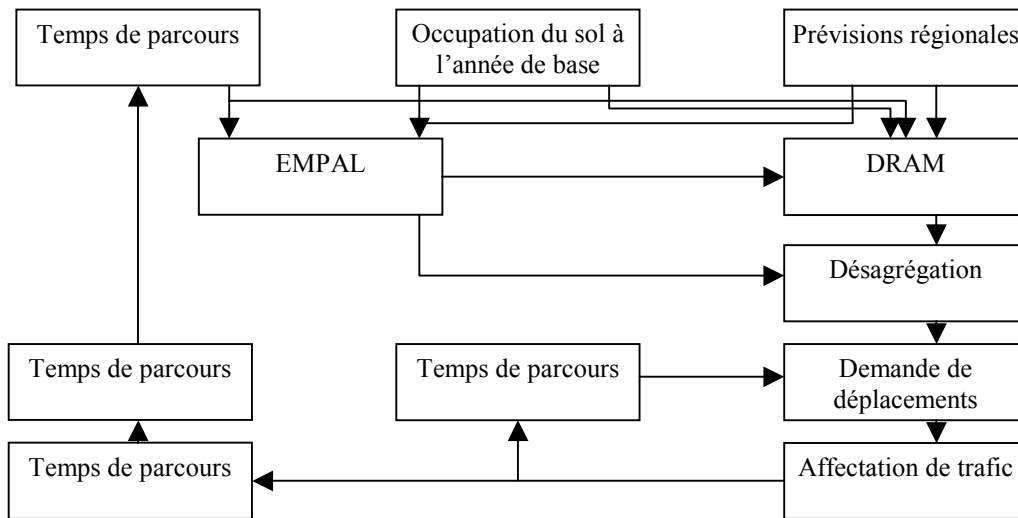


Figure 5 : Diagramme de DRAM / EMPAL (adapté de source [38])

La procédure de modélisation de trafic présentée dans la Figure 5 comprenant la production des matrices origine – destination est mieux détaillée dans le diagramme de la Figure 6. Une procédure semblable doit être développée pour notre projet.

Figure 6 présente le diagramme de fonctionnement de modèle ITLUP.

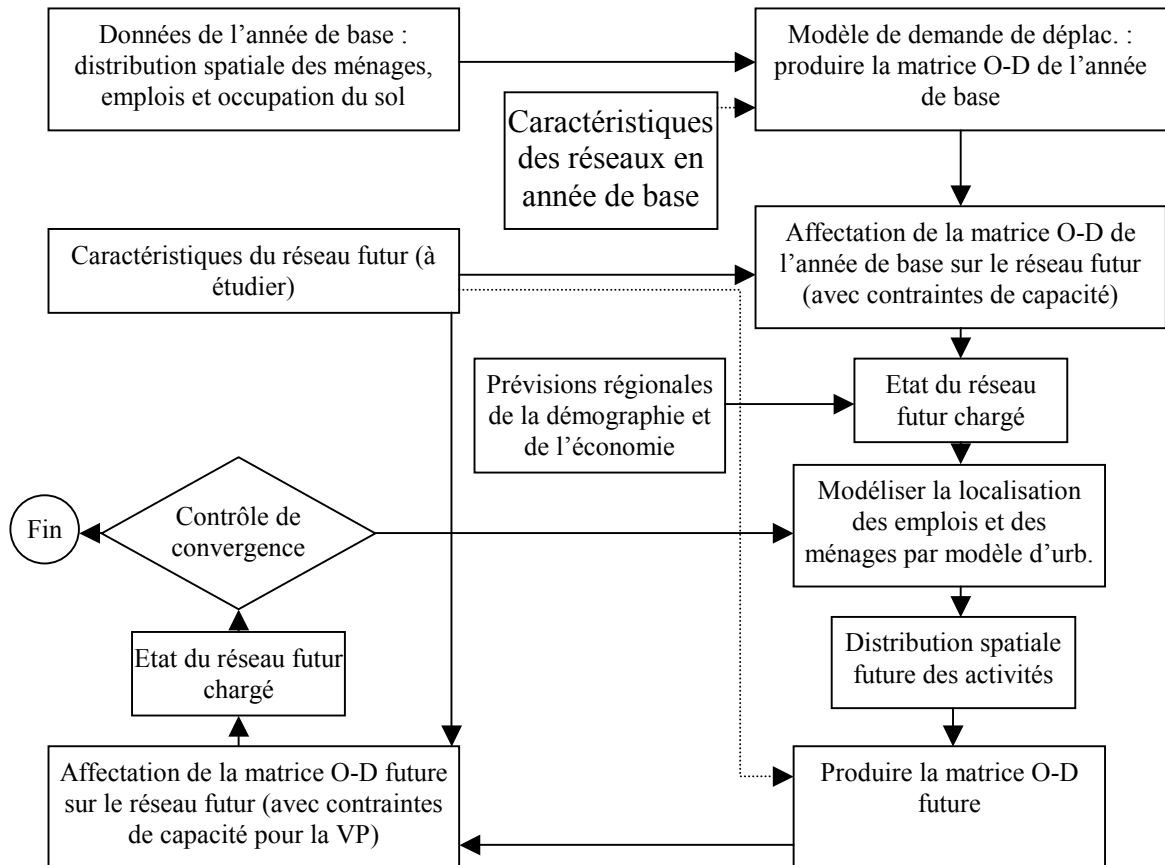


Figure 6 : Diagramme de DRAM / EMPAL (adapté de source [9])

Nom	DRAM / EMPAL
Origine	S. H. Putman Associates, USA
Portée	Études régionales, urbaines, de villes ayant une population de plus de 200 000 habitants
Structure du modèle	Localisation des ménages et de emplois
Processus de modélisation	DRAM et EMPAL fonctionnent séparément. Ils modélisent la localisation des ménages et des emplois. Les autres informations sont considérées comme exogènes par rapport à ces deux modules.
Choix de localisation des ménages	Oui
Choix de localisation des emplois	Oui
Classification des emplois	Par secteur
Développement de l'immobilier	Non
Classification de l'immobilier	Non
L'unité de mesure de l'immobilier	Acres
Prix de l'immobilier	Non
Élément géographique	Zones définies par l'utilisateur
Élément de choix	Agrégré sur la zone
Représentation des pouvoirs publics	Entrées exogènes comme les incréments définis par l'utilisateur
Base temporelle	Incrémental pour l'horizon allant jusqu'à 40 ans avec un pas de 5 ans
Modèle de trafic	Multimodal, multi-opérateur, statique, basé sur l'heure de pointe

Transport de fret	Non
Structure modulaire	Oui
Applications importantes	Colorado Springs, région de Hazelton, Los Angeles (USA)

Tableau 3: Caractéristiques de DRAM / EMPAL

### 3.3.4 TRANUS

Le logiciel TRANUS a été conçu par T. de la Barra et B. Perez. Il est commercialisé par la société Modelistica ([www.modelistica.com](http://www.modelistica.com)) à Caracas au Venezuela et distribué en Europe par la société Stratec en Belgique. L'équilibre de l'offre et de la demande par les prix, la méthode de choix discret de localisation par maximisation de l'utilité (modèle Logit) et le modèle de Léontieff (matrice input-output) pour les échanges économiques constituent ses trois fondements théoriques (de la Barra [14]).

Le modèle calcule à chaque horizon (d'une à plusieurs années) la localisation des ménages et des emplois, l'évolution du prix du foncier, les déplacements, et leurs affectations par mode sur les réseaux. Bien que le logiciel n'ait pas donné entière satisfaction sur l'agglomération de Lyon, la démarche a été jugée intéressante par le CERTU [21]. C'est le logiciel qui se rapproche le plus d'un point de vue théorique du modèle UrbanSim.

Le modèle TRANUS permet de décrire de manière cohérente l'impact des politiques foncières sur la condition de la mobilité et inversement de prédire l'impact des modifications des réseaux de transport et de politique de transport sur le prix du foncier et la localisation des agents (entreprises et ménages).

Nom	TRANUS
Portée	Études régionales, urbaines
Structure du modèle	Entrées - sorties spatiales, Équilibre
Processus de modélisation	Processus de choix du lieu de travail -> lieu de résidence -> lieu d'achat -> mode de transport
Choix de localisation des ménages	Oui
Choix de localisation des emplois	Oui
Classification des emplois	Agrégé, défini par utilisateur
Développement de l'immobilier	Non
Classification de l'immobilier	Agrégé, défini par utilisateur
L'unité de mesure de l'immobilier	m <sup>2</sup>
Prix de l'immobilier	Oui, équilibre
Élément géographique	Zones définies par utilisateur
Élément de choix	Agrégé sur la zone
Représentation des pouvoirs publics	Entrées exogènes comme les incréments définis par l'utilisateur
Base temporelle	Équilibre statique, incréments à définir par l'utilisateur
Modèle de trafic	Multimodal, Multi opérateur, Statique basé sur l'heure de pointe ou flux journalier
Transport de fret	Oui
Structure modulaire	Non
Applications importantes	État d'Oregon, Baltimore (USA), Valencia (Espagne), Bogota (Columbia)

Tableau 4: Caractéristiques de TRANUS

En ce qui concerne la mise au point du modèle intégré UrbanSim – METROPOLIS, nous envisageons, comme dans le modèle TRANUS, de développer un module spécifique permettant de calculer une accessibilité qui sera une entrée du modèle UrbanSim. En effet, l'accessibilité influence les décisions de localisation des agents. De plus, le modèle TRANUS traite la dimension temporelle de la manière suivante : le calcul de l'accessibilité à l'année  $t$  influence les processus de localisation des agents au cours de la période suivante ( $t+1$ ). En conséquence, l'impact d'une modification d'infrastructure pourra nécessiter plusieurs années avant de se faire ressentir complètement dans le système urbain.

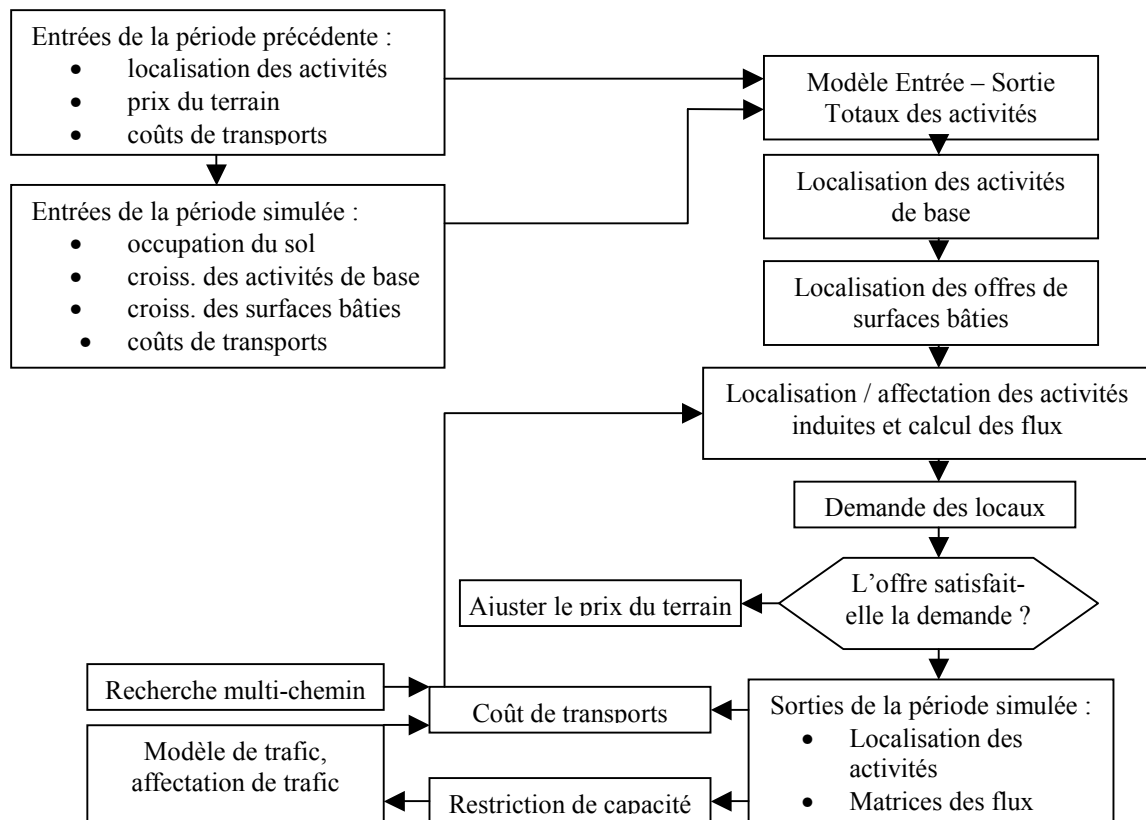


Figure 7 : Diagramme de TRANUS (adapté de source [9], [14])

### ***3.4 Synthèse***

La modélisation de l'urbanisation a débuté par les travaux de Von Thünen et les recherches sur les valeurs foncières au 18ème siècle. Elle s'est poursuivie au cours du 20ème siècle par les travaux des économistes comme Lowry [32] et Alonso [2]. Les modèles théoriques des années 60 sont suivis par les modèles empiriques dans les années 70. Ceux-ci sont allés jusqu'à l'utilisation de maximisation de l'entropie et la méthode de contrôle optimum. Depuis les années 70, l'élaboration des méthodes de choix discret a constitué une avancée majeure dans la modélisation de point de vue théorique et pratique. Enfin, au milieu des années 90, avec la législation américaine (lois sur l'air, ISTEA en 1991 et son extension TEA 21 en 1998) qui exige la prise en compte de l'évolution de l'occupation de sol dans les études de projets de transports, les logiciels existants sont revus et adaptés par des processus dits de « réingénierie ».

La conception d'UrbanSim prend en compte les expériences du passé. C'est un modèle dynamique qui reproduit l'évolution de l'occupation du sol au cours de la période simulée. Le prochain chapitre présentera en détails ce modèle.

## 4 Modèle d'urbanisation UrbanSim

### 4.1 Introduction

UrbanSim est apparu en 1999 aux Etats-Unis comme un nouveau logiciel répondant aux exigences des lois sur l'air ISTEA et TEA21. Ces lois exigeaient que les dimensions relatives à l'urbanisation, aux transports et à l'environnement soient prises en compte dans tout processus de planification régional. Or les modèles de trafic utilisés par les Metropolitan Planning Organizations (MPOs), équivalents des agences d'urbanisme en France, étaient jusqu'alors tous très classiques. Ils n'avaient pas été conçus pour prendre en compte les interactions multiples entre transports, urbanisation et environnement.

UrbanSim est conçu comme un ensemble de composants reflétant les choix des agents économiques (ménages, entreprises) et de leurs interactions avec le marché de l'immobilier et des bureaux en particulier.

Les choix des agents sont par exemple les choix de localisation résidentielle des ménages, les choix de localisation des emplois ou les choix de développement immobilier. Ces modules de localisation sont basés sur un quadrillage fin de l'aire d'étude, les choix des agents sont relatifs à chaque cellule de ce quadrillage. La structure de données d'UrbanSim permet de prendre en compte des données d'occupation du sol à un niveau infracommunal et donc de dépasser des approches classiques trop simplificatrices (découpage concentrique en 3 zones ou découpage par agrégation de communes).

D'autres composants vont simuler la mobilité des ménages et des emplois.

Il y a aussi des composants qui ne sont pas désagrégés, comme les modules de transitions économique et démographique. Ces modules sont agrégés, non spatiaux et liés en fait à des données macro-économiques exogènes. Ils permettront de caler le modèle au niveau global. Enfin, il existe un module de simulation du prix du foncier qui est un modèle de régression linéaire.

Les différents modules s'exécutent sur un pas d'un an, pendant toute la période de simulation (par exemple une période de vingt ans).

### 4.2 L'architecture d'UrbanSim

UrbanSim a une structure modulaire. Les modules ou groupes de modules sont identifiés dans la *Figure 8* par les **cases de t.1 à t.7**.

Les données d'entrée comprennent :

- le « **Data Store** » : « Stockage des données » ou les données de l'année initiale

- le « **Control Totals** » : « Totaux de contrôle », les données de prévisions macro-économiques qui serviront de contrôle ou de calage
- le « **Travel Model Outputs** » : « Sorties du modèle de trafic », les données sur les transports (accessibilité, matrices de temps généralisés de parcours par mode, utilités) issues du modèle de trafic. L'utilité d'un déplacement d'une zone à une autre est basée sur la « log somme » dans le modèle de choix modal, qui incorpore les temps de parcours et les coûts de tous les modes. La mesure de l'accessibilité augmente donc lorsqu'on ajoute des modes, ou lorsque les temps de parcours diminuent.
- le « **Scénario Assumptions** » : « Hypothèses de scénario », les scénarios de politiques d'urbanisation tenant compte des schémas d'aménagement locaux et régionaux.

Le modèle exécute des événements générés par les différents composants. Il est possible pour l'utilisateur d'entrer des événements spécifiques (« **User Specified Events** »).

Le module d'export de données (« **Export Model** ») permet le regroupement, l'agrégation et l'export des résultats dans des fichiers externes destinés à l'analyse et à l'importation dans un système d'information géographique.

Un composant spécial, le « **model coordinator** » ou « Coordinateur des modèles », gère les différents composants individuels et les entrées-sorties avec le Data Store.

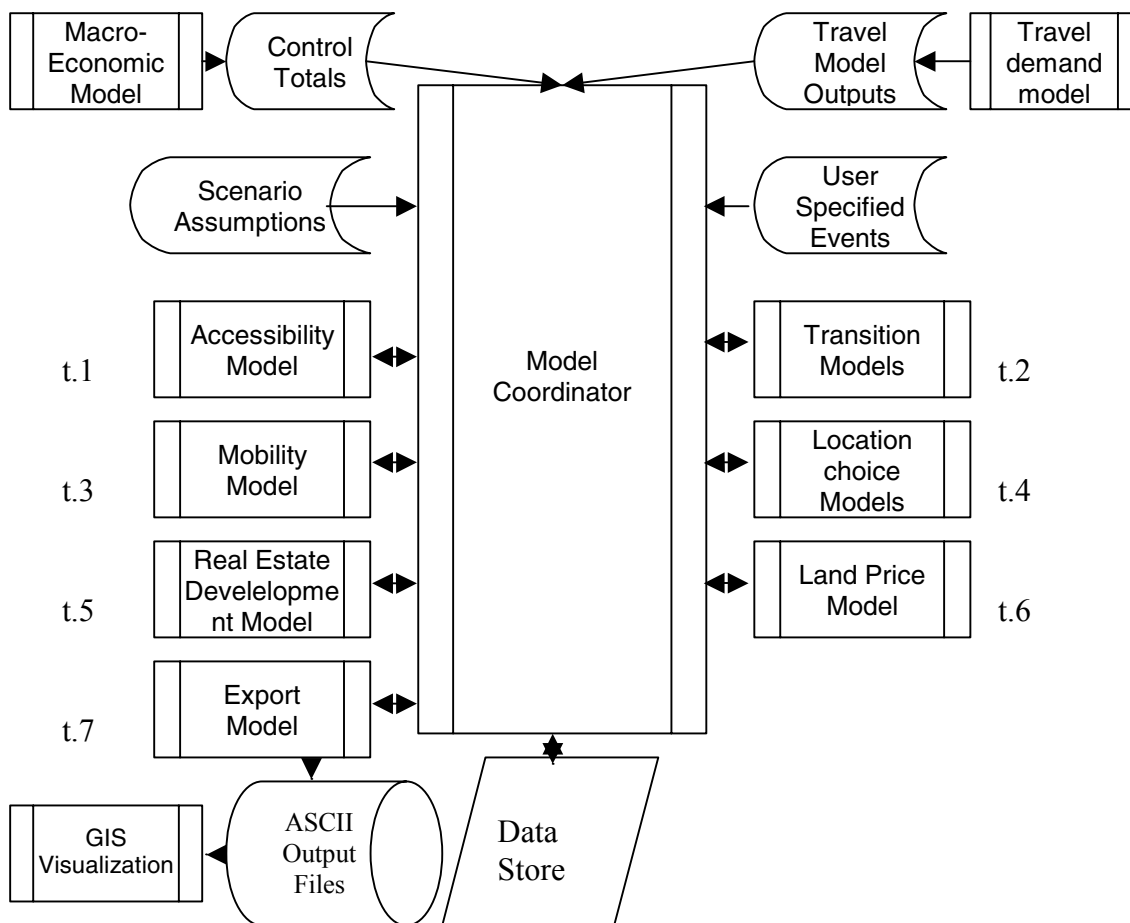


Figure 8 : Architecture d'UrbanSim (source : [www.urbansim.org](http://www.urbansim.org))

### 4.3 Les différents modules

Décrivons de manière plus détaillée les six premiers modules t.1 à t.6, chacun pouvant comporter des sous-modules :

#### 4.3.1 Le module d'accessibilité (« Accessibility model »)

De nombreuses études ont montré l'importance de l'**accessibilité** dans le choix de localisation résidentielle des ménages et des entreprises. La notion d'attractivité d'une zone permet de lui attribuer une valeur économique. La zone est ici une zone du découpage du modèle de trafic. L'attractivité est considérée comme un bien normal (dont la demande augmente avec le revenu), comme tout autre attribut du logement. Les individus attribuent une plus ou moins grande valeur à chaque localisation selon qu'il permet un accès plus ou moins aisé aux emplois et aux zones d'activités. On considère également que la demande pour ce bien varie selon les individus. Par exemple, les retraités n'ont pas le même intérêt pour une zone proche des emplois que les individus actifs.

La définition quantitative de la notion d'attractivité s'effectue à l'aide des variables suivantes :

- $D_j$  : le volume d'activité de la zone  $j$  (nombre d'emplois ou de surfaces commerciales par exemple).
- $L_{aij}$  : la Logsomme des agents qui utilisent un mode de transport  $a$  et qui se déplacent de la zone  $i$  vers la zone  $j$ . Elle définit l'**accessibilité** à la zone  $j$  à partir de la zone  $i$ . Elle est calculée par le modèle de trafic METROPOLIS (voir formule dans le paragraphe 5.2.6).
- $J$  : le nombre total des zones.

On peut maintenant définir l'attractivité de la zone  $i$  :

$$A_i = \sum_{j=1, J} D_j . e^{L_{aij}}$$

Ce calcul est effectué chaque année, à l'aide de la matrice issue du modèle de trafic et des données d'occupation du sol. Cela permet ensuite de créer des indicateurs d'accessibilité qui seront utilisés dans les modules de choix de localisation. Il n'y a donc pas d'interaction intra-annuelle entre transport et occupation du sol puisque en année  $t$ , l'attractivité est calculée en utilisant les données de l'année précédente ( $t-1$ ).

#### 4.3.2 Les modules de transition

Il y a deux sous-modules, l'un concernant l'économie, l'autre la démographie :

##### Le module de transition économique (« employment transition model »)

Le module de transition économique est éclaté selon différents secteurs d'activité. Selon le tissu économique local, UrbanSim permet de définir jusqu'à vingt secteurs différents. Les prévisions agrégées (ou macro-économiques) de l'activité et de l'emploi par secteur sont *exogènes* à UrbanSim.



Ces prévisions sont intégrées dans le module de transition économique pour calculer le taux de croissance de l'emploi par secteur et ainsi créer ou détruire des emplois. Ces emplois sont ensuite, dans le cas des créations, placés dans la base de données, en attente d'affectation par le module de localisation. Pour les emplois détruits, l'espace libéré est déclaré disponible dans la base de données.

Formellement, le module calcule le nombre d'emplois qui doivent être créés ou détruits dans le secteur  $s$ , l'année  $t$  :

$$\Delta J_{st} = C_{st} - |J_{st-1}|$$

avec  $C_{st}$  l'emploi exogène total dans le secteur  $s$ , en  $t$  et  $J_{st-1}$  l'ensemble de tous les emplois du secteur  $s$ , l'année  $t-1$ .

Ainsi,  $J_{st}$  est soit la réunion des emplois de l'année précédente avec les emplois créés soit l'ensemble des emplois de l'année précédente auquel on a soustrait les emplois détruits.

$$J_{st} = J_{st-1} \cup F_{st} \quad \text{si } \Delta J_{st} > 0$$

$$J_{st} = J_{st-1} \cup \emptyset \quad \text{si } \Delta J_{st} = 0$$

$$J_{st} = J_{st-1} \setminus F_{st} \quad \text{si } \Delta J_{st} < 0$$

$$J_{st} \subset J_a$$

où  $F_{st}$  est l'ensemble des flux d'emplois et  $J_a$  l'ensemble de tous les emplois possibles. Les emplois en flux sont soit pris aléatoirement dans  $J_a$  (création), soit tirés aléatoirement dans  $J_{st}$  (destruction).

On a maintenant un ensemble  $U_t$  d'emplois non localisés l'année  $t$ . Chaque année, si l'on crée des emplois, on a :

$$U_t = U_{t-1} \cup F_{st}$$

Si des emplois sont détruits, on doit les extraire de l'ensemble des emplois  $j$ , localisés en  $l$  (nommé  $P_t$ ):

$$P_t = P_{t-1} - \{(j, l) \in P_t \mid j \in F_{st}\}$$

De même, les emplacements précédemment occupés par des emplois relocalisés doivent être déclarés vacants.

### Le module de transition démographique (« household transition model »)

Ce module détermine les changements de la distribution des ménages, par type, selon le même mécanisme que le module de transition économique. Ces changements sont le résultat de mécanismes complexes, incluant, par exemple, mortalité, natalité, ... La théorie permettant de rendre compte de ces processus n'est pas décrite au sein d'UrbanSim qui se contente actuellement, comme pour le précédent module, de données exogènes pour déterminer ces changements.

Comme dans le cas du module de transition économique, les ménages créés sont placés dans une base de données, en attente de localisation. De même, l'espace libéré par les ménages détruits est rendu disponible.

Remarquons que UrbanSim considère comme agent le ménage et non pas l'individu. Dans la réalité, le choix de localisation résidentielle résulte d'un compromis entre les choix des différents membres d'un ménage. En revanche, UrbanSim prend en compte les caractéristiques du ménages (nombre d'actifs, nombre d'enfants, etc).

### 4.3.3 Les modules de mobilité

Il y a deux sous-modules :

La mobilité de l'emploi (« employment relocation choice model », ancien nom : « employment mobility model »)

La délocalisation des emplois est une décision prise par l'entreprise. L'unité d'analyse choisie est l'emploi individuel et non pas une entreprise. Le module détermine la probabilité que chaque emploi soit déplacé d'une cellule à l'autre, à l'issue d'une année  $t$ . Ces transitions reflètent la conjoncture sur le marché du travail. Les emplois ainsi déplacés seront relocalisés par le module de localisation (voir paragraphe suivant). L'espace libéré est déclaré vacant dans la base de données. Bien qu'il semble naturel que la décision de mobilité soit conditionnelle à l'espace de relocalisation (modèle imbriqué, conditionnel à la localisation), cela n'est pas pris en compte ici. En effet, cela pose des problèmes en termes de calage des données. Le choix de mobilité est considéré comme indépendant du choix de localisation.

Le modèle de mobilité de l'emploi fonctionne alors de la façon suivante : l'ensemble des emplois déplacés ( $M_{st}$ ) est déterminé par tirage aléatoire dans l'ensemble  $J_{at}$ , à l'aide du taux de mobilité exogène dans le secteur  $s$ .

La mobilité des ménages (« household relocation choice model », ancien nom : « household mobility model »)

La décision de mobilité des ménages est déterminée de la même manière que pour les emplois. Des taux exogènes de mobilité par type de ménage sont introduits, à partir des données démographiques.

### 4.3.4 Les modules de localisation

La localisation des emplois et des ménages est modélisée par deux sous-modules :

Le module de localisation de l'emploi (« employment location choice model »)

Ce module calcule la probabilité qu'un emploi nouveau (créé dans le module de transition économique) ou relocalisé (mobilité économique) soit localisé dans une cellule donnée. Chaque cellule contient un espace donné qui peut être occupé par des emplois. Cet espace total disponible pour l'emploi dépend de l'espace non résidentiel et de la densité d'usage. Il est également possible de localiser des emplois dans des zones résidentielles. L'ensemble des localisations disponibles de la zone  $l$ ,  $L_t^l$  ayant le type urbain  $d$ , a pour cardinal :

$$|L_t^l| = \frac{s_l}{\tau_{sd}} + \frac{h_l}{\tau_{hd}}$$

avec :

- $s_l$ : l'espace total non résidentiel en  $l$  (donnée dans la table « gridcell », voir paragraphe 4.5)
- $\tau_{sd}$ : le taux d'utilisation de l'espace dans type urbain  $d$  (donnée dans la table « sqft\_for\_non\_home\_based\_jobs »)
- $h_l$ : le nombre de logements en  $l$  (donnée dans la table « gridcell »)

- $\tau_{hd}$ : le taux de travail à domicile dans le type urbain  $d$  (donnée dans la table « residential\_units\_for\_home\_based\_jobs »).

Dans ce module de localisation des emplois (ainsi que pour la localisation des ménages), le stock d'espace disponible est fixé dans la période intra-annuelle de simulation. De plus, les agents sont preneurs de prix.

Parmi les variables pouvant influencer le choix de localisation, on trouve :

- l'accessibilité
- les externalités intra et inter-sectorielles
- l'âge des immeubles
- la densité des infrastructures
- la proximité d'un axe de voirie
- le niveau de desserte de la zone en transports en commun : desserte spatiale (présence d'une gare), desserte temporelle (fréquence)
- les prix.

L'ensemble de ces variables explicatives est introduit dans un modèle multinomial Logit (éventuellement un par secteur d'activité) afin de déterminer la probabilité qu'un emploi (placé aléatoirement dans une file d'attente) se localise à un endroit donné. Une fois cet emploi localisé, l'espace correspondant est supprimé et le module itère le processus pour l'emploi suivant en attente.

#### Le modèle multinomial Logit (ou la régression logistique)

Le modèle multinomial logit est la forme mathématique la plus utilisée dans les modèles de choix discrets. Il est basé sur l'hypothèse qu'un agent  $m$  (par exemple un ménage) associe une utilité  $U_{i,m}$  à chacune des alternatives et qu'il choisira alors celle qui lui procure la plus grande utilité. Cette utilité se décompose en une utilité déterministe  $u_{i,m}$  et en un terme aléatoire non observable  $\varepsilon_{i,m}$ , censé prendre en compte les erreurs de mesure relatives à la partie déterministe de l'utilité et les facteurs non observés.

On peut donc exprimer l'utilité stochastique par :

$$U_{i,m} = u_{i,m} + \varepsilon_{i,m} = \beta' X_i + \varepsilon_{i,m} \quad , \quad i \in \{\text{ensemble des alternatives}\}$$

où  $X_i$  représente le vecteur des *caractéristiques* du choix de l'alternative  $i$  et  $\beta$  représente le vecteur des coefficients à estimer par la méthode du maximum de vraisemblance.

Les termes d'erreur  $\varepsilon_{i,m}$  sont supposés distribués selon la double exponentielle, dont la fonction de répartition est donnée par :

$$F(w) = \frac{\exp(w)}{1 + \exp(w)}$$

La probabilité pour le ménage  $m$  de choisir l'alternative  $i$  est alors donnée par le modèle logit :

$$P_{i,m} = \frac{\exp(u_{i,m})}{\sum_j \exp(u_{j,m})}$$

Cet outil mathématique (et ses extensions) a été utilisé dans une grande série de problèmes de choix discret : choix de localisation résidentielle, de localisation des emplois, de développement immobilier, de choix modal ou de taux de motorisation, ...

Pour plus de détails, on se référera à McFadden [33].

### Le module de localisation des ménages (« household location choice model »)

Ce module a une structure similaire à celle du sous-module précédent. Grâce à un modèle logit multinomial, ce sous-module prédit la probabilité qu'un ménage nouveau ou déplacé se localise dans une cellule donnée. Il est possible de considérer différents types de ménages, selon le revenu, le nombre d'enfants ou la catégorie socio-professionnelle. On peut, à l'opposé, spécifier un seul modèle dans lequel ces variables interagissent avec les caractéristiques de logements. C'est cette approche qui est utilisée par UrbanSim.

Les variables explicatives, caractéristiques de la cellule, pourraient être :

- le coût du transport et du terrain (reflétant l'arbitrage entre les deux biens dans le modèle monocentrique)
- la densité urbaine
- le revenu moyen de la zone (effet d'imitation)
- les caractéristiques du logement (en particulier la disponibilité de places de stationnement privé) et de l'environnement
- le niveau de desserte de la zone en transports en commun : desserte spatiale (présence d'une gare), desserte temporelle (fréquence)

Caler ce type de modèle nécessite donc de travailler avec des démographes afin d'intégrer les variables pertinentes. Par exemple, les démographes nous disent qu'il existe deux moments du cycle de vie des ménages qui sont déterminants pour décider des investissements logement : la naissance des enfants et leur décohabitation. Plus généralement, les choix de déménagement sont guidés par le cycle de vie d'un ménage (mariage, naissance, divorce, décès ...). Une autre tendance est qu'autrefois, les Franciliens quittaient l'Ile-de-France pour la province au moment de leur retraite ou quelques années après. Mais cela est de moins en moins vrai. La mobilité résidentielle décroît avec l'âge. Le vieillissement des générations du baby-bom devrait donc entraîner une baisse de la fréquence des déménagements en Ile-de-France. Cette tendance pourra être modélisée car on dispose dans le fichier ménages du Recensement ou de l'Enquête Globale Transport de l'âge de la personne de référence du ménage.

Le logiciel UrbanSim étant relativement flexible, on pourra donc introduire des variables concernant les caractéristiques du ménage au moment du déménagement.

### **4.3.5 Le module de développement immobilier ou urbain (« real estate development model »)**

Ce module détermine la probabilité de réalisation de programmes immobiliers, que ce soit de la création ou de la réhabilitation. Chaque cellule, selon sa superficie affectée aux espaces non résidentiels et selon le nombre d'unités d'habitations, appartient à un type de développement (ou type urbain) donné.

Les données nécessaires à l'estimation du module de développement sont issues du calcul, pour chaque cellule et pour chaque année de son historique, de l'ensemble des transitions entre types de cellules. Le module peut ensuite estimer, à l'aide d'un modèle multinomial Logit, la probabilité qu'une cellule modifie son type (c'est à dire qu'un programme immobilier

permettant de passer d'un type à un autre se réalise). Ce changement est ensuite déterminé par un tirage aléatoire. La modification éventuelle est introduite dans la base de données. Un certain nombre d'éléments peuvent venir contraindre les changements de type : plans d'occupation des sols, contraintes physiques, techniques ou environnementales.

Les variables explicatives rendant compte du choix de type urbain sont les suivantes :

- les caractéristiques du site : programme de développement en cours, POS, ...
- caractéristiques urbaines,
- maillage routier,
- proximité d'un programme immobilier,
- espaces verts,
- caractéristiques régionales,
- temps de parcours vers les aéroports, vers des pôles d'emplois tels que la Défense.
- temps de parcours vers Paris,
- accessibilité,
- caractéristiques du marché,
- taux de surfaces résidentielles,
- prix du logement
- ...

Pour caler ce type de modèle par des régressions logistiques, il ne faut pas travailler comme des statisticiens purs et durs mais intégrer la vision des urbanistes en essayant de la formaliser en prenant en compte les « bonnes » variables. Par exemple, les urbanistes ont en général une approche des mutations de l'occupation du sol basée sur l'expérience et la connaissance du terrain. Ainsi, les mutations sont pour eux concentrés sur les espaces déjà très urbanisés et bien desservis par les TC. Elles prennent en général deux formes : une densification des zones d'habitat et des changements d'usage des espaces d'activités traditionnelles (exemple d'Issy-les-Moulineaux où les emprises industrielles se sont transformé en ZAC au fur et à mesure du desserrement parisien, ou encore exemple de l'opération de la Défense où une tertiarisation et un urbanisme volontariste et ambitieux du pôle d'affaires ont été décidés par l'Etat pour en faire un "Manhattan français", ce qui a permis aussi de structurer les mutations de ses environs).

Par ailleurs, il faut noter que les échelles de temps ou les bases temporelles sont très différentes entre le modèle de développement immobilier et les modèles de localisation des ménages et des emplois. Pour le premier, il faut plusieurs années, parfois une vingtaine d'années avant qu'un changement notable de type urbain soit constaté dans un quartier, tandis qu'un ménage déménage en moyenne tous les cinq-six ans en Ile-de-France. Cette réalité de la dynamique urbaine devra être bien reconstituée.

### 4.3.6 Le module de fixation du prix de terrain (« land price model »)

Les prix sont des indicateurs de l'adéquation entre l'offre et la demande de terrain, entre les différentes zones et les différents types. Les prix expliquent en partie les choix de localisation des entreprises et des ménages. Evidemment, il y aura toujours des sociétés qui recherchent des adresses de prestige pour leur « image » sans regarder les prix. Ce sont des cas particuliers non modélisables. Mais généralement, des prix trop élevés en centre-ville poussent les entreprises à chercher de nouveaux locaux en périphérie.

Parmi les variables à prendre en compte dans le module, il y a les variables d'accessibilité et de desserte. On sait que l'implantation d'une nouvelle gare ou station de transports en commun lourds augmente significativement les prix de l'immobilier dans les environs proches.

Les hypothèses suivantes sont prises :

- L'ensemble des agents est preneur de prix et les ajustements sont faits en réponse à la demande et l'offre agrégées.
- Les préférences et les déséquilibres entre offre et demande se reportent dans la valeur du terrain. La valeur d'un immeuble reflète uniquement son coût de remplacement.
- Il existe un taux de vacance structurel de long terme pour chaque type de propriété et la relation entre le taux de court terme et le taux de long terme influence le prix des propriétés.
- Le prix du terrain est déterminé par une approche en termes de prix hédoniques. Cela permet de capturer les interactions entre la demande et l'offre. Du fait que la régression hédonique utilise des variables actualisées par les différents modules, on peut également actualiser les prix de cette façon. En plus de cette détermination hédonique des prix relatifs, le prix absolu s'ajuste aux variations de l'offre et de la demande. Ces fluctuations sont reliées à la différence entre le taux de vacance structurel  $V_i^s$  et le taux de vacance courant  $V_{it}^c$ .

L'équation hédonique est alors :

$$P_{ilt} = \alpha + \delta \left( \frac{V_i^s - V_{it}^c}{V_i^s} \right) + \beta X_{ilt}$$

où  $P_{ilt}$  est le prix unitaire du terrain de type  $i$ , dans la zone  $l$ , l'année  $t$  et  $X_{ilt}$  un vecteur d'attributs de la localisation. Les prix sont actualisés annuellement, en fin de période d'activité et utilisés ensuite comme références pour l'année suivante.

Les caractéristiques entrant dans l'équation ci-dessus pourraient être :

- les caractéristiques du site : type, POS, contraintes environnementales
- accessibilité et desserte
- caractéristiques urbaines
- maillage, densité
- conditions du marché (taux de vacance)

Pour le cas de la Région Ile-de-France, il serait plus correct de modéliser le prix du bâti que le prix du terrain. Par ailleurs, on sera sans doute confronté à la difficulté de prise en compte des

cycles de spéculation immobilière (par exemple, la période 1988-1991) qui dépendent beaucoup de la croissance économique.

#### ***4.4 Exemple : choix de localisation d'un emploi***

On connaît le nombre total d'emplois par secteur sur l'ensemble de la Région à l'année de base n (par exemple 1990).

On connaît à l'année n+1 (1991) :

- le nombre d'emplois nouveaux, par la table « ANNUAL\_EMPLOYMENT\_CONTROL\_TOTALS » (voir paragraphe 7.2.5 pour la description des tables)
- le nombre d'emplois qui vont être relocalisés parmi les emplois existants à l'année n, par la table « ANNUAL\_RELOCATION\_RATES\_FOR\_JOBS »

Il s'agit de localiser les emplois nouveaux et les emplois relocalisés.

Pour un emploi i à localiser dans un secteur donné, on construit l'ensemble des cellules pouvant l'accueillir. Pour cela, on calcule pour chaque cellule le nombre d'emplois du secteur par la formule calculant  $L'_i$ , donnée au paragraphe 4.3.4.

Parmi ces cellules, on tire un échantillon aléatoire dont la taille est donnée par le paramètre de fraction fixé dans la table « SAMPLING\_RATES » .

Pour chaque cellule j de cet échantillon, on calcule l'utilité  $U_j$  de la cellule (pour l'emploi i) par une fonction linéaire dont les variables sont dans la table « EMPLOYMENT\_LOCATION\_CHOICE\_MODEL\_SPECIFICATION » et les coefficients pondérateurs dans la table **EMPLOYMENT\_LOCATION\_CHOICE\_MODEL\_COEFFICIENTS** ».

Puis on calcule la probabilité que l'emploi i se localise dans la cellule j par un logit multinomial : 
$$P_i = \frac{\exp(U_i)}{\sum_j \exp(U_j)}$$

L'affectation de l'emploi i dans une des cellules pour lesquelles la probabilité est non nulle est réalisée par une méthode de Monte-Carlo.

#### ***4.5 Les données d'entrée***

Les données d'entrée sont de quatre types :

- les données de base du modèle
- les constantes
- les données des scénarios
- les paramètres des différents modules

Ces données sont structurées en fichiers au format texte (.tab). Le Tableau 5 récapitule l'ensemble des fichiers à construire :

<b>Nom du fichier</b>	<b>Type de fichier</b>
DevelopmentEvents.tab	Scénario
EmploymentEvents.tab	Scénario
DevelopmentConstraints.tab	Scénario
LandUseEvents.tab	Scénario
DevelopmentConstraintEvents.tab	Scénario
TransitionTypes.tab	Scénario
AnnualEmploymentControlTotals.tab	Données
AnnualHouseholdControlTotals.tab	Données
GridCells.tab	Données
Households.tab	Données
Jobs.tab	Données
Zones.tab	Données
HouseholdCharacteristics.tab	Données
GridInfo.tab	Données
TravelData.tab	Données
RaceNames.tab	Constantes
SqftForNonHomeBasedJobs.tab	Constantes
Cities.tab	Constantes
DevelopmentTypeGroupDefinitions.tab	Constantes
DevelopmentTypes.tab	Constantes
PlanTypes.tab	Constantes
UrbanSimConstants.tab	Constantes
EmploymentAdHocSectorGroupDefinitions.tab	Constantes
Counties.tab	Constantes
DevelopmentTypeGroups.tab	Constantes
EmploymentAdHocSectorGroups.tab	Constantes
ResidentialUnitsForHomeBasedJobs.tab	Constantes
EmploymentSectors.tab	Constantes
AnnualRelocationRatesForHouseholds.tab	Paramètres
AnnualRelocationRatesForJobs.tab	Paramètres
HouseholdLocationChoiceCoefficients.tab	Paramètres
DeveloperModelCoefficients.tab	Paramètres
EmploymentLocationChoiceCoefficients.tab	Paramètres
ResidentialLandShareCoefficients.tab	Paramètres
LandPriceCoefficients.tab	Paramètres

*Tableau 5 : Liste des fichiers de la base de données d'entrée*

Le secteur d'étude doit être découpé en mailles ou cellules carrées, les plus fines possibles, chaque cellule est identifiée par une clé unique (GRID-ID).

Les données d'entrée à UrbanSim sont à recueillir dans ce maillage, pour l'année initiale (par exemple 1990) et l'année de calage (par exemple 1999).

Donnons 3 exemples de structures de table. Ces tables seront construites pour le cas de la Région Ile-de-France (voir partie 7).



- Le Tableau 6 présente la structure du fichier **gridcell.tab**. Ce fichier comporte pour chaque cellule du maillage de l'année initiale et de l'année de calage (ou actuelle) les **correspondances avec d'autres découpages, des attributs d'occupation du sol, des valeurs foncières, des contraintes environnementales.**

<b>Champ</b>	<b>Type</b>	<b>Unité</b>	<b>Valeur</b>
GRID_ID	Entier		
RELATIVE_X	Entier		
RELATIVE_Y	Entier		
ZONE_ID	Entier		
CITY_ID	Entier		
COUNTY_ID	Entier		
PLAN_TYPE_ID	Entier		
IS_OUTSIDE_URBAN_GROWTH_BOUNDARY	Entier		
FRACTION_RESIDENTIAL_LAND	Flottant		
RESIDENTIAL_UNITS	Entier		
GOVERNMENTAL_SQFT	Entier		
INDUSTRIAL_SQFT	Entier		
COMMERCIAL_SQFT	Entier		
RESIDENTIAL_IMPROVEMENT_VALUE	Entier		
RESIDENTIAL_LAND_VALUE	Entier		
NONRESIDENTIAL_LAND_VALUE	Entier		
GOVERNMENTAL_IMPROVEMENT_VALUE	Entier		
INDUSTRIAL_IMPROVEMENT_VALUE	Entier		
COMMERCIAL_IMPROVEMENT_VALUE	Entier		
YEAR_BUILT	Entier		
DEVELOPMENT_TYPE_ID	Entier		
DISTANCE_TO_ARTERIAL	Entier		
DISTANCE_TO_HIGHWAY	Entier		
PERCENT_WATER	Entier		
PERCENT_STREAM_BUFFER	Entier		
PERCENT_FLOODPLAIN	Entier		
PERCENT_WETLAND	Entier		
PERCENT_SLOPE	Entier		
PERCENT_OPEN_SPACE	Entier		
PERCENT_PUBLIC_SPACE	Entier		
PERCENT_ROADS	Entier		

*Tableau 6 : caractéristiques d'une cellule*

- Le Tableau 7 montre la structure du fichier **jobs.tab**. Il contient la localisation de chaque **emploi** dans le maillage en cellules carrées pour l'année initiale et l'année de calage (ou actuelle).

<b>Champ</b>	<b>YEAR</b>	<b>JOB_ID</b>	<b>GRID_ID</b>
Type	Entier court	Entier	Entier
Valeurs			
	(...)		

*Tableau 7 : localisation des emplois*

- Enfin, le Tableau 8 montre la structure du fichier **household.tab**. Il contient la localisation de chaque **ménage** pour l'année initiale et l'année de calage (ou actuelle).

Champ	YEAR	HOUSEHOLD_ID	GRID_ID
Type	Entier court	Entier	Entier
Valeurs			
	(...)		

Tableau 8 : localisation des ménages

## 4.6 Les paramètres des modules et leurs estimations

### 4.6.1 Le calage des paramètres

Le calage d'UrbanSim consiste à estimer les paramètres du modèle à partir de données désagrégées sur une période constituée de deux années pour lesquelles il existe le maximum de données. Le modèle sera ensuite appliqué, avec les mêmes paramètres, à un horizon futur selon différents scénarios à définir.

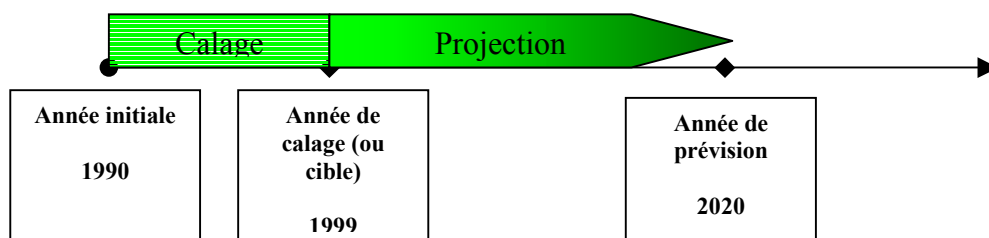


Figure 9 : cours du temps dans UrbanSim

Nous avons hésité à choisir comme année initiale 1982 (année de recensement). Du point de vue calage, cela aurait bien sûr présenté un avantage car la période 1982-1999 est suffisamment longue pour pouvoir observer des changements notables de l'occupation du sol. Le problème est que l'IAURIF ne dispose pas de données complètes sur l'année 1982, contrairement à l'année 1990. La question a donc été rapidement tranchée.

Comme on l'a dit dans la partie 4.3, la base de données d'entrée d'UrbanSim est constituée des données d'entrée ( $x_i$ ), des paramètres ( $\alpha_r$ ), certains constantes et les définitions des scénarios à appliquer. Les données d'entrée décrivant la situation des ménages, des emplois et l'occupation du sol servent à calculer des variables qui, avec les paramètres, définissent des fonctions internes d'UrbanSim ( $U_{k=1,K}$ ) du type suivant :

$$x_{i=1,I} \rightarrow v_{j=1,J}(x_{i=1,I}) \rightarrow U_{k=1,K}(\alpha_{r=1,R}, v_{i=1,I})$$

Les paramètres du modèle, dans un premier temps, doivent être estimés par l'utilisateur avec un logiciel externe d'économétrie (SAS, LIMDEP, ALOGIT, STATA ...). Pour chaque module, les informations sont écrites dans un fichier texte. Chaque ligne de ce fichier comprend un identifiant, la valeur de la variable dépendante et les valeurs des variables indépendantes. Il est possible de constituer l'échantillon d'estimation de manière personnalisée ce qui nous permet d'utiliser les ménages pour lesquelles nous disposons des informations plus détaillées et plus fiables. Les valeurs des paramètres seront prises en compte pour l'année de base par UrbanSim en incluant la section <create-estimation-data-tables/> dans le fichier des paramètres. Ces valeurs estimées seront ensuite ajustées par une procédure d'optimisation.

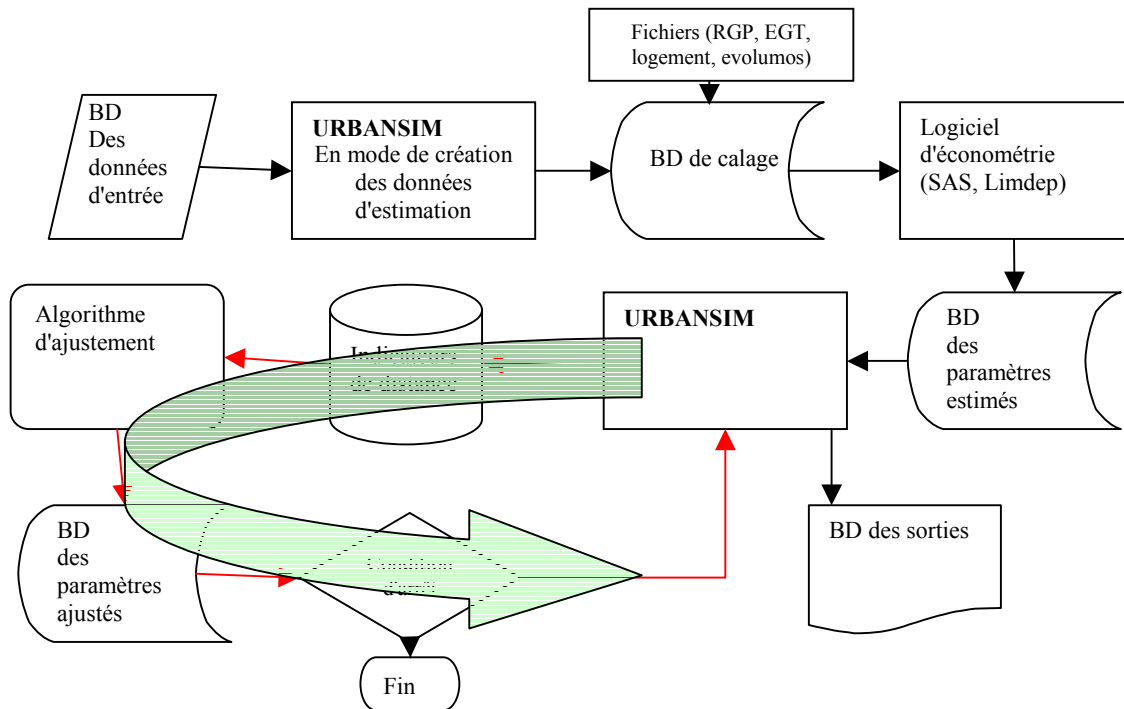


Figure 10 : la procédure de calage d'UrbanSim

La Figure 10 présente ce processus que nous allons développer par la suite.

La première phase de calage consistera à construire des fichiers de calage, c'est-à-dire **d'observations désagrégées** issues d'enquêtes ménages et logements, du Recensement et de données sur l'évolution du mode d'occupation du sol.

#### 4.6.2 Estimation

Il y a trois groupes de paramètres :

- **Les paramètres des distributions statistiques**

Ce sont les taux de transitions démographique et économique et les taux de relocalisation qui seront calculés en fonction des totaux de contrôle.

- **Les paramètres des régressions linéaires**

Ce sont notamment les coefficients du modèle hédonique du prix de terrain.

- **Les paramètres des modèles de choix discrets**

Ils comprennent ceux des modèles de choix de localisation et ceux des modèles de développement immobilier.

La majorité des paramètres à estimer se trouve dans le troisième groupe. Leur estimation nécessite effort de modélisation et un grand volume de calculs. En plus, dans le cas où nos données sont recueillies pour deux années distantes, les valeurs des variables explicatives ne correspondront plus aux années de choix. Cela est une source d'erreur supplémentaire. Par exemple, dans le cas de notre aire d'étude (la région Ile-de-France), nous avons des données sur les années 90 et 99 mais nous ne pouvons pas connaître les valeurs de toutes les variables pour les années intermédiaires où des choix seront faits. L'incorporation de toute information plus proche du moment des choix pourra améliorer la qualité d'estimation mais celle-ci nécessitera des interventions de l'opérateur à l'exécution du logiciel.

Concernant le modèle de développement immobilier, il faut ajouter que si les changements de type d'occupation du sol au niveau des cellules ne sont pas assez nombreux, la qualité d'estimation diminue.

### **4.6.3 Contrôle et ajustement des estimations**

Pour les raisons précédemment évoquées, l'application des paramètres estimés n'aboutira pas aux résultats exactement égaux aux données de l'année de calage (ou année cible). C'est pourquoi, nous devons mesurer la qualité d'estimation et essayer d'ajuster les paramètres estimés afin de diminuer cet écart.

Le Pr Waddell et son équipe ont imaginé un processus d'ajustement automatique de ces paramètres. La méthode est décrite en détails dans le paragraphe 4.6 concernant l'étude de cas Eugène-Springfield.

## ***4.7 Etude de cas : Eugene-Springfield***

La première application d'UrbanSim a été réalisée sur la ville d'Eugène-Springfield dans l'Oregon (Etats-Unis) en 1998.

Les données étaient les suivantes :

- 77 535 ménages
- 186 280 individus

- 75 250 emplois
- un découpage du modèle de trafic en 271 zones
- un maillage en 15441 cellules carrées de dimension 150 \*150 mètres.

Les paramètres ont été estimés pour la période entre l'année initiale **1980** et l'année de calage **1994**. Cette période est assez longue pour voir des modifications significatives de l'occupation du sol. Les simulations visaient à reproduire la situation à l'horizon 2020. La base de données correspondant à l'année 1994 est mise en disponibilité pour les essais avec le logiciel.

Il nous a semblé important d'étudier dans le détail cette étude de cas. Elle nous a permis une première prise en main du logiciel UrbanSim.

Les données d'entrée sont présentées en annexe 1. Elles concernent l'occupation du sol, les emplois et les ménages.

Les valeurs des différents coefficients estimés pour les modules de choix de localisation des emplois, de choix de localisation des ménages, de développement immobilier et de fixation du prix de terrain sont présentées en annexe 2.

### Méthode d'ajustement

Les paramètres qui affectent la dynamique du modèle ainsi que ceux qui concernent les échanges sur le marché, les évolutions des prix et les actions des promoteurs immobiliers, sont explorés de manière heuristique afin de minimiser un certain indicateur de la distance entre le modèle et la réalité. D'autres paramètres surnommés "Une-fois" comme les taux de relocalisation ne font pas l'objet de calcul heuristique et leurs effets potentiels d'amélioration des résultats seront étudiés par quelques tests de sensibilité.

Les indicateurs de la distance sont définis au niveau des zones de modèle de trafic. Quatre indicateurs sont proposés afin de mesurer la distance. Avec  $X_z^{CM}$  étant la valeur simulée et  $X_z^{CR}$  la valeur réelle de la variable X concernant la zone z à l'année cible, on pose :

$$C_1 = \sum_{z=1,Z} \frac{(X_z^{CM} - X_z^{CR})^2}{(X_z^{CR})^2}$$

$$C_2 = \frac{1}{Z} \sum_{z=1,Z} (X_z^{CM} - X_z^{CR})$$

$C_3 = R^2$  de l'estimation de  $X^C$  par le modèle

$C_4 =$  un coefficient de corrélation entre les valeurs simulées et observées qui est entre -1 et 1.

En remplaçant  $X^C$  qui concerne l'année cible par la variable  $D = X^C - X^D$ , quatre autres mesures sont déduites :

$$C_5 = \sum_{z=1,Z} \frac{(D_z^M - D_z^R)^2}{(D_z^R)^2}$$

$$C_6 = \frac{1}{Z} \sum_{z=1,Z} (D_z^M - D_z^R)$$

$C_7 = R^2$  de l'estimation de D par le modèle

$C_8$  = un coefficient de corrélation entre les différences simulées et observées dont la valeur est comprise entre -1 et 1.

Le X mentionné ci-dessus peut être chacune des 12 variables du Tableau 9 qui sont classées en 5 catégories.

Variable	Catégorie	Identifiant r
Population totale	Population $j=1$	1
Nombre total de ménages		2
Nombre de ménages par tranche de revenu		3
Nombre de ménages par taille		4
Nombre total d'emplois	Emploi $j=2$	1
Nombre d'emplois par secteur		2
La surface par poste d'occupation du sol	Sols $j=3$	1
Nombre de logements par catégorie de l'occupation résidentielle du sol	Bâtis $j=4$	1
La surface de bâti par catégorie de l'occupation non résidentielle du sol		2
Prix de terrain par catégorie de l'occupation du sol	Valeurs $j=5$	1
Valeur de bâti par logement ou par unité de surface et par catégorie de l'occupation du sol		2
Valeur totale de bâti par catégorie de l'occupation du sol		3

Tableau 9 : Les variables de la base de données de contrôle

Les indicateurs de chaque catégorie seront agrégés.

Par exemple, pour la catégorie j :

$$C_{ij} = \sum_r C_{ijr}$$

et enfin nous aurons huit mesures de la distance, les  $C_{i=1,8}$

**La fonction objectif** à minimiser est considérée comme la somme de ces 8 indicateurs.

Afin d'optimiser la fonction objectif avec 77 paramètres, l'équipe du Pr Waddell a utilisé la méthode heuristique "Simulated Annealing" [1]. Pour l'exemple d'Eugène, 525 simulations en tout ont été lancées.

L'algorithme d'optimisation est schématisé à la figure 11.

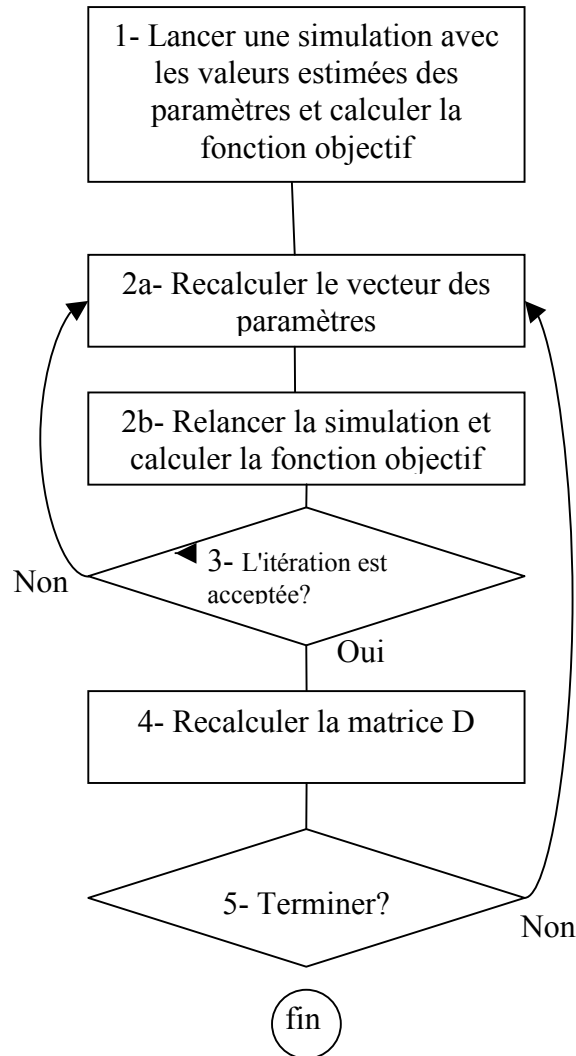


Figure 11 : l'algorithme d'optimisation, "Simulated Annealing »

La méthode est formalisée par l'algorithme suivant :

- 1- Faire une simulation avec les valeurs initiales,  $V_0$  étant le vecteur des paramètres issu de l'estimation par un logiciel d'économétrie externe.
- 2- Lancer une simulation avec  $V_{i+1} = V_i + Du$  où  $u$  est un vecteur de nombres aléatoires entre  $-1$  et  $1$ , et  $D$  est une matrice diagonale définissant la variation maximale autorisée des variables.
- 3- Contrôler la fonction objectif pour savoir si elle est améliorée ou pas. Dans le cas d'une dégradation du résultat, ce vecteur peut être quand même accepté avec une probabilité de  $p = \exp\left(-\frac{\delta f}{T\bar{d}}\right)$  où  $\delta f$  est la variation de fonction objectif,  $\bar{d}$  est le pas moyen et  $T$  est un paramètre au choix de l'utilisateur.
- 4- Pour une itération acceptée, mettre à jour la matrice  $D$  comme  $D_{i+1} = (1 - \alpha)D_i + \alpha\omega R$  où  $\alpha$  et  $\omega$  sont les paramètres à choisir par l'utilisateur et  $R$  la matrice diagonale comprenant les changements des variables pour l'itération acceptée.

5- Repartir à l'étape 2.

Afin de compenser l'inexactitude due à la construction de la fonction objectif par la somme d'indicateurs non normalisés, l'équipe américaine a analysé également les valeurs des huit mesures. Ils ont constaté que dans le cas d'Eugène, l'optimisation aboutit à une distance plutôt raisonnable pour les quatre premières mesures mais que par contre la distance des quatre dernières n'est pas assez satisfaisante.

## 4.8 Synthèse

- Il y a donc en tout 7 modules (plus 2 sous-modules) dans UrbanSim , chacun requérant des données spécifiques. Ils sont cependant interdépendants car les sorties de l'un sont les entrées des autres, le déroulement étant séquentiel et l'ordre reste au choix de l'utilisateur dans un fichier XML.

Exemple de séquences :

```
<accessibility-model />
<employment-location-choice-model />
<employment-relocation-choice-model />
<household-relocation-choice-model />
<household-transition-model />
<employment-transition-model />
<household-location-choice-model />
<land-price-model />
<developer-model />
```

- Par ailleurs, UrbanSim utilise une structure hiérarchique pour organiser les différentes simulations appelées « scénarios ».

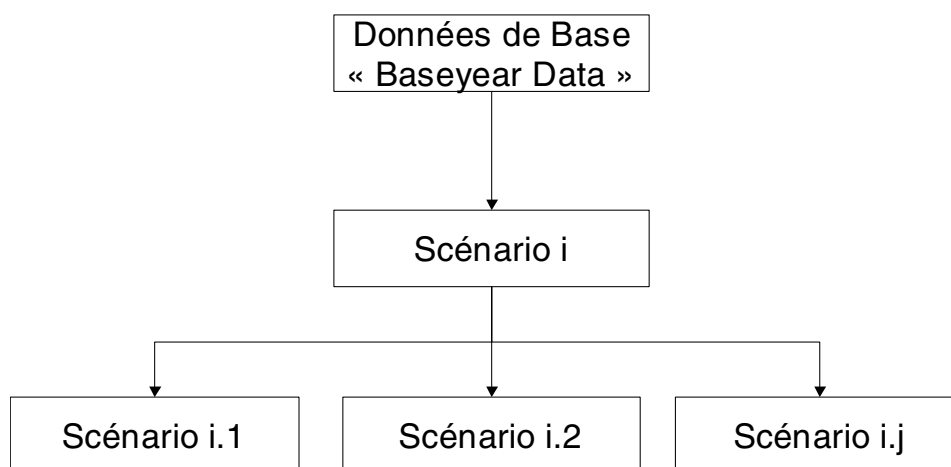


Figure 12 : les scénarii dans UrbanSim



Exemple :

Baseyear Data : données socio-économiques de la région Ile-de-France en 1990

Scenario i : ouverture de la tangentielle nord en 2012

Scenario i.1 : augmentation de 50% de la fréquence de la tangentielle nord en 2015

Scenario i.2 : augmentation de 100% de la fréquence de la tangentielle nord en 2015

Une simulation UrbanSim consiste à faire tourner le moteur UrbanSim sur un « scénario ». Les résultats de la simulation sont stockés dans une base de données unique (notée Output Database).

- En ce qui concerne l'interface utilisateur, UrbanSim utilise la plate-forme Eclipse développée par IBM. Mais dans la dernière version, la 2.8, cette plate-forme a été abandonnée pour revenir à un système de commandes en ligne, peut-être plus archaïque mais offrant plus de stabilité.
- Pour le cas d'Eugène-Springfield, le bilan en termes de corrélation entre la situation observée en 1994 et la situation reconstituée par simulation en 1994 a été fait à deux niveaux géographiques :
  - Au niveau des 15 441 cellules : les résultats du modèle ne sont pas censés être utilisés à un niveau aussi fin. Cependant, la corrélation a été analysée et elle s'est avérée bonne.
  - Au niveau des 271 zones du modèle de trafic : la corrélation est excellente pour les populations. C'est moins bon pour les emplois. Selon les auteurs, ce n'est pas étonnant car les emplois sont beaucoup plus concentrés et volatils que les populations.

Le modèle n'a pas pu prédire des événements exceptionnels, comme l'implantation d'un centre commercial régional (« mall ») ou la disparition d'une grande usine dans la région. Ces deux événements n'avaient jamais été envisagés avant 1980. Cela n'a rien aussi d'étonnant et on voit bien là les limites de la modélisation. De nombreux facteurs humains, de l'ordre du politique (les fameux « jeux d'acteurs »), sont impossibles à intégrer dans un modèle. De même, des facteurs difficilement mesurables peuvent influencer plus ou moins les comportements de mobilité. Pour caricaturer, comment intégrer le fait qu'un cyclone en Floride peut avoir une incidence sur le report modal en Ile-de-France (l'explication étant que le cours du baril de pétrole a flambé à cause de la présence d'une plate-forme off-shore dans le golfe du Mexique) ...

Il existe un moyen dans UrbanSim de prendre en compte des événements exceptionnels prévus dans le futur. Il suffit de définir dans les fichiers de type « events » les nombres d'emplois, de ménages, le changement de type urbain et l'année pour les cellules où se produisent ces événements.

Ainsi, l'exemple de l'application Eugène-Springfield montre que la dynamique de l'urbanisation (localisation des ménages et des activités) n'est pas simple à reconstituer, sur un territoire 70 fois plus petit que la Région d'Ile-de-France. On peut donc nourrir de petites inquiétudes et s'attendre à être confronté à de réelles difficultés de calage au cours de l'année 2 de notre recherche.

Malgré cela, le choix d'UrbanSim nous paraît le mieux approprié pour notre projet. Si on devait caractériser UrbanSim en quelques mots, on dirait qu'il est :

- modulaire (modules de choix discret)
- calable sur des données désagrégées
- flexible au niveau du nombre de zones (pour notre projet, on a un découpage de la Région Ile-de-France en 50 000 cellules ! Voir partie 7)
- dynamique
- relativement complet d'un point de vue théorique
- ouvert quant à la prise en compte des variables en entrée
- pratique car il y a une séparation claire entre les données et les modules
- évolutif de par ses développements permanents à l'Université de Washington.

Il faut noter enfin qu'UrbanSim est, au moment où ce rapport est finalisé, en pleine restructuration. La dernière version, la 2.8, propose non plus 9 modules mais 15 !

Par ordre alphabétique, il y a :

- « Accessibility Model »
- « Developer Model »
- « Developer Model With Vacancy Adjustment »
- « Employment Home-Based Location Choice Model »
- « Employment Non-Home-Based Location Choice Model »
- « Employment Relocation Choice Model »
- « Employment Transition Model »
- « Household Location Choice Model »
- « Household Relocation Choice Model »
- « Household Transition Model »
- « Land Price Adjusting Location Choice Models »
- « Land Price Model »
- « Model Variable Logger Model »
- « Residential Land Share Model »
- « Scaling Procedure for Jobs Model »

C'est en particulier le module « Employment Location Choice Model » qui a été entièrement refondé pour prendre en compte le cas des emplois à domicile de plus en plus nombreux aux Etats-Unis.

Pour en savoir plus, il faut se référer à la documentation en ligne du site [urbansim.org](http://urbansim.org).

Les nouveautés seront décrites dans le rapport de la deuxième année.

## 5 Modèle de trafic METROPOLIS

### 5.1 Introduction

Le logiciel METROPOLIS (utilisé par le laboratoire THEMA de l'Université de Cergy-Pontoise, l'EPFZ en Suisse, l'Université d'Alberta et des Universités américaines) a été choisi dans notre projet pour être intégré au logiciel UrbanSim. A ce jour, il constitue essentiellement un modèle d'affectation dynamique routière. Il ne possède pas encore de modèle de demande de déplacements ni de modèle d'affectation en transports en commun.

METROPOLIS utilise pour l'aire d'étude qu'est la région Ile-de-France les matrices de déplacements de l'IAURIF en 572 zones (auxquelles il faudrait ajouter 34 zones externes, soit en tout 606 zones). Dans le cadre de cette recherche, nous avons donc décidé de reprogrammer le modèle de demande de l'IAURIF dans le logiciel METROPOLIS.

D'abord, faisons quelques rappels généraux sur les modèles de trafic et l'aspect innovant de METROPOLIS par rapport aux modèles statiques. Pour plus de détails sur METROPOLIS, on pourra se reporter à l'annexe 4.

Les planificateurs en matière de transports disposent depuis de nombreuses années de modèles permettant de décrire les comportements individuels ou agrégés en matière de transports urbains. L'objectif de ces modèles est de reconstituer et de prédire les déplacements puis les trafics dans un réseau de transports urbains. Les éléments constitutifs de ces modèles sont la demande de transport (matrices origine-destination) et l'offre de transports (réseau routier et réseau de transports collectifs). Le modèle détermine alors un équilibre dans lequel l'offre égalise la demande sur chacun des réseaux.

Depuis les premiers modèles de trafic statiques agrégés jusqu'aux derniers développements, de nombreuses améliorations ont été introduites en termes de fonctionnalités. La principale innovation est l'utilisation des modèles dynamiques, dans la lignée des travaux de Vickrey [45]. Ainsi, le modèle METROPOLIS introduit des éléments tels que :

- les choix de mode, d'heure de départ et d'itinéraire des usagers,
- le calcul de situation d'équilibre entre l'offre et la demande,
- la simulation totalement dynamique du trafic (seconde après seconde), en période de pointe et en période creuse,
- des données compatibles avec celles des modèles statiques en terme d'offre et de demande,
- des lois de congestion flexibles arbitraires définies par l'utilisateur.

Ainsi, la description de l'équilibre dans le réseau de transports est beaucoup plus réaliste. En contrepartie, elle requiert des capacités de calcul relativement importantes.

Décrivons maintenant le modèle de demande de l'IAURIF et la manière de le reprogrammer dans METROPOLIS.

## 5.2 *Elaboration des matrices origine - destination*

### 5.2.1 Introduction

Il s'agit de construire des matrices origine-destination de déplacements. Les éléments d'une matrice représentent le nombre de déplacements, d'une zone  $i$  vers une zone  $j$ , pour un motif donné de déplacement. La région géographique considérée est découpée en  $n$  zones (il est possible aussi de considérer que l'une des  $n$  zones représente l'ensemble des zones de déplacements qui ne sont pas localisées dans la zone géographique étudiée).

Nous commencerons par introduire les motifs de déplacement, ce qui conduira à une classification des *motifs de déplacements*. Ensuite, nous modélisons les émissions (nombre de déplacements générés à partir d'une zone) et les attractions (nombre de déplacements attirés par une zone). De ces effectifs pour l'émission et l'attraction, calculés pour un motif donné de déplacement, nous en déduisons les éléments de la *matrice Origine - Destination*, ou O-D (dont l'élément de matrice  $i,j$  représente le nombre de déplacements pour un motif de déplacement donné, de la zone  $i$  vers la zone  $j$ ).

Ensuite, ces déplacements sont répartis selon un mode de transport : le mode public (transport en commun) et le mode privé (voiture particulière) à l'aide d'un modèle binaire de choix discret (le modèle Logit binaire dans le cas considéré ici). Enfin, une procédure d'affectation permet d'identifier le chemin (ou les chemins) parcouru par l'individu (successions de tronçons de route ou trajets de bus, tramway, métro, RER et train).

Le résultat de cette procédure est ce que l'on appelle *le modèle classique à 4 étapes*. Nous commençons donc par introduire le modèle classique à 4 étapes de l'IAURIF sur la Région Ile-de-France : modèle de génération (calculs des émissions et attractions), modèle de distribution (association des émissions et des attractions pour produire des déplacements d'une zone origine vers une zone destination), modèle de choix modal (entre transports en commun et voiture particulière) et modèle d'affectation (c'est-à-dire le modèle de choix d'itinéraire). Nous présentons d'abord ces quatre étapes dans un cadre unifié (voir Quinet, 1998). Ensuite nous considérons l'extension de cette approche classique au cas dynamique, c'est-à-dire aux situations dans lesquelles le niveau de congestion est fonction de l'heure de la journée. Dans ce cas, les usagers font face à un choix supplémentaire : le choix de l'heure de départ. Nous expliquons comment intégrer cette dimension supplémentaire et construire ce que nous nommerons le « modèle à 5 étapes ». La cohérence de cette approche sera enfin discutée.

De manière schématique, nous pouvons décrire le modèle à 4 (et 5) étapes de la manière suivante :

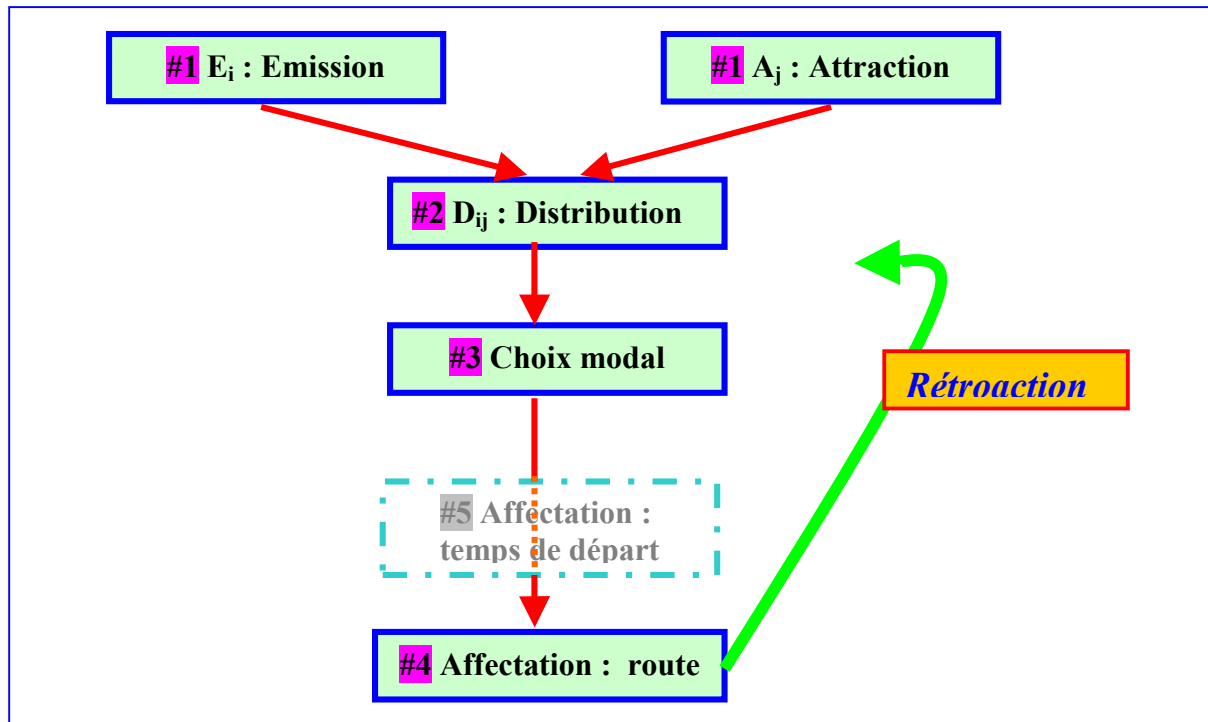


Figure 13 : représentation synthétique du modèle à 5 étapes

### 5.2.2 Segmentation des motifs

L'espace est segmenté en  $n$  zones. L'objet des modèles de demande de transport est de décrire la mobilité des usagers. S'il s'agit d'une zone de départ, nous l'indiquons par  $i$ , et s'il s'agit d'une zone d'arrivée, nous l'indiquons par  $j$ .

Un déplacement est défini par un *motif origine* et un *motif destination*. Un motif de déplacement est défini comme un couple orienté : motif origine  $\rightarrow$  motif destination. Nous segmentons l'ensemble des motifs en  $M$  motifs de déplacement. Le cas que nous traiterons comporte  $M = 9$  motifs origine ou destination, qui sont définis comme suit :

- $m=1$  Domicile : noté « DO »
- $m=2$  Lieu de travail : noté « TR »
- $m=3$  Affaires professionnelles : noté « APR »
- $m=4$  Affaires personnelles : noté « APE »
- $m=5$  Achats : noté « ACH »
- $m=6$  Loisirs : noté « LOI »
- $m=7$  Ecole/ Enseignement supérieur : noté « ECO »
- $m=8$  Santé : noté « SA »
- $m=9$  Divers : noté « DIV »

On exclut les déplacements de  $m$  vers  $m$ , de sorte que nous avons  $(M) \times (M-1)$  motifs de déplacements différents. Par exemple, les déplacements du domicile « DP » au lieu de travail « TR » correspondent au motif de déplacement  $\{DO \rightarrow TR\}$ . Ceci signifie dans l'exemple considéré 72 motifs de déplacement (ou en ne comptant pas le sens de déplacement, cela correspond à 36 couples non ordonnés de motifs de déplacement).

Considérons les motifs de déplacement à partir du domicile. On peut les énumérer :

Domicile- Travail :	DO→TR
Domicile-Affaires professionnelles	DO→APR
Domicile-Affaires personnelles	DO→APE
Domicile-Achats	DO→ACH
Domicile-Loisirs	DO→LOI
Domicile-Ecole, enseignement supérieur	DO→EC
Domicile-Santé	DO→SA
Domicile-Divers	DO→DIV

En définitive, nous avons la matrice motif origine – motif destination tous modes motorisés suivante :

	Domicile	Travail	Affaires Prof.	Affaires Perso.	Achats	Loisirs	Ecole	Santé	Divers
Domicile		DO→ TR	DO→ APR	DO→ APE	DO→ ACH	DO→ LOI	DO→ EC	DO→ SA	DO→ DIV
Travail									
Affaires Prof.									
Affaires Perso.									
Achats									
Loisirs									
Ecole									
Santé									
Divers									

Tableau 10 : matrice motif-origine – motif destination

A chaque motif, le modèle à 4 (ou 5) étapes associe une matrice de déplacements de zone à zone (désagrégation spatiale) aux émissions et attractions. Il y a par conséquent autant de matrices Origine-Destination (O-D) de zone à zone que d'éléments de matrice non nuls dans la matrice motif origine – motif destination, représentée dans le tableau 10.

### 5.2.3 Génération

#### Emissions

Les émissions correspondent au nombre de déplacements émis à partir d'une zone pour un motif donné de déplacement. Bien évidemment, chaque activité est potentiellement génératrice de déplacements. Par exemple, un hôpital génère les déplacements des employés, des livreurs, des malades, des médecins et des infirmières, des visiteurs, etc. Nous nous intéressons ici aux déplacements motorisés (voiture particulière VP et transport en commun

TC) produits par une zone donnée  $i$  contribuant aux effectifs d'un motif donné. Pour le déplacement de motif Domicile vers le Travail, noté DO→TR, (rappelons que DO est 1 et TR est 2), les émissions  $E_i^{1,2}$  à partir de la zone  $i$  ( $i=1 \dots n$ ) sont données par :

$$E_i^{1,2} = \Theta^{1,2} X_i^{1,2},$$

où  $\Theta^{1,2}$  représente le paramètre d'émission (qui peut être évalué par l'Enquête Globale Transports ou EGT) pour le motif Domicile → Travail (noté DO→TR) et où  $X_i^{1,2}$  représente la variable d'émission correspondant à ce motif, à partir de la zone  $i$ . Par exemple,  $X_i^{1,2}$  représente la population active dans la zone  $i$  et  $\Theta^{1,2}$  représente le nombre moyen de déplacements vers le lieu de travail depuis le domicile, par jour ouvrable, générés par un actif (par exemple nous aurons  $X_i^{1,2} = 10.000$  et  $\Theta^{1,2} = 1,2$ ). Ce paramètre  $\Theta^{1,2}$  est considéré comme uniforme sur l'ensemble de l'Ile-de-France (il n'est pas indicé par  $i$ , relatif à la zone d'émissions). Mais on pourrait affiner en calculant des  $\Theta^{1,2}$  par département de résidence (ce qui a été fait dans le modèle global de l'IAURIF).

De manière plus générale, on aura pour l'ensemble des déplacements à partir du domicile :

$$E_i^{1,m} = \Theta^{1,m} X_i^{1,m}, m = 2 \dots M; i = 1 \dots n.$$

Ceci complète la modélisation de la première ligne de la matrice. Il reste maintenant à écrire des expressions pour les autres lignes de la matrice.

Une manière simple de procéder revient à postuler pour tous les déplacements de  $m$  vers  $r$  des relations de même type que celle postulée pour les déplacements domicile → travail. Pour l'ensemble des déplacements, nous considérons la formulation unique suivante :

$$E_i^{p,m} = \Theta^{p,m} X_i^{p,m}, p, m = 1 \dots M, p \neq m, i = 1 \dots n.$$

Ceci nécessite la connaissance de  $M \times (M-1)$  paramètres  $\Theta^{p,m}$  et de  $M \times (M-1) \times n$  variables,  $X_i^{p,m}$ . Ceci représente en tout  $M \times (M-1) \times (n+1)$  données : pour  $M = 9$ , cela signifie  $72 \times 573$  soit environ 41000 unités d'informations pour les 572 zones internes du modèle IAURIF, ce qui nous semble excessif.

Ces chiffres sont bien sûr théoriques (il s'agit de bornes supérieures), puisqu'en pratique tous les motifs de déplacement ne sont pas pris en compte. Nous présentons dans l'annexe 3 la description des motifs de déplacement proposée par l'IAURIF et la description des motifs origine- motifs destination retenus en pratique.

Nous examinons aussi ci-dessous deux procédures de simplification qui peuvent être utilisées en vue de réduire la quantité d'informations nécessaire :

### ***Simplifications pour évaluer les émissions***

a) On peut utiliser une approche moins gourmande en données. Par exemple, on peut se baser sur les déplacements de motifs domicile-travail (de  $m = 1$  vers  $m = 2$ ) pour décrire les autres déplacements vers le lieu de travail, mais ayant des motifs origines différents. Plus généralement, nous supposons que :

$$E_i^{p,m} = \Gamma^{p,m} E_i^{1,m}, p = 2 \dots M, m = 1 \dots M, p \neq m, i = 1 \dots n.$$

Dans ce cas, lorsque les déplacements à partir du domicile sont décrits, il est nécessaire de déterminer  $(M-1) \times (M-1)$  inconnues pour définir les paramètres  $\Gamma^{m,r}$  (ce qui fait un maximum de  $8 \times 8 = 64$  paramètres à déterminer dans notre exemple). On peut simplifier encore davantage cette approximation de la manière suivante en supposant :

$$E_i^{p,m} = \Gamma^m E_i^{1,m}, p = 2 \dots M, m = 1 \dots M, p \neq m, i = 1 \dots n$$

Dans ce cas, les  $(M-1)$  paramètres supplémentaires inconnus sont les  $\Gamma^m$ .

b) Une autre solution est possible. On peut supposer que seules les variables au lieu d'origine comptent, ainsi que les caractéristiques du déplacement. C'est-à-dire, chaque ménage, par exemple, génère un déplacement pour le lieu de travail chaque jour et deux déplacements pour motifs affaires personnelles (à partir du domicile). Les différences entre les motifs à la destination  $r$  sont dans ce cas capturées uniquement par les paramètres  $\Theta^{m,r}$ . Dès lors  $X_i^{1,m} = X_i^m; m = 1 \dots M, (r \neq 1), i = 1 \dots n$ . Plus généralement, pour cette approximation, nous avons :

$$E_i^{p,m} = \Theta^{p,m} X_i^p, p, m (p \neq m) = 1 \dots M, i = 1 \dots n.$$

### Attractions

Pour les attractions, on procède d'une manière similaire. Pour le déplacement domicile vers le lieu de travail, noté DO  $\rightarrow$  TR, les attractions  $A_j^{1,2}$  de la zone de destination  $j$  ( $j = 1 \dots n$ ) sont données par :

$$A_j^{1,2} = \Phi^{1,2} Y_j^{1,2}$$

où  $\Phi^{1,2}$  représente un paramètre à extraire de l'EGT et où  $Y_j^{1,2}$  représente une variable explicative de l'attraction de la zone  $j$ , en l'occurrence les emplois. Le modèle le plus général, et le plus gourmand en variables et en paramètres est :

$$A_j^{p,m} = \Phi^{p,m} Y_j^{p,m}, p, m (p \neq m) = 1 \dots M, i = 1 \dots n.$$

Ici, comme dans le cas des émissions, l'approche complète nécessite la connaissance de  $(M-1) \times (M-1)$  paramètres  $\Phi^{p,m}$  et de  $(M-1) \times (M-1) \times n$  variables,  $Y_j^{p,m}$ , puisqu'il y a  $n$  zones, soit  $(M-1) \times (M-1) \times (n+1)$  inconnues.



## Les lois de conservation

Les attractions et les émissions pour un motif de déplacement sont reliées par les lois de conservation suivantes :

$$\sum_{i=1}^n E_i^{p,m} = \sum_{j=1}^n A_j^{p,m} = D^{p,m}$$

où  $D^{p,m}$  représente le nombre de déplacements correspondant au motif  $p \rightarrow m$ . Dans notre exemple, cela signifie 72 contraintes. Notons que le nombre total de déplacements, noté  $D$  est donné par :  $\sum_{p,m=1}^M D^{p,m} = D$ .

### 5.2.4 Distribution

Le module distribution consiste à associer à des émissions et des attractions un volume de déplacements d'une origine (zone d'émission) à une destination (zone d'attraction). Nous considérons d'abord l'approche classique retenue ici. Nous décrivons ensuite brièvement une autre manière d'aborder le problème de la distribution.

#### Approche classique

Nous désirons dans ce cas calculer le nombre de déplacements  $M_{ij}^r$  de la zone  $i$  vers la zone  $j$  pour le motif noté par un indice unique, noté « r » par simplicité d'écriture. Notons ce nombre par :

$$M_{ij}^r = k_i l_j E_i^r A_j^r f(c_{ij})$$

où  $E_i^r$  représente les émissions de la zone  $i$ ,  $A_j^r$  les attractions de la zone  $j$  et où  $f(c_{ij})$  représente une fonction des coûts entre  $i$  et  $j$ . Nous avons  $f(.) > 0$  et  $f'(.) \leq 0$ .

On considère la spécification suivante pour la fonction de coût :

$$f(c_{ij}) = \frac{\exp(-\theta c_{ij})}{(c_{ij})^\nu}$$

où  $\theta$  et  $\nu$  sont deux paramètres non négatifs à estimer. Si  $\theta = 0$  et  $\nu = 2$ , nous avons le modèle gravitaire (de type entropique, introduit par A. Wilson, à l'Université de Leeds). Si  $\nu = 0$ , nous avons un modèle de type choix discrets (cf. D. McFadden).

Par exemple, si l'on considère les estimations de l'IAURIF, nous obtenons :

$$f(c_{ij}) = c_{ij}^{1.18} \exp(-0.08c_{ij}).$$

Cette fonction, représentée ci-dessous, montre que la distance a un impact globalement décroissant (de manière exponentielle) en fonction du facteur  $c_{ij}$  (quelque fois appelé facteur d'impédance).

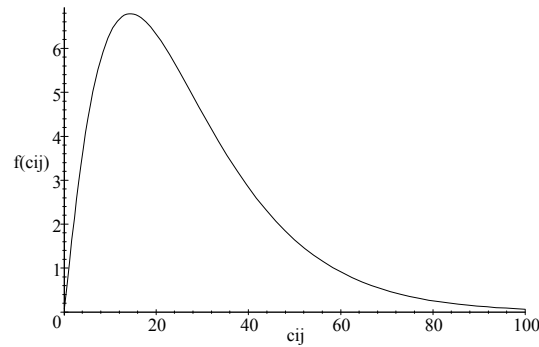


Figure 14 : Représentation de la fonction  $f(c_{ij})$ , en fonction de l'impédance  $c_{ij}$  mesurée en minutes

Notons que cette fonction est croissante pour des temps de trajet de 0 à 15 minutes (environ).

La matrice 0-D (Origine-Destination) a la forme suivante :

$$\begin{pmatrix} M_{1,1} & \dots & M_{1,j} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,j} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & \dots & M_{n,j} & \dots & M_{n,n} \end{pmatrix}$$

*Matrice origine-destination pour un motif de déplacement donné*

Les termes  $k_i$  et  $l_j$  sont des facteurs de normalisation inconnus qui garantissent les identités suivantes:

$$\begin{cases} \sum_{i=1}^n M_{ij}^r = A_j^r \\ \sum_{j=1}^n M_{ij}^r = E_i^r \end{cases}$$

Nous avons de plus la loi de conservation suivante :  $\sum_{j=1}^n A_j^r = \sum_{i=1}^n E_i^r = D^r$

La formulation proposée dans la littérature impose une forte hypothèse de symétrie, puisque cela revient à supposer que les déplacements entre  $i$  et  $j$  ne dépendent que des coûts entre  $i$  et  $j$ . Il s'agit d'une hypothèse de symétrie forte, mais que nous ne remettons pas en question ici car elle est très largement acceptée dans la profession.

Nous considérons les trois options suivantes pour la spécification des coûts:

$$c_{ij} = \begin{cases} d_{ij} \\ t_{ij}^0 \\ t_{ij}^v \end{cases}$$

Nous avons ainsi respectivement  $c_{ij}$  égal à la distance euclidienne entre  $i$  et  $j$ , au temps de parcours sans congestion, sur le réseau routier et au temps de parcours, mesuré (par exemple expérimentalement ou par simulation) avec congestion entre la zone  $i$  et la zone  $j$ . Notons que la version dans laquelle les usagers anticipent rationnellement le niveau de congestion (avant de se localiser en un endroit et d'effectuer des déplacements) n'est pas considéré ici, étant donné, que stricto sensu, cette approche nécessite une anticipation de la part des individus de l'ensemble des états futurs possible, ce qui nous semble difficile à calculer. Les coûts de la zone  $i$  vers la zone  $j$  ne sont pas nuls, et permettent de décrire les déplacements intra zones. En général, ces déplacements intra zones ne sont affectés par les logiciels d'affectation statiques (DAVISUM) ou dynamique (METROPOLIS). Par exemple, en pratique, on peut prendre pour  $c_{ii}$ , 50% du coût minimum entre  $i$  et les autres zones.

La première relation de conservation devient :

$$\sum_{i=1}^n M_{ij}^r = l_j A_j^r \sum_{i=1}^n k_i E_i^r f(c_{ij}) = A_j^r$$

Soit :

$$l_j = \left[ \sum_{v=1}^n k_v E_v^r f(c_{vj}) \right]^{-1}$$

En conséquence, par remplacement, nous avons :

$$M_{ij}^r = \frac{k_i}{\sum_{v=1}^n k_v E_v^r f(c_{vj})} E_i^r A_j^r f(c_{ij})$$

La seconde relation de conservation s'écrit maintenant, après substitution :

$$\sum_{j=1}^n M_{ij}^r = \sum_{j=1}^n \frac{k_i}{\sum_{v=1}^n k_v E_v^r f(c_{vj})} E_i^r A_j^r f(c_{ij}) = E_i^r$$

soit :

$$k_i \sum_{j=1}^n \frac{A_j^r f(c_{ij})}{\sum_{v=1}^n k_v E_v^r f(c_{vj})} = 1$$

On remarque que le membre de gauche de cette équation est une fonction croissante de  $k_i$ , tandis que le membre de droite de cette équation est constant, de sorte que la relation de récurrence suivante est envisagée pour trouver les facteurs de normalisation inconnus  $k_i$ . La vitesse de convergence de la récurrence (où  $w$  représente le pas) dépend d'un facteur d'ajustement  $\lambda$ , avec  $\lambda < 1$ , et représente le paramètre de convergence, qui doit être pris suffisamment petit, mais pas trop (pour ne pas ralentir le processus).

$$k_i(w+1) = k_i(w) + \lambda \left\{ \frac{1}{\sum_{j=1}^n \frac{A_j^r f(c_{ij})}{\sum_{v=1}^n k_v(w) E_v^r f(c_{vj})}} - k_i(w) \right\}$$

On peut aussi résoudre ce problème de point fixe pour trouver les facteurs de normalisation  $k_1 \dots k_n$  par la méthode de Newton.

D'un point de vue numérique, on a procédé d'une manière différente qui s'avère être beaucoup plus rapide que la méthode exposée ci-dessus :

Il s'agit à nouveau d'une procédure itérative. Notons par  $w$  le numéro de l'itération.

- ETAPE 1 : Initialiser  $w$  à 1 et fournir les valeurs initiales de  $k$ . En pratique on prend  $k_i(1) = 1, i = 1 \dots n$ .
- ETAPE 2 : Calculer les valeurs de  $l_i, i = 1 \dots n$ , à l'itération  $w+1$ , en utilisant les relations suivantes :

$$l_j(w+1) = \left[ \sum_{v=1}^n k_v(w) E_v^r f(c_{vj}) \right]^{-1}$$

- ETAPE 3 : Calculer les valeurs de  $k_i, i = 1 \dots n$ , en utilisant les relations suivantes :

$$k_i(w+1) = \left[ \sum_{v=1}^n l_v(w+1) A_v^r f(c_{iv}) \right]^{-1}$$

- ETAPE 4 : Calculez  $\text{Max}_{i=1..n} |k_i(w+1) - k_i(w)| = B_k(w)$ ,  $\text{Max}_{i=1..n} |l_i(w+1) - l_i(w)| = B_l(w)$  et  $\text{Max} |B_l(w) - B_k(w)| = B(w)$ . Si  $B(w) < \text{Seuil}$  : Stop. Sinon :  $w \rightarrow w+1$  et aller à l'ETAPE 2.

Des expériences de calculs effectués pour la matrice de l'IAURIF (572x572) montrent que cette procédure converge avec une précision de  $10^{-6}$  après 5 à 6 itérations.

### Approche concurrente

Du point de vue de l'économiste, le problème de la distribution est un problème d'appariement entre l'offre et la demande. En effet, les émissions peuvent être interprétées comme une offre et les attractions peuvent être interprétées comme une demande. Dans ce cas, les émissions et les attractions doivent être traitées comme des fonctions des prix et non pas comme des constantes (ce qui est fait dans l'approche considérée ci-dessus). Nous pouvons alors, une fois ces fonctions d'offre et de demande spécifiées, recourir à la technique de résolution des problèmes d'équilibre général calculables.

Il nous est possible de fournir très simplement l'idée sous-jacente à cette approche basée sur l'ajustement des prix. Pour simplifier l'exposé, nous conservons l'hypothèse (restrictive) selon laquelle les émissions et les attractions sont données.

Considérons une émission pour le motif domicile - travail ( $r = \text{DOM} \rightarrow \text{TR}$ ). Il s'agit par conséquent de modéliser le choix du lieu de travail d'un individu qui quitte son domicile localisé par le zone  $i$ . On peut utiliser le modèle de choix discret suivant :

$$P_{j|i}^r = \frac{\exp\left(\frac{\omega_j^r - c_{ij}}{\mu^r}\right)}{\sum_{k=1}^n \exp\left(\frac{\omega_k^r - c_{ik}}{\mu^r}\right)}$$

où  $\omega_1^r \dots \omega_n^r$  sont  $n$  paramètres à estimer (on a posé  $c_{ii} = -\infty$ , puisque les déplacements étudiés sont nécessairement entre deux zones différentes). On peut interpréter les variables  $\omega_1^r \dots \omega_n^r$  comme des prix à la destination (des salaires).

Dans ce cas, nous avons :

$$M_{ij}^r = E_i^r P_{j|i}^r$$

et

$$\sum_{j=1}^n M_{ij}^r = \sum_{j=1}^n E_i^r P_{j|i}^r = E_i^r \sum_{j=1}^n P_{j|i}^r = E_i^r,$$

de sorte que la seconde loi de conservation est automatiquement satisfaite. En ce qui concerne la première loi de conservation, elle s'écrit :

$$\sum_{i=1}^n M_{ij}^r = \sum_{i=1}^n E_i^r P_{j|i}^r = \sum_{i=1}^n \frac{E_i^r \exp\left(\frac{\omega_j^r - c_{ij}}{\mu^r}\right)}{\sum_{k=1}^n \exp\left(\frac{\omega_k^r - c_{ik}}{\mu^r}\right)} = A_i^r$$

Etant donnée la condition selon laquelle la somme des émissions est égale à la somme des attractions, ce système comporte  $(n-1)$  équations et  $(n-1)$  inconnues :  $\omega_2 \dots \omega_n$ . Il est ainsi possible d'utiliser des processus walrasiens d'ajustement des prix pour déterminer le vecteur inconnu des prix. Une procédure (« cob web ») a été proposée par de Palma et Marchal (2003).

### 5.2.5 Choix modal

Une partie des individus est captive, et une autre partie des usagers a le choix entre un mode de transport privé et un mode de transport public, que nous notons par simplicité : voiture particulière (VP) et transports en commun (TC). Le modèle peut facilement être étendu à un nombre quelconque de modes. Dans une version simplifiée du modèle, on suppose que la fraction des individus captifs du bas est fixe et égale à  $\xi_{ij}$  et allant de la zone  $i$  vers la zone  $j$ . (les individus captifs des TC ne jouent aucun rôle, de sorte qu'ils ne sont pas mentionnés dans notre analyse). Considérons une matrice O/D de déplacement motorisés fixe et donnée (résultat de l'étape de distribution). Prenons le cas d'un individu qui va de l'origine  $i$  vers la destination  $j$ . La probabilité qu'un individu choisisse la voiture est  $1$  s'il est captif de la voiture,  $0$  si il est captif des transports en commun. On suppose que cette probabilité est donnée par un modèle logit binaire entre les TC et la VP.

Le modèle de choix modal pour un individu non captif de la voiture (ou des TC) est donné par un modèle logit binaire. La probabilité que cet individu allant de  $i$  vers  $j$  emprunte sa voiture ( $v$ ) pour le motif  $r$  est donné par:

$$P_{ij}^v = \frac{1}{1 + \exp\left(\frac{1}{\mu^r} \frac{\alpha^v t_{ij}^v - \alpha^b t_{ij}^b - \pi^r}{\alpha^v t_{ij}^v + \alpha^b t_{ij}^b}\right)}$$

où nous avons utilisé les notations suivantes :

$t_{ij}^v$  ( $t_{ij}^b$ ) représente le temps de parcours privé (public), pour l'élément  $i, j$  de la matrice O-D.

$\alpha^v$  ( $\alpha^b$ ) représente le coût du temps de parcours pour les transports privés (publics),

$\mu^r$  représente un facteur d'échelle du modèle logit binaire, associé au motif  $r$ .

$\pi^r$  représente la pénibilité associée au transport en motif de déplacement  $r$ .

Notons que dans ce modèle il y a de l'hétéroscédasticité (la variance des termes d'erreurs n'est pas la même pour les deux modes). Ceci n'a pas d'importance pour le cas binaire, puisque seule compte la différence des termes d'erreur, mais ceci doit être pris en compte dans le cas multinomial (par exemple, pour le choix entre la voiture, le bus ou le métro).

Pour un individu choisi de manière aléatoire dans la population (pouvant par conséquent être ou non captif des transports en commun), nous avons :

$$P_{ij}^r(v) = \xi_{ij}^r + \frac{1}{1 + \exp\left(\frac{1}{\mu^r} \frac{\alpha^v tt_{ij}^v - \alpha^b tt_{ij}^b - \pi^r}{\alpha^v tt_{ij}^v + \alpha^b tt_{ij}^b}\right)}$$

Dans la pratique de notre travail de modélisation, nous supposons qu'une fraction constante  $\xi$  des individus est captive (cette fraction est indépendante de la zone d'origine et de destination). Une autre fraction de la matrice aura le choix entre les transports privés et les transports publics en fonction des conditions de circulation entre ces modes. Cette fraction  $(1 - \xi)$  est décrite par un modèle logit binaire. Par convenance, la valeur du temps pour les transports individuels peut être prise égale à l'unité tandis que la valeur des temps de parcours

pour les transports en commun, exogène, peut aussi être multipliée par le facteur  $\alpha^b / \alpha^v$ . Grâce à cette normalisation, nous pouvons réécrire le modèle proposé par l'IAURIF de la manière suivante :

$$P_{ij}^r(v) = \xi^r + \frac{1}{1 + \exp\left(\frac{1}{\mu^r} \frac{tt_{ij}^v - tt_{ij}^b - \hat{\pi}^r}{tt_{ij}^v + tt_{ij}^b}\right)}$$

Dans METROPOLIS, on utilise un modèle de choix modal logit binaire pour les usagers non captifs.

Notons que les temps de parcours peuvent dépendre des motifs (puisque les heures de départ sont différentes). On peut aussi être plus ou moins sensible à la congestion en fonction des motifs de déplacement.

### 5.2.6 Choix de l'heure de départ

Dans ce cas, l'arbre de décision de l'individu est un modèle logit emboîté à trois étapes. Ce modèle peut être formalisé en utilisant les valeurs incluses, qui ont une représentation explicite lorsque les termes d'erreurs sont indépendamment et identiquement distribués et distribués selon la loi double exponentielle.<sup>3</sup>

Dans un modèle dynamique, les usagers ont aussi le choix de l'heure de départ. Dans ce cas, le modèle à quatre étapes est modifié comme indiqué précédemment. Les temps de parcours intervenant dans le module de distribution seront remplacés par les valeurs incluses et il en est de même pour les temps de parcours intervenant dans le modèle de choix modal.

<sup>3</sup> Dans le cas de METROPOLIS, nous considérons un modèle logit ordonné. Les probabilités de choix sont les mêmes que celles présentées ci-dessus, mais le mode de simulation de ce modèle est différent puisqu'il suffit d'une seule variable aléatoire pour modéliser le modèle logit (ou probit) ordonné.

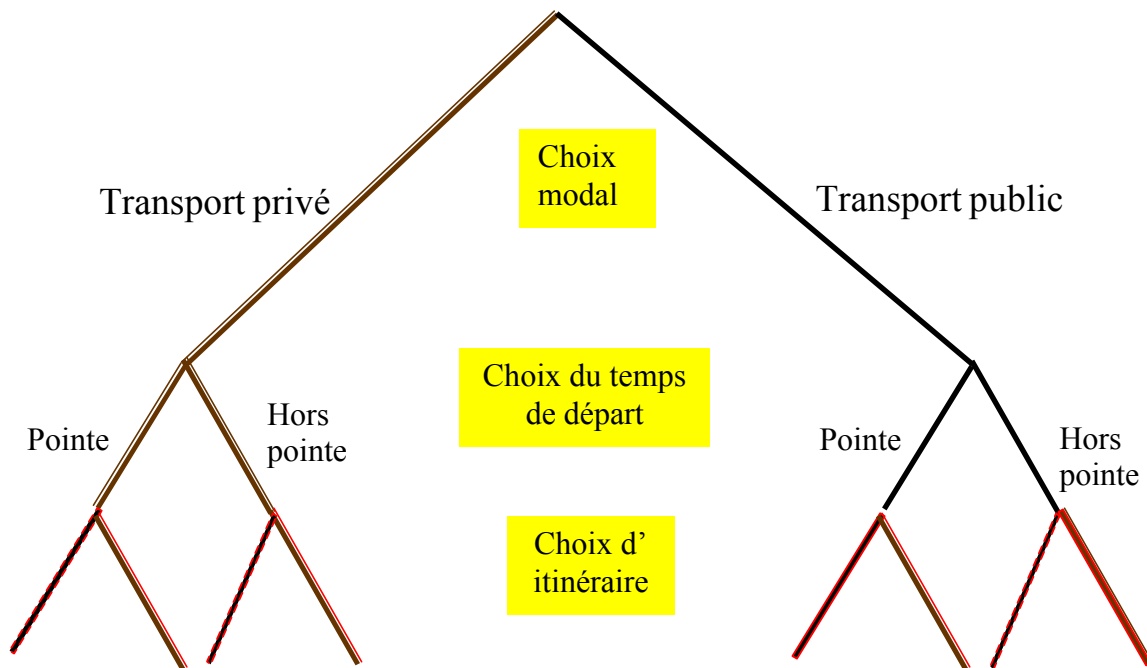


Figure 15 : arbre de décision dans les modèles dynamiques

Considérons le déplacement de la zone  $i$  vers la zone  $j$ . Le coût généralisé de déplacement  $C_{ij}^r(t)$  pour le motif  $r$ , la zone d'origine  $i$  et la zone de destination  $j$  et un temps de départ  $t$  est défini par l'équation suivante (le mode VP ou TC est donné, mais nous ne le spécifions pas, pour ne pas alourdir les notations) :

$$C_{ij}^r(t) = \alpha tt_{ij}^r(t) + \beta [\hat{\tau}^r - t - tt_{ij}^r(t)]^+ + \gamma [t + tt_{ij}^r(t) - \hat{\tau}^r]^+ + \tau_{ij}^r(t)$$

où  $tt_{ij}^r(t)$  représente le chemin le plus court (ou optimal s'il y a une tarification  $\tau_{ij}^i(t)$  fonction de l'heure de la journée) pour le couple origine  $i$  - destination  $j$  et pour le mode  $r$  et où  $\hat{\tau}^r$  représente le temps d'arrivée désirée à la destination pour le motif  $r$ .

Le nombre d'individus quittant  $i$  pour se rendre en  $j$  pour le motif  $r$  (et le mode considéré) pendant l'intervalle de temps  $(t, t + \Delta t)$  est donné par :

$$M_{ij}^r = \frac{\exp\left(\frac{C_{ij}^r(t)}{\mu_d^r}\right) \Delta t}{\int_0^t \exp\left(\frac{C_{ij}^r(u)}{\mu_d^r}\right) du}$$



où les temps de départ se situent dans l'intervalle  $[0, T]$  ; le paramètre d'hétérogénéité du modèle de choix de l'heure de départ (logit multinomial) est noté  $\mu_d^r$ . La valeur incluse  $\Xi_{ij}^r$  correspondant au couple  $\{i, j\}$  et au mode  $r$  est une mesure de surplus et peut être calculée de manière exacte dans le cas du logit continu. Il est donné par la formule suivante :

$$\Xi_{ij}^r = \mu_d^r \log \left\{ \int_0^T \exp \left( \frac{C_{ij}^r(u)}{\mu_d^r} \right) du \right\}$$

qui correspond à la formule dite du **Logsomme**.

Dans ce cas, les matrices statiques qui sont le résultat de cette procédure sont affectées de 4 séries de valeurs de paramètres.

- La valeur unitaire du temps : moyenne  $\alpha$ , variance  $\sigma(\alpha)$ , et type de distribution  $D(\alpha)$ .
- Coût unitaire des arrivées précoces : moyenne  $\beta$ , variance  $\sigma(\beta)$ , et type de distribution  $D(\beta)$ .
- Coût unitaire des arrivées tardives : moyenne  $\gamma$ , variance  $\sigma(\gamma)$ , et type de distribution  $D(\gamma)$ .
- Heure d'arrivée désirée à la destination : moyenne  $\hat{\tau}$ , variance  $\sigma(\hat{\tau})$ , et type de distribution  $D(\hat{\tau})$ .

A priori les valeurs des paramètres dépendent des motifs de déplacements. Cependant, des hypothèses sont nécessaires pour simplifier l'estimation des paramètres.

## 5.2.7 Affectation

Enfin, étant donné le choix modal et le choix de l'heure de départ, l'individu choisit son itinéraire sur le réseau correspondant au mode sélectionné. Il s'agit de résoudre un problème de minimisation d'un coût généralisé de déplacement, pour lequel des algorithmes rapides existent. Par exemple, nous utiliserons des algorithmes polynomiaux dans le logiciel METROPOLIS. Ces algorithmes étendent au cas dynamique les procédures classiques de type Dijkstra. Dans le cas dynamique, considéré en priorité ici, le temps de parcours pour traverser un tronçon de route dépend de l'heure d'arrivée au début de ce tronçon. En présence d'une tarification, les algorithmes de chemin le plus court, qui prennent en compte le temps de parcours, des pénalités à l'arrivée, ne sont plus polynomiaux, mais comme l'ont montré de Palma, Hansen et Labbé (1990), des procédures bi-critères peuvent dans ce cas être utilisées.

L'affectation nécessite de modéliser les réseaux de transports. Les graphes correspondant à ces réseaux utilisés dans METROPOLIS (routier et transports en commun) sont ceux du modèle de l'IAURIF. Ils sont décrits dans la partie 7 (« Construction des bases de données »).

## Rétroaction

Nous nous proposons de tester la convergence du modèle en considérant une procédure itérative en 7 étapes décrite ci-dessous :

- ETAPE 1 : Calculs des émissions et des attractions
- ETAPE 2 : Evaluation des temps de parcours à vide de zone à zone
- ETAPE 3 : Enregistrement des temps de parcours effectifs (qui sont les temps de trajet à vide à la première itération) et les Logsommes effectifs
- ETAPE 4 : Calcul de la matrice O-D par motif de déplacement sur la base des temps de parcours effectifs (ETAPE 3) : résolution de l'algorithme appelé « Réalité »
- ETAPE 5 : Calcul du choix modal sur la base des Logsommes effectifs calculés à l'ETAPE 3 et calcul des matrices O-D en VP par type de motifs sur la base des matrices O-D calculées dans l'ETAPE 4
- ETAPE 6 : Affectation dans METROPOLIS de la matrice VP calculée à l'ETAPE 5
- ETAPE 7 : Soit arrêt (si critère de convergence satisfait) soit allez en ETAPE 3.

Cette procédure itérative permet de calculer une matrice origine-destination voiture particulière cohérente.

Il est possible d'introduire UrbanSim au niveau de l'ETAPE 1, mais cette extension ne sera considérée que dans des recherches futures si elle s'avère justifiée.

### ***5.3 Modèle d'affectation en transports en commun***

A ce stade de la recherche, nous envisageons deux méthodes possibles :

- développer notre propre algorithme d'affectation en transports en commun
- utiliser le logiciel de l'IAURIF, DAVISUM, pour calculer la matrice de temps généralisés sur le réseau de transports en commun. Cette matrice ne dépend que des missions et du graphe TC et non de la capacité du réseau, donc elle est indépendante de la demande de déplacements et donc de l'évolution de l'urbanisation. On pourrait par conséquent, en phase de pré-traitement, faire tourner le modèle TC de l'IAURIF au préalable à différents horizons (en tenant compte des projets et des nouvelles missions) puis extraire les matrices de temps généralisés. Ces matrices issues de ces simulations seraient ensuite entrées de manière exogène dans METROPOLIS à chaque étape de simulation. Cette méthode a le mérite d'être en pratique rapidement mise en œuvre mais se fait au détriment de l'automatisation souhaitée.

Cette question fondamentale sera abordée au cours de la deuxième année du projet.

Il faut savoir que dans la prochaine version de DAVISUM (9.24) prévue pour la fin de l'année 2004, le logiciel METROPOLIS sera intégré comme une des fonctionnalités de DAVISUM. En effet, un partenariat de développement en commun a été mis en place entre la société PTV (concepteur et éditeur de DAVISUM) et M. de Palma. On aura alors dans un même outil le logiciel METROPOLIS pour la route et DAVISUM pour les transports en commun.

Au niveau de l'exécution du modèle de trafic, il est inutile de le faire tourner chaque année, contrairement à UrbanSim. Il faut le faire les années de mise en service d'un projet important.

Ainsi, si on s'intéresse aux transports en commun, les projets suivants ont été réalisés entre 1990 et 1999 :

- Entre 1990 et 1993 : T1 et TVM
- Entre 1993 et 1996 : T2 et la liaison La Verrière-La Défense
- Entre 1996 et 1999 : Meteor, Eole, la gare de Saint-Denis-Stade de France (ligne D), la station Saint-Denis-Université (ligne 13), le déplacement des gares Plaine Voyageurs (ligne B), Bibliothèque François-Mitterrand (ex boulevard Masséna) et le prolongement du RER C de Montigny-Beauchamp à Pontoise.

On peut donc envisager 2 matrices intermédiaires de temps TC entre 1990 et 1999 : la matrice de 1993 et la matrice de 1996.

On fera de même pour le mode VP.

## 5.4 Synthèse

La procédure que nous n'avons pas décrite ici est la procédure de calage des paramètres intervenant dans les modèles de génération, distribution, choix modal et dans le modèle de choix de l'heure de départ. La raison en est que les procédures décrites ci-dessus seront automatisées, tandis que les procédures de calages restent encore de l'ordre de « l'opération manuelle », et sont très peu décrites dans la littérature car elles relèvent essentiellement du savoir-faire et de l'expérience du praticien.

Nous n'avons pas discuté des rétroactions. Une manière rigoureuse de traiter cette procédure d'évaluation des matrices O-D consiste à utiliser pour les étapes de choix modal et de distribution les temps de parcours issus d'une affectation statique (DAVISUM, par exemple) ou dynamique (METROPOLIS, par exemple). Dans ce cas, les temps de parcours sont calculés de manière endogène. Cependant, trois critiques peuvent être adressées à cette procédure :

- elle est lourde d'un point de vue numérique, car elle nécessite la résolution d'un point fixe,
- elle suppose probablement une rationalité exagérée de la part des individus,
- elle entraîne des difficultés au niveau de l'estimation économétrique, car elle implique des problèmes d'endogénéité.

Néanmoins, elle offre une manière rationnelle de calculer une solution stationnaire dans un système de transport. Dans le cas du modèle à long terme UrbanSim-METROPOLIS, nous utiliserons cette procédure pour calculer les conditions initiales de notre système dynamique. En d'autres termes, nous supposerons que la première année, la matrice des déplacements est évaluée en fonction des temps de parcours obtenus par l'affectation de cette même matrice origine-destination.

## **6 Architecture du modèle intégré UrbanSim-METROPOLIS**

### ***6.1 Introduction***

Le modèle intégré est constitué de deux modèles principaux, celui d'urbanisation « UrbanSim » et celui de trafic « METROPOLIS ». Ces deux fonctionnent actuellement de manière individuelle. Du côté d'UrbanSim, la matrice des Logsommes représentant le surplus des usagers se déplaçant entre des paires de zones est introduite comme une donnée exogène au début de chaque cycle de simulation. Du côté de METROPOLIS, il faut introduire une matrice O-D qui devra être produite à partir des données concernant la distribution géographique de la population et des emplois issue du modèle d'urbanisation. Le modèle intégré est constitué des programmes nécessaires pour calculer des matrices O-D ainsi que les transformations nécessaires sur les données échangées entre les deux modèles.

La simulation dynamique d'une période donnée nécessite le lancement successif et répété des deux modèles principaux ainsi que les scripts précédemment décrits qui est un travail long et difficile à faire manuellement. Le lancement manuel de ces composants peut également induire beaucoup d'erreurs. C'est pourquoi nous développons un module prenant en charge l'ordonnancement et la gestion de toutes les étapes d'une simulation ainsi que certains paramètres de contrôle du modèle intégré tels que la période de simulation, les années de l'exécution du modèle d'urbanisation et celui du modèle de trafic et les totaux de contrôle. Ce module, qui bénéficiera d'une interface graphique, pourra produire des indicateurs globaux telles que les mesures d'efficacité qui aideront l'opérateur à suivre les résultats de simulation sans entrer dans les détails. Nous menons actuellement des réflexions sur ce sujet pour les introduire au cours de la deuxième année du projet.

### ***6.2 Structure du modèle intégré***

On propose le schéma suivant pour coupler METROPOLIS et UrbanSim (1°, 2°, 3° sont l'ordre d'exécution des programmes). Le cycle des simulations commence par le lancement du modèle de trafic pour obtenir les indicateurs d'accessibilité. Ensuite, ces résultats seront intégrés dans le scénario de cycle *i* sur lequel UrbanSim sera lancé en deuxième étape. Enfin, le module « Scenario Recreator » se charge de transformer les résultats de la dernière simulation pour créer un nouveau scénario. Le premier scénario est essentiellement basé sur les données de l'année de base.

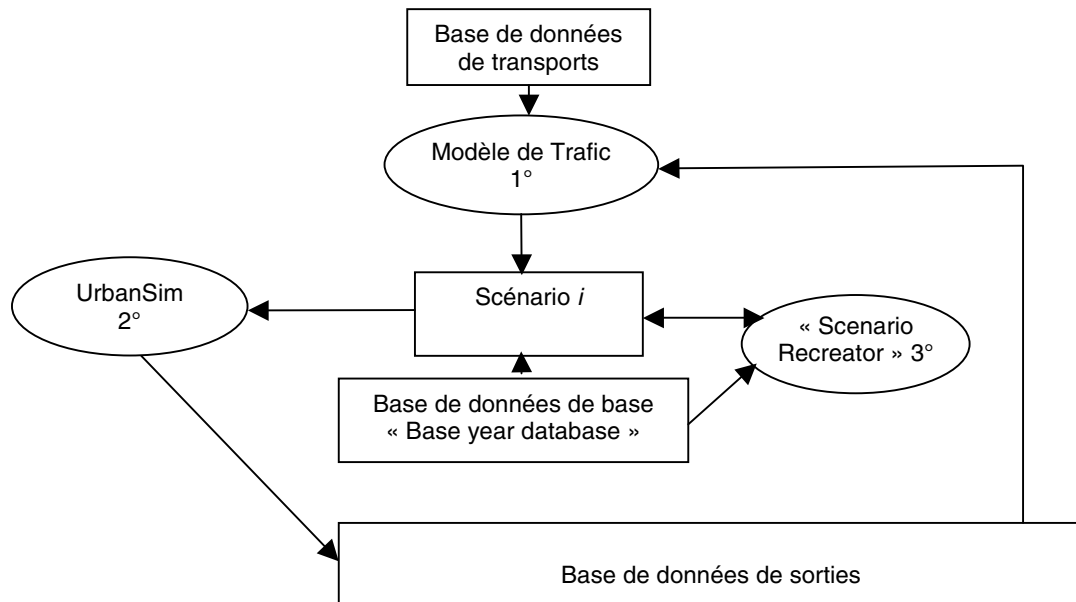


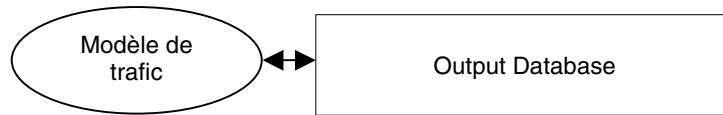
Figure 16 : schéma de l'intégration d'un modèle de trafic et d'UrbanSim

On itère les étapes 1, 2 et 3 sur le scénario  $i$  jusqu'à atteindre la fin de la période souhaitée. Par exemple, pour une simulation de 2000 à 2020 on pourrait avoir :

Année de Base	1999
Scénario 1	2000 à 2004
Scénario 2	2005 à 2009
Scénario 3	2010 à 2015

### 6.3 Echanges de données

Nous commençons par l'échange de données entre le modèle de trafic et la base de données d'UrbanSim. Cet échange concerne deux types de données : les logsommes et les temps de parcours de chaque zone vers certains points prédéfinis de la région tels que le centre de Paris, les centres d'affaires ou les aéroports de la région. Les sorties de METROPOLIS sont stockées dans une base de données gérée par MySQL. Il existe une table pour la logsomme par type d'usager. Les données de ces tableaux doivent être recopiées dans une seule table dans la base d'UrbanSim où il existe une colonne par type d'usager.



Ensuite, les sorties d'UrbanSim, qui comportent notamment la nouvelle distribution spatiale des ménages et des emplois, doivent produire les nouvelles matrices O-D pour alimenter METROPOLIS. Nous avons deux approches possibles pour résoudre ce problème :

- Application du modèle à 5 étapes pour chaque année de simulation :  
La difficulté consiste à générer les émissions et attractions des différentes années pour tous les motifs de déplacement, alors que UrbanSim ne connaît que les données population-emplois. On pourrait néanmoins étendre la base de données pour mémoriser les autres données (écoles, hôpitaux, commerces etc...).
- Application du modèle à 5 étapes pour les premières années de simulation, puis utilisation de la méthode classique dite de Fratar pour projeter les matrices origine-destination

Nous choisissons la première solution pour lequel un script est inséré dans le modèle intégré. Ce script est basé sur les formules estimées par l'IAURIF à partir des résultats de l'Enquête Globale des Transports (EGT).

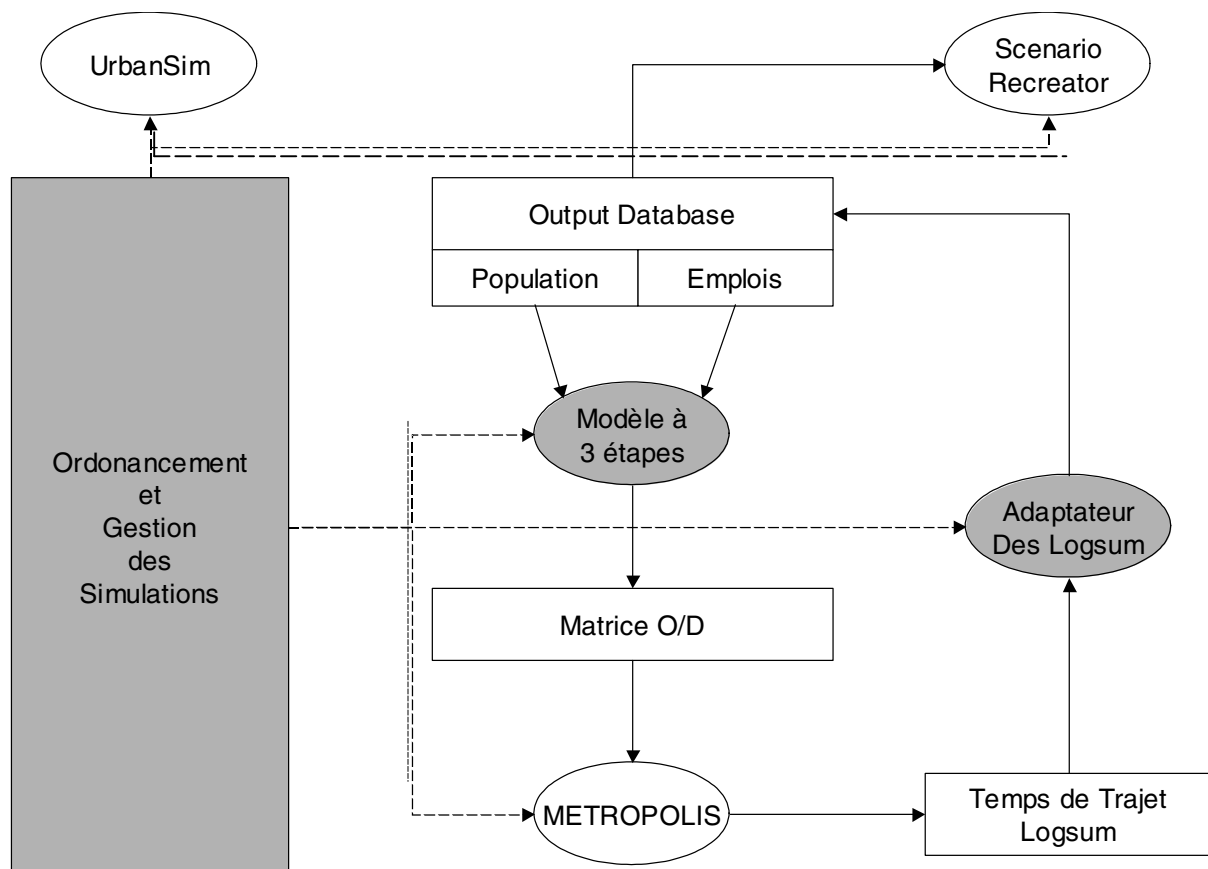


Figure 17 : Echanges entre UrbanSim et METROPOLIS

## **6.4 Description Technique**

Ces modules ont été développés en Windev pour l'interface graphique et les calculs. MySQL sera utilisé comme système de gestion de bases de données.

### **6.4.1 Module d'ordonnancement et gestion des simulations**

Ce module doit permettre de décrire complètement les étapes d'une simulation :

- Données de départ
- Année initiale
- Année cible
- Années d'exécution de METROPOLIS
- Scénarios intermédiaires
- Paramètres

Il doit ensuite permettre l'exécution automatisée et ordonnée des différents modules, à savoir :

- UrbanSim
- METROPOLIS
- ScenarioRecreator
- Modèle de demande à 3 étapes
- Adaptateur des Logsommes

### **6.4.2 Production des matrices Origine-Destination**

Ce module convertit les résultats de UrbanSim en matrices Origine-Destination. Il est exécuté en mode « batch » sur la base des paramètres passés par le module d'ordonnancement.

La description mathématique complète a été donnée au chapitre précédent.

### **6.4.3 Adaptateur des Logsommes**

Ce module convertit les résultats de METROPOLIS au format UrbanSim. Il calcule aussi certains temps de parcours importants pour UrbanSim :

- Temps de parcours de chaque zone aux aéroports
- Temps de parcours de chaque zone à Paris
- ...

Les Logsommes sont pris en compte dans les 9 modules d'UrbanSim.

## 6.5 Description de l'interface graphique du module d'ordonnancement et de gestion

Cette interface a été développée au cours de la première année du projet avec le logiciel WinDev. Elle est présentée ci-après.

### 6.5.1 Gestion des paramètres

L'écran principal de l'interface permet de visualiser et de modifier les paramètres de base du projet (onglets « Project », « Scenario », « Base Year », «Travel Model » )

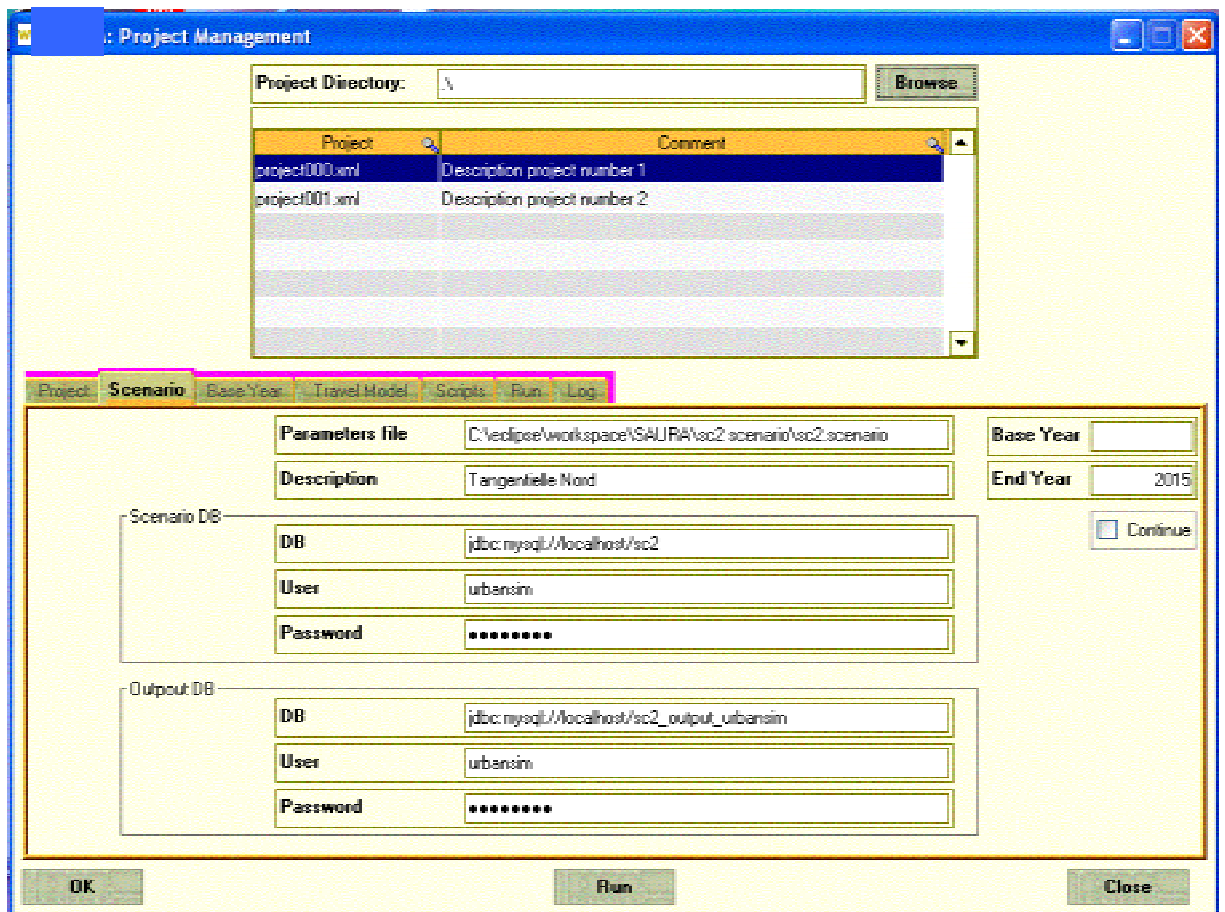


Figure 18 : Fenêtre de gestion des paramètres du modèle intégré



## 6.5.2 Gestion avancée des paramètres

L'écran principal de l'interface permet de visualiser et de modifier des fichiers « script », ceux-ci sont principalement de type :

- « xml » standard
- « sql » au format des commandes Mysql acceptant des paramètres de substitution au format « \$nom\_du\_paramètre », « nom\_du\_paramètre » étant un paramètre défini dans les scripts du projet ou du scénario
- « txt » (édition de fichier texte standard : txt, bat, ...)

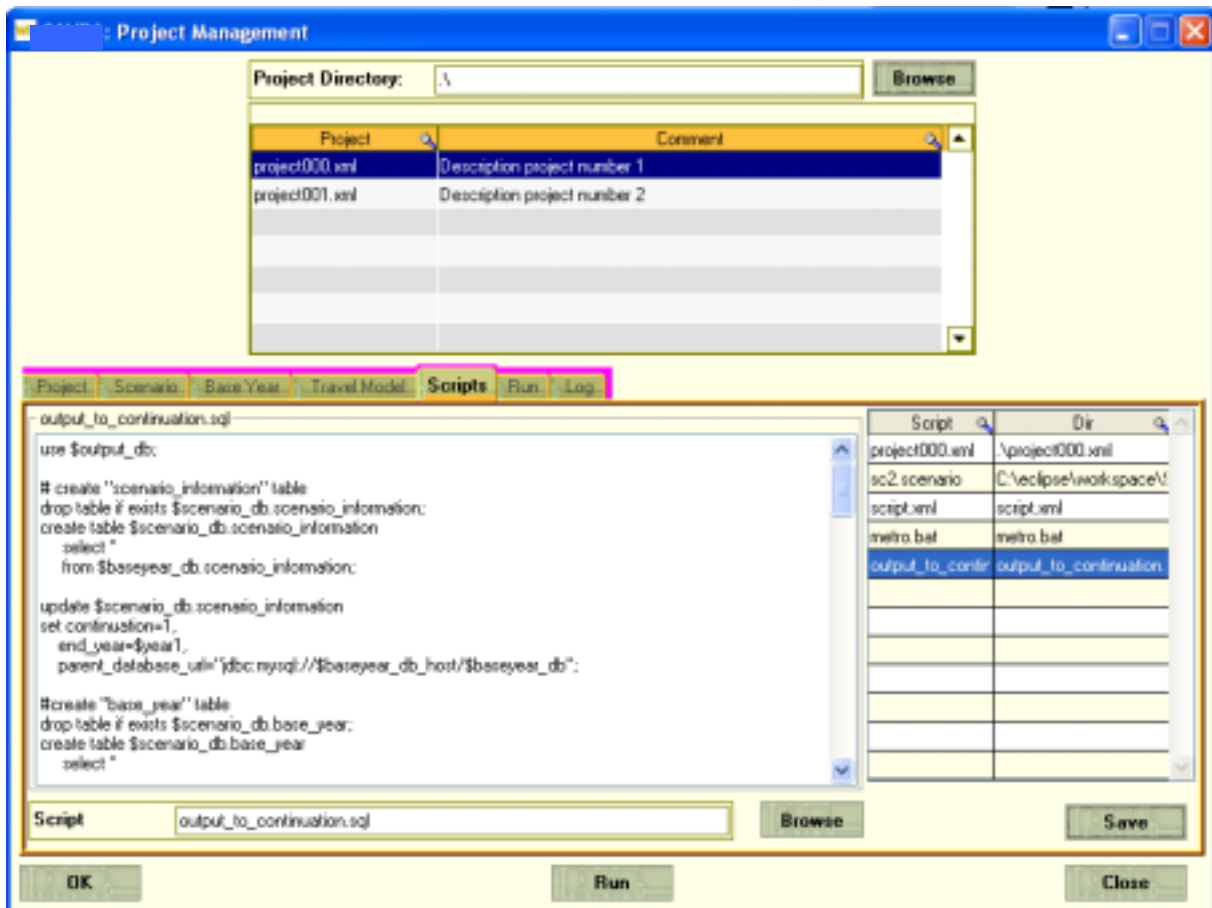


Figure 19 : Fenêtre de la gestion avancée des scripts

### 6.5.3 Gestion de l'exécution

L'écran principal de l'interface permet de visualiser les étapes d'exécution et leur déroulement (onglet «Run »)

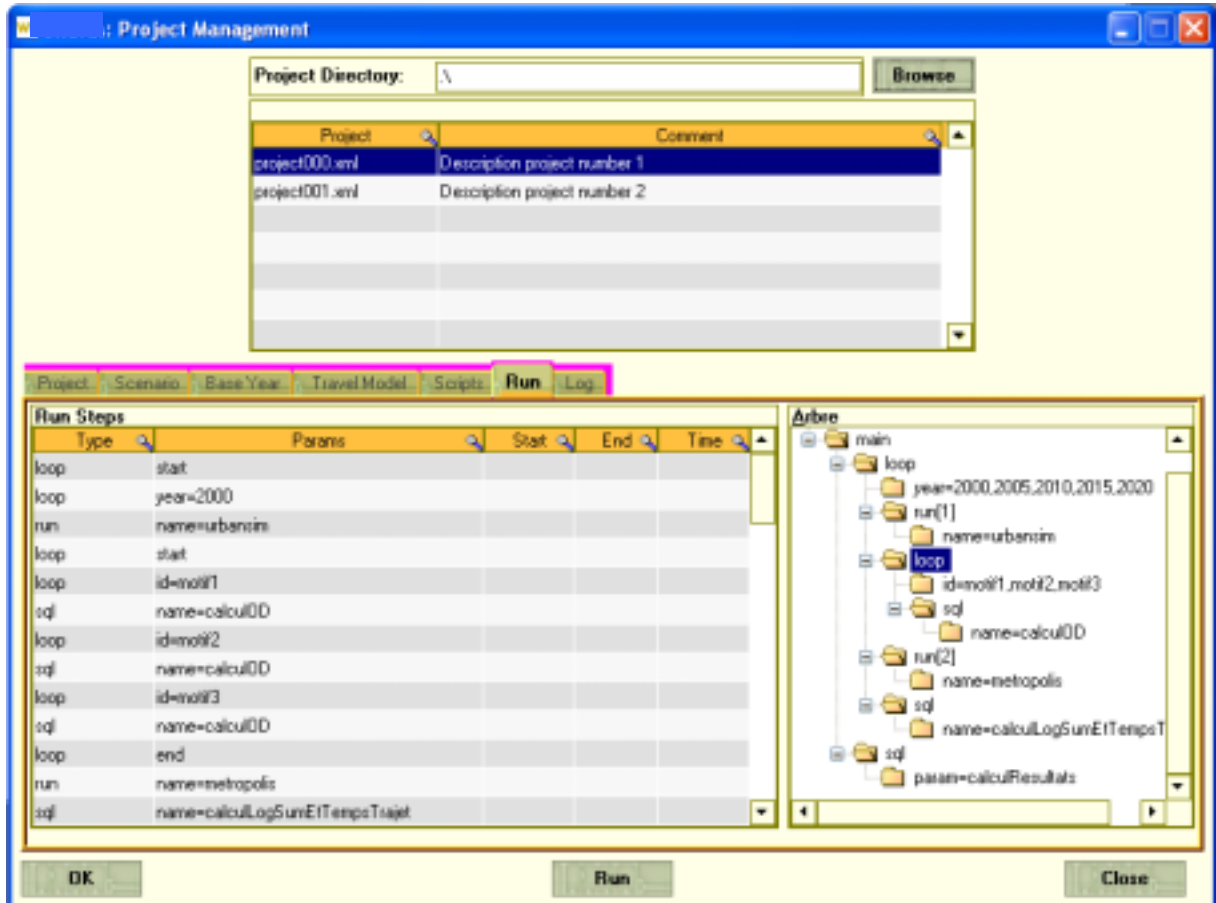


Figure 20 : Fenêtre de la gestion d'exécution des modules

L'onglet « Log » permet de visualiser les messages des différentes étapes.

## 6.6 Synthèse

L'interface a donc été conçue pour permettre des échanges entre UrbanSim et METROPOLIS de manière automatisée. Elle permettra d'automatiser le processus suivant :

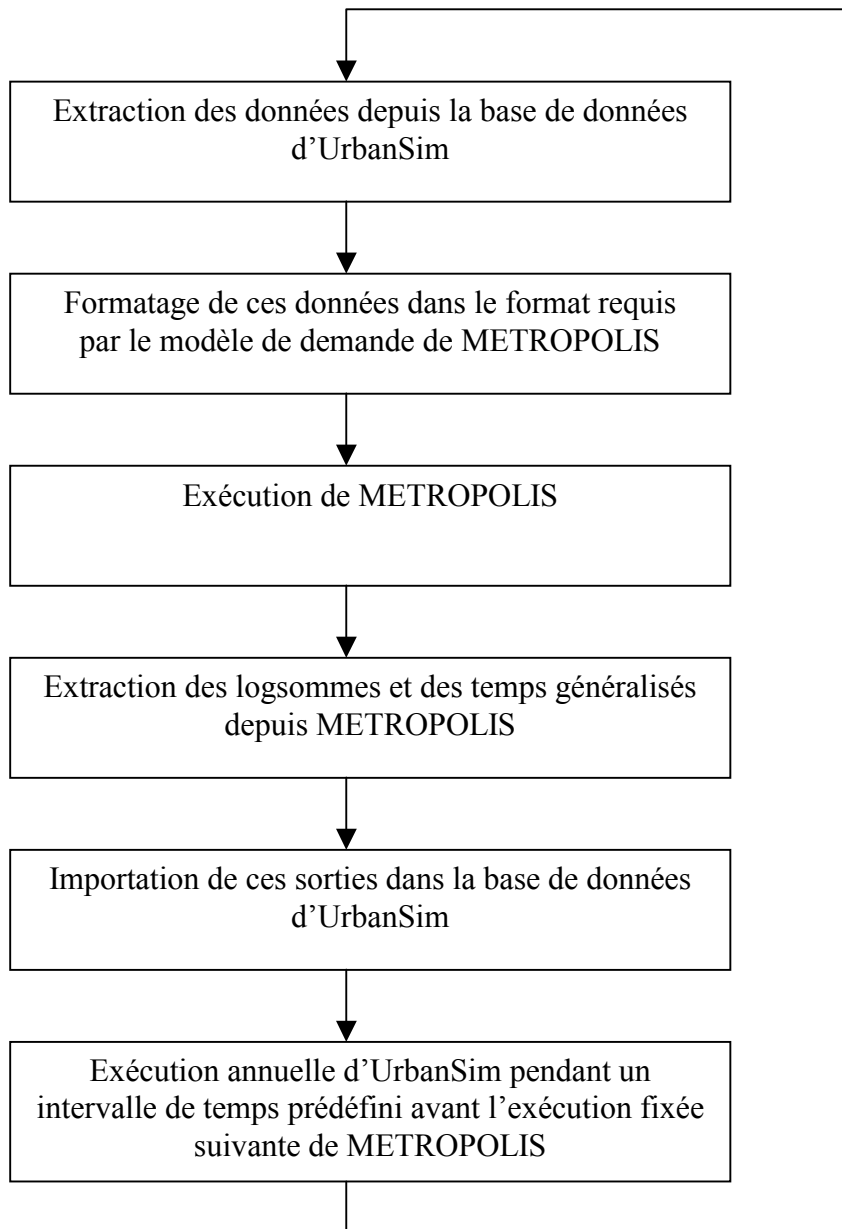


Figure 21 : Enchaînement des tâches

Au cours de la deuxième année de notre recherche, cette architecture sera complétée, testée et améliorée.

## 7 Construction des bases de données Ile-de-France

### 7.1 Introduction

L'aire d'étude est l'ensemble de la Région d'Ile-de-France. La surface est de 12 000 km<sup>2</sup> pour une population totale de 11 millions d'habitants (dont 2 millions à Paris) et 5 millions d'emplois en 1999, année du dernier Recensement général de la population. Pour l'application au projet de la Tangentielle nord, on pourra envisager d'extraire un secteur d'étude avant de faire tourner le programme. Nous avons décidé de recueillir les données sur toute la région pour se donner la possibilité d'appliquer ultérieurement le modèle intégré SIMAURIF à tout autre secteur d'étude.

UrbanSim requiert en entrée une masse de données très importante dans un zonage bien défini. Ce zonage est un **quadrillage très fin en mailles carrées**, le nombre de mailles (ou cellules) étant au choix de l'utilisateur. UrbanSim travaille à un niveau fin car il calcule des variables d'accessibilité locale pour chaque maille. De plus, on sait qu'au sein d'une même commune, il existe de très fortes disparités au niveau des types urbains. Travailler à un niveau infracommunal permet donc de prendre en compte cette hétérogénéité.

Les différents modules d'UrbanSim requièrent également des bases de données de calage qui seront construites à partir d'enquêtes diverses. Pour donner un exemple, le choix résidentiel dépend de nombreux facteurs socio-économiques intrinsèques au ménage, du cycle de vie du ménage, de l'offre foncière, de l'accessibilité aux transports et bien d'autres facteurs encore. Les Anglo-Saxons ont défini le concept de « pull/push » pour décrire les facteurs qui interviennent dans le choix de localisation résidentielle (on fuit son logement actuel pour un autre, ou on est attiré par un autre). Les interactions entre ces facteurs sont nombreuses et on peut même se poser la question de la rétroaction (dans quelle mesure la mobilité résidentielle des ménages influe-t-elle sur le marché de l'offre ?). Seule une enquête logement par panel de type longitudinal peut nous permettre d'appréhender les phénomènes de la mobilité résidentielle et donc de ne retenir que les facteurs déterminants. Pour cela, on essaiera de construire des fichiers d'observations décrivant les « trajectoires » résidentielles spatiales et temporelles d'un échantillon de ménages. Ils permettront d'une part de calibrer les paramètres des modèles de choix discrets et les modèles de régression linéaire, d'autre part de valider les modèles. Les enquêtes ménages utilisées dans notre projet sont décrites dans la partie 7.3.

Le Professeur P. Waddell préconise la construction des bases de données et de calage en suivant une méthodologie en 14 étapes :

1. Définition du secteur d'étude
2. Génération d'un maillage
3. Attribution d'un identifiant unique à chaque cellule
4. Intersection du maillage avec le Mode d'Occupation du Sol (MOS) et le zonage du modèle
5. Affectation des caractéristiques du MOS et données communales au maillage

6. Affectation des limites au maillage
7. Affectation des contraintes environnementales au maillage
8. Affectation d'un type d'urbanisation à chaque cellule du maillage
9. Synthèse d'un fichier ménages et localisation sur le maillage
10. Affectation des données d'emploi au maillage
11. Mise en correspondance entre surfaces non résidentielles et emploi sur le maillage
12. Elaboration d'un diagnostic et résolution des incohérences de la base de données
13. Chargement de la base dans le SGBD SQL
14. Vérification de l'intégrité de la base de données

Nous avons suivi cette méthodologie à la lettre mais pour la clarté du rapport, nous avons choisi de présenter les résultats de cette collecte et de la mise en forme, à partir de la structure même des différents fichiers informatiques.

On a donc distingué **la base de données en entrée** d'UrbanSim (7.2) et **la base de données de calage** (7.3).

## ***7.2 Base de données en entrée d'UrbanSim***

### **7.2.1 Structure de données**

Elle a été décrite dans le paragraphe 4.4. Rappelons qu'il y a quatre types de tables :

- données
- constantes
- paramètres
- scénarios

On s'intéresse ici aux types de tables « données » et « constantes ». Les deux autres types de tables seront construits au cours de la deuxième année du projet (les tables « paramètres » seront remplis une fois le calage effectué, les tables « Scénarios » seront constitués lors de l'élaboration des scénarios de l'application du modèle intégré à la Tangentielle nord).

Il y a 9 tables de type données. Nous allons les décrire un par un et la manière dont nous les avons construits. Les trois plus importants sont gridcells.tab, households.tab et jobs.tab.

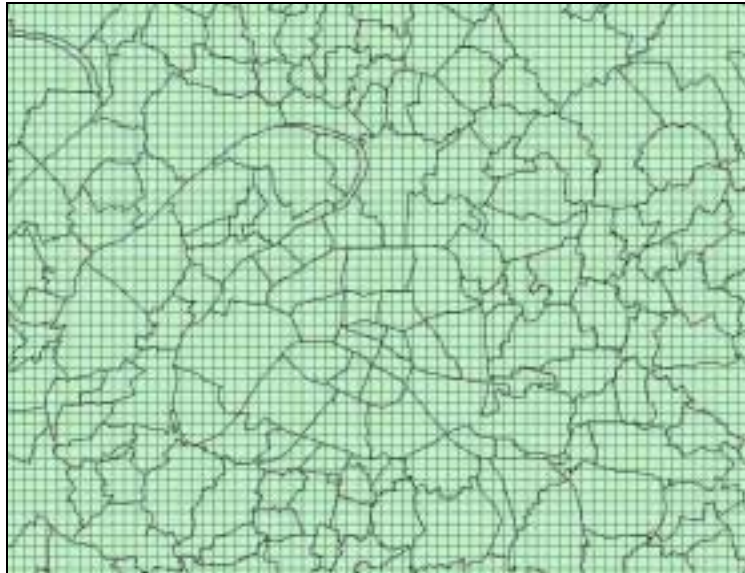
La plupart des colonnes de ces tables ont été créées avec Arc-Info et Arc-Gis, 2 logiciels utilisés à l'IAURIF, qui permettent de consulter, produire, mettre à jour et analyser les bases de données géoréférencées du Système d'Information Géographique Régional (SIGR). Une fois toutes les tables construites, les tables seront transférées dans le Système de Gestion de Bases de Données (SGBD) MySQL.

Les bases de données sont constituées à 2 années : 1990 et 1999. Ce qui signifie que la plupart des traitements décrits ci-après ont été effectuées deux fois !

## 7.2.2 Table « gridcells »

→La première étape a consisté à générer un maillage.

La région a été découpée en carrés de 500 m de côté. Il y a ainsi 49 237 cellules carrées.



Carte n°2 : Extrait du maillage, avec limites de communes

Ce maillage constitue une couverture Arc-Info ayant une topologie de polygones. Il a été créé avec la fonction GRID d'Arc-Info.

Un identifiant unique par cellule est attribué automatiquement par Arc-Info. Il va de 1 à 49237, constituant la première colonne (ou variable) appelée **GRID-ID** de la table gridcells.tab. Remarquons pour l'aspect pratique que compte tenu du nombre de cellules, le fichier est gérable sur le tableur Excel !

Les coordonnées du centroïde (**RELATIVE\_X** et **RELATIVE\_Y**) de chaque cellule sont aussi calculées automatiquement par Arc-Info dans le système de coordonnées Lambert II.

Les trois premières colonnes étant obtenues, il s'agit maintenant de remplir la table gridcells.tab par les 28 autres colonnes.

Champ		Type	Unité	Valeur
GRID_ID	1	Entier		De 1 à 49237
RELATIVE_X	2	Entier		Lambert II
RELATIVE_Y	3	Entier		Lambert II
ZONE_ID	4	Entier		De 1 à 572
CITY_ID	5	Entier		Code INSEE
COUNTY_ID	6	Entier		N° département
PLAN_TYPE_ID	7	Entier		0, 1 ou 2
IS_OUTSIDE_URBAN_GROWTH_BOUNDARY	8	Entier		0 ou 1
FRACTION_RESIDENTIAL_LAND	9	Flottant		De 0 à 1
RESIDENTIAL_UNITS	10	Entier		Nb logements
COMMERCIAL_SQFT	11	Entier	m2	
INDUSTRIAL_SQFT	12	Entier	m2	

GOVERNMENTAL_SQFT	13	Entier	m2	
COMMERCIAL_IMPROVEMENT_VALUE	14	Entier	Euros	
INDUSTRIAL_IMPROVEMENT_VALUE	15	Entier	Euros	
GOVERNMENTAL_IMPROVEMENT_VALUE	16	Entier	Euros	
NONRESIDENTIAL_LAND_VALUE	17	Entier	Euros	
RESIDENTIAL_IMPROVEMENT_VALUE	18	Entier	Euros	
RESIDENTIAL_LAND_VALUE	19	Entier	Euros	
YEAR_BUILT	20	Entier	Années	
DISTANCE_TO_ARTERIAL	21	Entier		
DISTANCE_TO_HIGHWAY	22	Entier		
PERCENT_WATER	23	Entier		de 0 à 100
PERCENT_STREAM_BUFFER	24	Entier		
PERCENT_FLOODPLAIN	25	Entier		
PERCENT_WETLAND	26	Entier		
PERCENT_SLOPE	27	Entier		
PERCENT_OPEN_SPACE	28	Entier		
PERCENT_PUBLIC_SPACE	29	Entier		
PERCENT_ROADS	30	Entier		
DEVELOPMENT_TYPE_ID	31	Entier		Formule

Tableau 11 : Caractéristiques d'une cellule

→ La deuxième étape a consisté à croiser le maillage avec différentes couches disponibles du SIGR :

- le Mode d'Occupation du Sol (MOS)
- le découpage du modèle de l'IAURIF en 606 zones
- le découpage administratif en 1300 communes de la région
- le zonage du SDRIF de 1994
- la couverture des îlots INSEE de population

### 1) Le MOS

Le MOS est l'inventaire de l'occupation du sol sur toute la région en 83 postes, regroupables en 49, 21, 10 ou 3 postes. Il est disponible aux années 1990 et 1999.

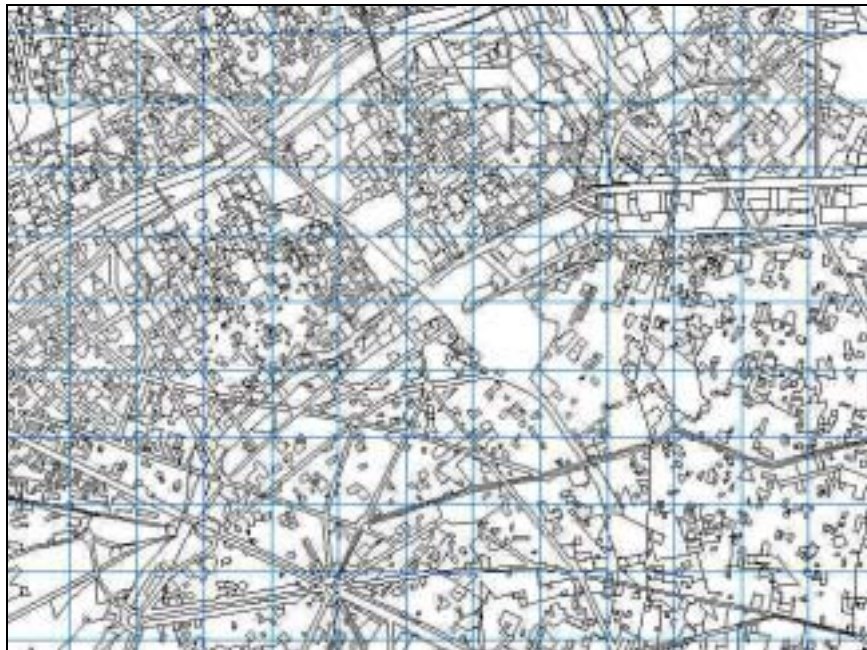
Voici la nomenclature en 83 postes :

code	Libellé	code	Libellé
1	BOIS OU FORÊTS		
2	COUPES OU CLAIRIÈRES EN FORÊTS	43	CENTRES COMMERCIAUX
3	PEUPLERAIES	44	HYPERMARCHES
4	TERRES LABOURÉES	45	GRANDS MAGASINS
5	SURFACES EN HERBE A CARACTÈRE AGRICOLE	46	STATIONS-SERVICE
6	VERGERS, PÉPINIÈRES	47	AUTRES COMMERCES
7	MARAÎCHAGE, HORTICULTURE	48	BUREAUX
8	CULTURES INTENSIVES SOUS SERRES	49	INSTALLATIONS SPORTIVES COUVERTES
9	EAU FERMÉE (ÉTANGS, LACS...)	50	CENTRES ÉQUESTRES
10	COURS D'EAU	51	PISCINES COUVERTES
11	SURFACES EN HERBE NON AGRICOLES	52	PISCINES EN PLEIN AIR
12	CARRIÈRES, SABLIERES	53	AUTODROMES
13	DÉCHARGES	54	ENSEIGNEMENT DE PREMIER DEGRE
14	ESPACES RURAUX VACANTS (MARAIS,FRICHES...)	55	ENSEIGNEMENT SECONDAIRE
15	BERGES	56	ENSEIGNEMENT SUPÉRIEUR
16	PARCS LIES AUX ACTIVITÉS DE LOISIRS	57	ENSEIGNEMENT AUTRE

17	PARCS OU JARDINS	58	HÔPITAUX, CLINIQUES
18	JARDINS FAMILIAUX	59	AUTRES ÉQUIPEMENTS DE SANTÉ
19	JARDINS DE L'HABITAT INDIVIDUEL	60	CIMETIERES
20	JARDINS DE L'HABITAT RURAL	61	MAIRIES
21	JARDINS DE L'HABITAT CONTINU BAS	62	MARCHÉS PERMANENTS
22	TERRAINS DE SPORT EN PLEIN AIR	63	LIEUX DE CULTE
23	TENNIS DÉCOUVERTS	64	GRANDS ÉQUIPEMENTS CULTURELS
24	BAIGNADES	65	ÉQUIPEMENTS DE PROXIMITÉ
25	CAMPING, CARAVANING	66	SIÈGES D'ADMINISTRATIONS TERRITORIALES
26	PARCS D'ÉVOLUTION D'ÉQUIPEMENTS SPORTIFS	67	ÉQUIPEMENTS DE MISSION SECURITE CIVILE
27	GOLFS	68	INSTALLATIONS RADIO-ÉLECTRIQUES
28	HIPPODROMES	69	ADMINISTRATIONS AUTRES
29	TERRAINS VACANTS EN MILIEU URBAIN	70	PRODUCTION D'EAU
30	HABITAT INDIVIDUEL	71	ASSAINISSEMENT
31	ENSEMBLES D'HABITAT INDIVIDUEL IDENTIQUE	72	ÉLECTRICITÉ
32	HABITAT RURAL	73	GAZ
33	HABITAT CONTINU BAS	74	PÉTROLE
34	HABITAT COLLECTIF CONTINU HAUT	75	INFRASTRUCTURES AUTRES
35	HABITAT COLLECTIF DISCONTINU	76	EMPRISES DE TRANSPORT FERRÉ
36	PRISONS	77	AUTOROUTES
37	HABITAT AUTRE	78	VOIES DE PLUS DE 25M D'EMPRISE
38	ACTIVITÉS EN TISSU URBAIN MIXTE	79	PARKINGS DE SURFACE
39	GRANDES EMPRISES D'ACTIVITÉ	80	PARKINGS EN ÉTAGES
40	ZONE ET LOTISSEMENT AFFECTES AUX ACTIVITES	81	GARES ROUTIÈRES, DÉPÔTS DE BUS
41	ENTREPOSAGE A L'AIR LIBRE	82	INSTALLATIONS AÉROPORTUAIRES
42	ACTIVITÉ DE PRODUCTION ANIMALE	83	CHANTIERS

Tableau 12 : nomenclature du modèle d'occupation des sols (MOS). Source : IAURIF

La fonction d'intersection d'Arc-Info a été appliquée entre la couverture de maillage et la couverture du MOS qui comporte 370 000 polygones. Elle permet de générer une table de 49237 lignes avec pour chacune la répartition en surface au sol des 83 postes du MOS.



Carte n° 3 : Extrait de la couverture intersection entre le maillage (en bleu) et MOS (en noir)

Les 83 postes ont été regroupés en 49, 21, 10 et 3 postes pour faciliter les tâches d'extraction et de calcul.

La **colonne 8 (IS\_OUTSIDE\_URBAN\_GROWTH\_BOUNDARY)** est le résultat d'un calcul sur les postes du « MOS en 3 postes » :

- 0 si code « MOS en 3 postes » =2 ou 3 (espaces urbains)
- 1 si code « MOS en 3 postes » =1 (espace rural)

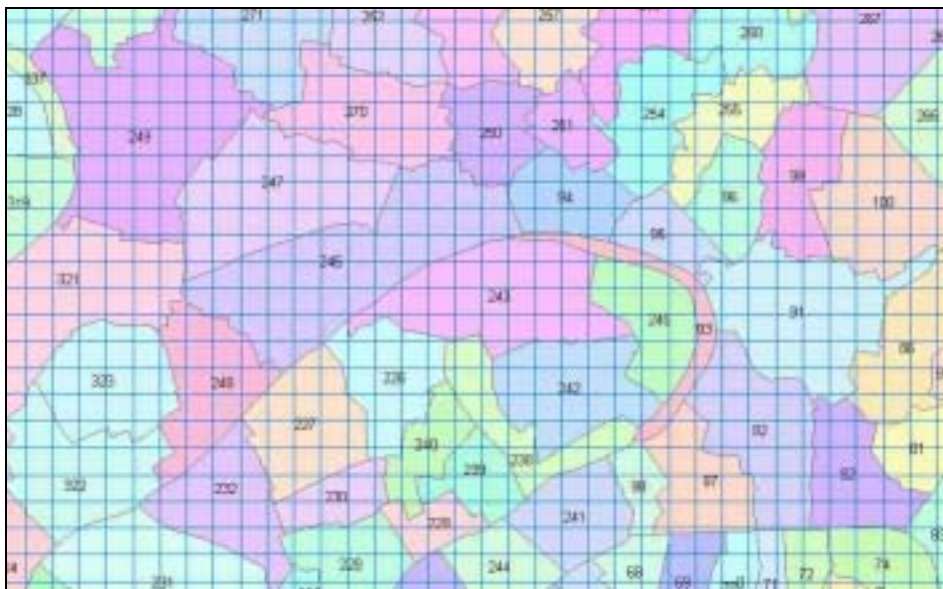


La **colonne 9 (FRACTION\_RESIDENTIAL\_LAND)** est la part pour chaque cellule de surface résidentielle. Elle résulte d'un calcul sur le « MOS en 21 postes » : somme des aires des postes 9 (habitat individuel), 10 (habitat collectif) et 11 (autre habitat) / surface d'une cellule

## 2) Le découpage du modèle en 606 zones

Le découpage du modèle de l'IAURIF comporte 606 zones (il existe aussi un ancien découpage en 522 zones). Il y a 80 zones pour Paris, 165 zones pour la petite couronne (départements 92, 93, 94), 327 zones pour la grande couronne (départements 77, 78, 91, 95) et 34 zones externes à la région. Une zone est soit un quartier pour Paris, soit une commune soit un regroupement de communes. Une cinquantaine de communes importantes ont été découpées en deux.

On a donc intersecté la couverture de maillage et la couverture du découpage en 606 zones. Cette opération permet d'affecter à chaque cellule carrée la zone (**colonne 4 : ZONE-ID**) dans laquelle elle se trouve. Si une cellule est à cheval sur deux zones, c'est la zone de plus grande surface intersectée qui lui est attribuée.



Carte n° 4 : Extrait de la couverture intersection entre le maillage (en bleu) et le zonage en 606 zones

## 3) Les communes

La Région comprend 1300 communes dont les 20 arrondissements de Paris. Chaque commune a un code INSEE à 5 chiffres, les deux premiers étant le numéro du département.

On a intersecté la couverture de maillage et la couverture des communes. Cette opération permet d'affecter à chaque cellule carrée le code INSEE (**colonne 5 : CITY-ID**) dans laquelle elle se trouve. Si une cellule est à cheval sur deux communes, c'est la commune de plus





Carte n° 6 : Extrait de la couverture intersection entre le maillage et le zonage SDRIF 94

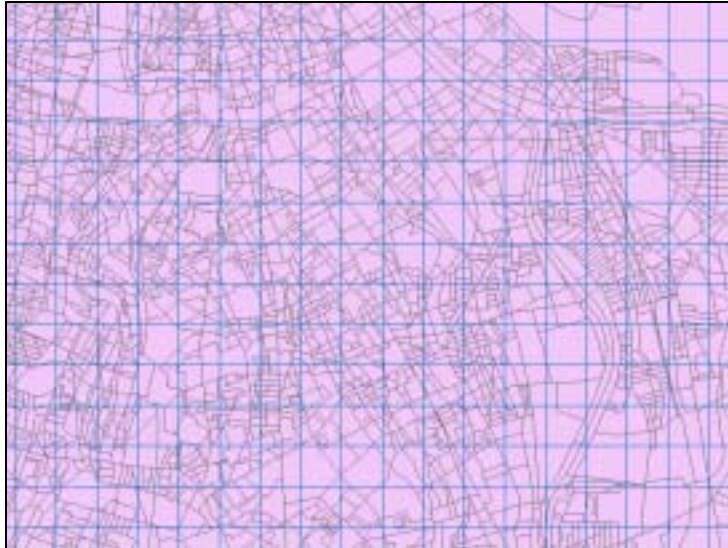
Le croisement des deux couches permet d'affecter à chaque cellule le type 1 ou 2, ou 0 si ce n'est aucun des deux. C'est la **colonne 7 : PLAN\_TYPE\_ID**.

Notons que dans le cadre de la révision du SDRIF qui va débuter en novembre 2004, si une carte de destination des sols était élaborée avant la fin de la deuxième année de notre recherche, alors on intégrerait ce nouveau zonage.

### 5) Les îlots INSEE de densité de population

L'IAURIF dispose de la couche des îlots INSEE. Ce thème contient la description graphique des îlots des recensements de 1982, 1990 et 1999 avec des données relatives à la population et au logement connues à l'îlot. Le croisement entre le maillage et cette couche fournit la **colonne 8 : RESIDENTIAL\_UNITS**. C'est le nombre de logements par cellule, comprenant les logements occupés et les logements vacants. Lorsqu'un îlot est à cheval sur deux cellules, on affecte le nombre de logements au prorata de la surface de la cellule intersectée.

La couverture des îlots INSEE comporte environ 52 000 polygones.



Carte n° 7 : Extrait de la couverture intersection entre le maillage et les îlots INSEE

→ La troisième étape a consisté à remplir les colonnes 11 à 17 qui concernent des données sur les surfaces et les loyers des commerces, des activités secondaires et des activités tertiaires.

### 1) Les surfaces de commerces

L'IAURIF dispose d'un fichier alpha-numérique de surfaces de vente par commune de plus de 900 m<sup>2</sup>. Ces données ne sont pas géolocalisées. Il a fallu trouver un moyen de répartir ces données sur le maillage. Le principe en est le suivant : on identifie parmi les postes du MOS en 83 postes ceux concernant les commerces puis on distribue au prorata de la surface des polygones de type commerces du MOS les surfaces de vente à la commune. Il est nécessaire de faire ce traitement car le MOS ne donne que des surfaces au sol alors que le fichier de surface de vente tient compte des étages (la 3<sup>ème</sup> dimension) dans les équipements commerciaux (les commerces sont identifiés dans le MOS par les postes 43 à 47). Cette méthode permet de remplir la **colonne 11 : COMMERCIAL\_SQFT**.

### 2) Les surfaces d'activités secondaires

Pour la **colonne 12 (INDUSTRIAL\_SQFT)**, on n'a pas de données plus précises que le MOS, ce qui est déjà inespéré. Pour tenir compte de la 3<sup>ème</sup> dimension (les SHON qui sont les surfaces de planchers), on pondère les trois postes activités du MOS par des coefficients moyens d'occupation du sol (COS) qui donnent les densités réelles : 38 (activités en tissu urbain mixte), 39 (grandes emprises d'activités) et 40 (zone et lotissement affectés aux activités)

Ainsi, on calcule INDUSTRIAL\_SQFT par  $1,5 * \text{poste } 38 + 0,5 * \text{poste } 39 + 1 * \text{poste } 40$ . On aurait pu affiner encore les COS par département mais cela n'a pas encore été fait.

### 3) Les surfaces d'activités tertiaires (bureaux)

Pour la **colonne 13 (GOVERNMENTAL\_SQFT)**, on a procédé comme pour les commerces. L'IAURIF dispose d'un fichier de surfaces de bureaux à la commune (source : ORIE) tandis que le MOS identifie les surfaces au sol affectées aux bureaux (postes 48 et 66). On a donc

distribué les surfaces de bureaux connues à la commune sur les deux postes du MOS dans chaque cellule.

Il faut aborder ce domaine de l'immobilier d'entreprise de manière humble car c'est affaire de spécialiste. On se contente pour l'instant de recueillir des données de base sur les années 1990 et 1999. Mais il apparaît indispensable de travailler par la suite avec des experts et des professionnels de l'immobilier d'entreprise pour mieux comprendre les raisons des variations des loyers. Il est clair que le marché des bureaux est très dépendant de la conjoncture économique. Lorsque les entreprises sont soucieuses de faire des économies, et donc de réduire leurs frais généraux, elles reportent leurs projets de déménagements ou d'agrandissement. Il serait aussi intéressant de pouvoir distinguer dans les données les nouveaux bureaux dits « en blanc », c'est-à-dire ceux qui sont construits sans avoir préalablement trouvé de locataire.

#### 4) Les loyers de surfaces d'activités secondaires

L'IAURIF dispose de la base de données de loyers de surfaces d'activités par m<sup>2</sup> et par an à la commune (source : GRECAM). Ce fichier n'est pas exhaustif sur les 1300 communes de la Région, et disponible pour deux années (1992 et 2002) qui sont légèrement décalées par rapport à nos deux années de recueil (1990 et 1999). Il est renseigné sur seulement 150 communes, ce qui est peu satisfaisant. Mais il n'y a rien de mieux. Dans ce cas où des données sont manquantes et existantes à des dates décalées, on utilise deux méthodes :

- il existe un découpage de la Région d'Ile-de-France en 94 zones élaboré par la DREIF et l'IAURIF dans les années 90. Il correspond pour une grande partie aux limites des Syndicats Intercommunaux d'Etude et de Programmation (SIEP) et des Schémas Directeurs Locaux (SDL). On suppose que chaque zone regroupe des communes sur un certain nombre de critères urbains. On utilise ce découpage pour attribuer à une commune non renseignée le loyer moyen de la zone dans laquelle cette commune se trouve.
- Pour passer de 1992 à 1990 et de 2002 à 1999, on applique un coefficient de correction qui est l'indice du logement.

La **colonne 15 (INDUSTRIAL\_IMPROVEMENT\_VALUE)** est la valeur totale en Euros de chaque cellule. Il faut donc multiplier la colonne 12 (en m<sup>2</sup>) par la valeur du loyer (Euros/m<sup>2</sup>).

#### 5) Les loyers de bureaux

L'IAURIF dispose de la base de données de loyers de bureaux par m<sup>2</sup> et par an à la commune (source : enquêtes GRECAM). Ce fichier est lui aussi incomplet par rapport aux 1300 communes de la Région, et disponible pour deux années (1992 et 2002) comme pour les loyers de surfaces d'activités. Le fichier est renseigné sur 240 communes seulement. Pour le compléter dans une première étape, on applique la méthode décrite précédemment avec le découpage en 94 zones. Cette méthode apparaît quelque peu frustrée en regard de l'analyse des spécialistes du thème des prix de location des bureaux en Ile-de-France. Il est prévu dans une seconde étape que l'on affine spatialement cette donnée de manière plus réaliste. Il faut savoir que l'éventail de prix est très large. On s'appuiera donc sur des études réalisées par l'IAURIF sur la géographie des prix. Par exemple, si on veut schématiser, le haut de gamme se trouve

en zone centrale, à l'intérieur d'un périmètre regroupant les principaux centres d'affaires de Bercy à La Défense et à Saint-Lazare. En première couronne, les écarts de prix sont très marqués. Les niveaux de loyers sont très élevés dans la partie médiane des Hauts-de-Seine et le long des lignes du RER en périphérie sud et est, mais assez bas en périphérie nord de Paris. Dans la seconde couronne, la carte des prix se superpose au réseau ferré et aux Villes nouvelles (Saint-Quentin et Marne-la-Vallée arrivant en tête). A noter aussi l'apparition dans les années 90 de deux pôles tertiaires importants, qui se sont fortement développés sur les nœuds de communication : Massy et le secteur de Roissy.

**La colonne 16 GOVERNMENTAL\_IMPROVEMENT\_VALUE** est ensuite calculée en multipliant la colonne 13 par le loyer de bureaux de la cellule.

### 6) Les loyers de surfaces commerciales

Il n'y a aucune donnée sur ce thème en Ile-de-France. On a dû prendre ce qui existait : les loyers des surfaces d'activités secondaires et les loyers de bureaux :

- Pour Paris et les communes de la vallée de la Seine (puissant axe d'activités, un tiers des 36 communes) dans les Hauts-de-Seine : loyer de surfaces commerciales =  $0,5 \times$  loyer de bureaux
- Pour les autres départements et le reste des Hauts-de-Seine : loyer de surfaces commerciales =  $1,2 \times$  loyer de surfaces d'activités secondaires.

Cette simplification a été proposée par un expert de l'IAURIF en développement local. Même si elle est peu satisfaisante, elle a le mérite de permettre au modèle UrbanSim de prendre en compte la variable des loyers de surfaces commerciales.

**La colonne 14 COMMERCIAL\_IMPROVEMENT\_VALUE** est ensuite calculée en multipliant la colonne 11 par le loyer de surfaces commerciales de la cellule.

Dernière précision : tous les loyers de 1990 sont ramenés en Euros constants à la valeur 1999.

### 7) La valeur foncière totale du non résidentiel

**La colonne 17 (NONRESIDENTIAL\_LAND\_VALUE)** est la somme des colonnes 14, 15 et 16.

**→ La quatrième étape a consisté à remplir les colonnes 18, 19 et 20 qui concernent les données sur les valeurs foncières de l'immobilier résidentiel et l'âge moyen.**

**La colonne 18 RESIDENTIAL\_IMPROVEMENT\_VALUE** correspond à la valeur foncière du bâti résidentiel.

L'IAURIF dispose d'une donnée très intéressante et particulièrement difficile à obtenir : le fichier des transactions de la Chambre Notariale de Paris. Ce fichier donne, pour chaque année de 1990 à 2002, le prix moyen par commune de la Région des ventes immobilières, en quatre catégories : appartement ancien (plus de 5 ans), appartement neuf (moins de 5 ans), maison ancienne (moins de 5 ans), maison neuve (plus de 5 ans).

Mais ce fichier n'est pas complet. Il arrive qu'en 1990 et 1999, sur 1300 communes, certaines n'ont pas connu de ventes immobilières. Pour compléter le fichier, on cherche l'année la plus proche où une information existe et on corrige par le taux d'indice du logement.

Lorsqu'aucune donnée n'existe sur la période 1990–2002, on utilise le zonage SDRIF en 94 zones, qui permet d'attribuer le prix moyen de la zone à la commune incluse dans cette même zone.

Par ailleurs, on effectue une requête sur le RGP pour créer une table donnant pour chaque commune la répartition en 4 catégories. Puis on applique ces taux aux cellules appartenant à la commune.

La **colonne 19 RESIDENTIAL\_LAND\_VALUE** correspond à la valeur foncière du terrain. C'est une variable importante aux Etats-Unis, sans doute plus qu'en France. Ici, le prix du terrain est compris dans le prix des logements, à moins de faire construire soi-même. Pour calculer cette variable, on a pour l'instant pris des hypothèses simples, qui pourront être affinées lorsqu'on abordera le calage du module du prix du foncier.

- Si pavillon,  $RESIDENTIAL\_LAND\_VALUE = RESIDENTIAL\_IMPROVEMENT\_VALUE * 20\%$ ,
- Si appartement,  $RESIDENTIAL\_LAND\_VALUE = RESIDENTIAL\_IMPROVEMENT\_VALUE * 10\%$ ,

Quant à la **colonne 20 YEAR\_BUILT**, elle est calculée en prenant, dans la base RGP, la moyenne des années d'achèvement des logements (variable AACHI).

**→ La cinquième étape a consisté à calculer les variables de distances des cellules au réseau routier.**

Pour chacune des 49271 cellules, on calcule la :

- la distance du centroïde à la route nationale la plus proche : **Colonne 21 (DISTANCE\_TO\_ARTERIAL)** :
- la distance du centroïde à l'autoroute ou voie rapide la plus proche : **Colonne 22 (DISTANCE\_TO\_HIGHWAY)** :

Ces calculs sont effectués grâce à la fonction NEAR d'Arc-Info qui est appliquée aux réseaux routiers de 1990 et 1999 du SIGR. Le réseau routier du SIGR est un réseau linéaire calé sur le MOS, hiérarchisé en 6 classes : autoroutes, voies rapides, routes nationales, routes départementales, voies communales importantes et autres. Il comporte environ 200 000 arcs bidirectionnels en 1999 (ce n'est pas le même réseau que celui du modèle qui, lui, est un graphe hamiltonien).

Il est prévu au cours de la deuxième année d'introduire dans la table gridcells, au même niveau que les deux variables précédentes de desserte par le réseau routier, des variables de desserte spatiale (nombre de gares, stations, arrêts de bus par cellule) et temporelle (nombre de passages par mode à l'heure de pointe) en transports en commun. Ces deux variables ont déjà été calculées au niveau de chaque commune.

**→ La sixième étape a consisté à remplir les colonnes 23 à 30. Elles portent sur les contraintes environnementales et sur les emprises de transports.<sup>4</sup>**

L'idée sous-jacente à l'introduction de ces variables est qu'il existe des contraintes à l'évolution du type urbain d'une cellule. Par exemple, si celle-ci est située en zone inondable, elle est non constructible et aura donc peu de chances de voir son type urbain évoluer. Les

<sup>4</sup> Ces colonnes ont été remplies par Laurence Buelens, géomaticienne à l'IAURIF

cellules situées en zones contraintes sont considérées comme figées et sont exclues de l'univers de choix dans les modules de localisation résidentielle et des emplois.

UrbanSim propose 8 variables de contraintes environnementales, qui sont spécifiques au cas américain. En Ile-de-France, nous avons d'autres contraintes, d'autres servitudes que nous pourrions intégrer parmi les colonnes d'UrbanSim. Ce sera fait au cours de la deuxième année. En particulier, on prendra en compte les Zones naturelles d'intérêt écologique, faunistique et floristique (ZNIEFF) et les Zones de protection du patrimoine architectural urbain et paysager (ZPPAUP).

On a donc rempli les 8 colonnes par défaut. Chaque valeur de surface est rapportée à la surface de la cellule pour avoir un pourcentage :

### 1) Colonne 23 : PERCENT\_WATER

On a extrait le poste 4 (eau) du MOS en 21 postes.

### 2) Colonne 24 : PERCENT\_STREAM\_BUFFER

On s'intéresse aux berges des cours d'eau. Le SIGR comporte une couche renseignée sur l'hydrographie en Ile-de-France. Le linéaire hydrographique est hiérarchisé en 1 = permanent > 7,5 m ; 2 = permanent < 7,5 m ; 3 = intermittent ; 4 = abandonné ou sec ; 5 = fictif. En extrayant les types 1 et 2, puis en créant des buffers de 30 mètres de part et d'autre des cours d'eau sélectionnés, on obtient une couche de polygones représentant les berges. Puis on fait une intersection entre le maillage et cette couche des berges. Le résultat donne la surface de berges par cellule.

### 3) Colonne 25 : PERCENT\_FLOODPLAIN

Il existe une couche dans le SIGR appelée PHEC (« Plus hautes eaux connues », source : DIREN) représentant le plan de prévention des risques d'inondations (PPRI). On fait une intersection entre le maillage et cette couche. Le résultat donne la surface de zones inondables par cellule.

### 4) Colonne 26 : PERCENT\_WETLAND

Il existe une couche dans le SIGR appelée ECOZONE. Les codes 23, 27 et 36 fournissent les surfaces de zones humides.

### 5) Colonne 27 : PERCENT\_SLOPE

Cette variable est la pente. Le SIGR dispose d'un Modèle Numérique de Terrain (MNT) de la Région. On a fait un croisement entre cette couche et le maillage, ce qui a permis d'attribuer une pente moyenne à chaque cellule. Cette information permettra de dire au modèle qu'une cellule dans laquelle le terrain est trop escarpé aura peu de chances de voir son type urbain évoluer.



### 6) Colonne 28 : PERCENT\_OPEN\_SPACE

Le concept d'« open space » est typiquement américain et s'oppose à « developed space ». Les urbanistes américains sont confrontés au problème de préservation de l'open space, sachant que l'urbanisation se fait au détriment de cet open space. On a du mal à trouver un équivalent en Région Ile-de-France. On a simplement essayé d'identifier les postes du MOS pouvant correspondre à ce concept. Ce sont les postes 16 à 29 du MOS en 83 postes.

### 7) Colonne 29 : PERCENT\_PUBLIC\_SPACE

Ce sont les postes 1, 6, 7 et 14 du MOS en 21 postes. En introduisant cette variable, on suppose qu'une cellule de type espace public (bois, parcs et jardins, ...) est bien protégée et a peu de chances de se voir transformer un jour en chantier.

### 8) Colonne 30 : PERCENT\_ROADS

On a étendu la surface des emprises routières (PERCENT\_ROADS est le nom de la variable par défaut) à toutes les autres emprises de transport au sol : emprises de transport ferré, parking de surface et en étages, gares routières, dépôts de bus, installations aéroportuaires, autres. L'ensemble correspond aux postes 75 à 82 au MOS en 83 postes. Au passage, on peut noter qu'UrbanSim est tout à fait flexible puisqu'on n'est pas obligé de renseigner uniquement la surface d'emprises routières. Il suffit de renommer la variable PERCENT\_ROADS par PERCENT\_TRANSPORTS dans la table gridcells et partout où il apparaît dans le code source.

**→ La dernière étape a consisté à calculer une variable synthétique caractérisant le type urbain de chaque cellule.**

Elle concerne la dernière colonne, la **colonne 31 (DEVELOPMENT\_TYPE)**.

Par défaut, UrbanSim propose 25 types urbains différents, allant du résidentiel bas à la zone industrielle.

A chaque cellule est attribué un type urbain. Un type est fonction des variables tels que le nombre de logements, la surface au sol, l'espace non résidentiel, le % d'espace non urbanisable et la mixité. Pour attribuer le type 25 (non urbanisable), une autre formule a été écrite, en fonction des pourcentages d'eau, de zones humides, de zones inondées, de pente et d'espace public.

	A	B	C	D	E	F	G
1	DEVELOPMENT_TYPE_ID	NAME	MIN_UNITS	MAX_UNITS	MIN_SOFT	MAX_SOFT	PRIMARY_USE
2	1	R1	1	1	0	999	Residential
3	2	R2	2	4	0	999	Residential
4	3	R3	5	9	0	999	Residential
5	4	R4	10	14	0	2499	Residential
6	5	R5	15	21	0	2499	Residential
7	6	R6	22	30	0	2499	Residential
8	7	R7	31	75	0	4999	Residential
9	8	R8	76	65000	0	4999	Residential
10	9	M1	1	9	1000	4999	Mixed Use
11	10	M2	10	30	2500	4999	Mixed Use
12	11	M3	10	30	5000	24999	Mixed Use
13	12	M4	10	30	25000	49999	Mixed Use
14	13	M5	10	30	50000	9999999	Mixed Use
15	14	M6	31	65000	5000	24999	Mixed Use
16	15	M7	31	65000	25000	49999	Mixed Use
17	16	M8	31	65000	50000	9999999	Mixed Use
18	17	C1	0	9	5000	24999	Commercial
19	18	C2	0	9	25000	49999	Commercial
20	19	C3	0	9	50000	9999999	Commercial
21	20	I1	0	9	5000	24999	Industrial
22	21	I2	0	9	25000	49999	Industrial
23	22	I3	0	9	50000	9999999	Industrial
24	23	GV	0	9	5000	9999999	Government
25	24	Vacant Developable	0	0	0	0	Vacant Developable
26	25	Undevelopable	0	0	0	0	Vacant Undevelopable

Tableau 13 : Table development\_types par défaut. Source : [www.urbansim.org](http://www.urbansim.org)

Voici les formules proposées par UrbanSim, compte tenu des notations suivantes :

SC : surfaces commerciales

SA : surfaces activités secondaires

SB : surfaces de bureaux

STOT = SC + SA + SB

NL = nombre de logements

#### TYPE URBAIN=

- 1 si STOT >= 0 et STOT <= 999 et NL = 1
- 2 si STOT >= 0 et STOT <= 999 et NL >= 2 et NL <= 4
- 3 si STOT >= 0 et STOT <= 999 et NL >= 5 et NL <= 9
- 4 si STOT >= 0 et STOT <= 2499 et NL >= 10 et NL <= 14
- 6 si STOT >= 0 et STOT <= 2499 et NL >= 22 et NL <= 30
- 7 si STOT >= 0 et STOT <= 4999 et NL >= 31 et NL <= 75
- 8 si STOT >= 0 et STOT <= 4999 et NL >= 76 et NL <= 65000
- 9 si SC >= 1000 et SC <= 4999 et NL >= 1 et NL <= 9
- 10 si SC >= 2500 et SC <= 4999 et NL >= 10 et NL <= 30
- 11 si SC >= 5000 et SC <= 24999 et NL >= 10 et NL <= 30
- 12 si SC >= 25000 et SC <= 49999 et NL >= 10 et NL <= 30
- 13 si SC >= 50000 et SC <= 9999999 et NL >= 10 et NL <= 30
- 14 si SC >= 5000 et SC <= 24999 et NL >= 31 et NL <= 65000
- 15 si SC >= 25000 et SC <= 49999 et NL >= 31 et NL <= 65000
- 16 si SC >= 50000 et SC <= 9999999 et NL >= 31 et NL <= 65000
- 17 si SC >= 5000 et SC <= 24999 et NL >= 0 et NL <= 9
- 18 si SC >= 25000 et SC <= 49999 et NL >= 0 et NL <= 9

- 19 si SC  $\geq$  50000 et SC  $\leq$  9999999 et NL  $\geq$  0 et NL  $\leq$  9
- 20 si SA  $\geq$  5000 et SA  $\leq$  24999 et NL  $\geq$  0 et NL  $\leq$  9
- 21 si SA  $\geq$  25000 et SA  $\leq$  49999 et NL  $\geq$  0 et NL  $\leq$  9
- 22 si SA  $\geq$  50000 et SA  $\leq$  9999999 et NL  $\geq$  0 et NL  $\leq$  9
- 23 si SB  $\geq$  5000 et SB  $\leq$  9999999 et NL  $\geq$  0 et NL  $\leq$  9
- 24 si STOT = 0 et NL = 0
- 25 si %eau  $\geq$  50 ou %espace ouvert  $\geq$  50 ou %zones humides  $\geq$  50 ou %zones inondables  $\geq$  50 ou % espace ouvert  $\geq$  50 ou présence d'une base militaire = "Oui" ou présence d'une forêt = "Oui" ou % espace non urbanisable  $\geq$  75

Pour l'instant (septembre 2004), nous n'avons pas encore défini les types urbains dans le cas de la Région Ile-de-France. Ce travail sera fait en collaboration avec des urbanistes de l'IAURIF au cours de l'année 2 de la recherche.

### 7.2.3 Table « households »

La table « households » (comme la table « jobs ») est entièrement désagrégée. C'est-à-dire qu'elle comporte autant d'enregistrements qu'il y a de ménages (d'emplois) en Ile-de-France. Aux Etats-Unis, il n'y a pas de recensement général de la population exhaustif mais des recensements partiels portant en général sur 10% de la population (PUMS : Public Use Microdata Sample). C'est pourquoi l'équipe du Pr Waddell utilise une méthode de synthèse d'un fichier de ménages sur son aire d'étude à partir de recensements partiels par la méthode de Monte-Carlo.

En France, on dispose des RGP de 1990 et 1999 qui sont exhaustifs. On n'a donc pas à générer des fichiers de ménages virtuels.

La création de la table « households » (ménages) est donc effectuée en grande partie à l'aide des données présentes dans le Recensement Général de la Population (RGP). Les informations utilisées au sein du RGP pour créer la table sont :

- Le nombre d'individus dans le ménage
- Le nombre d'actifs dans le ménage
- Le nombre de voitures dans le ménage
- Le code îlot

Cependant, le RGP ne contient pas l'information de revenu des ménages (entité fiscale). Pour l'année 1999 (RGP 99), les informations relatives au revenu vont être déduites de la base de données de la Direction Générale des Impôts (DGI) qui donne des informations sur la distribution des revenus au niveau communal (1300 communes). Pour certaines communes, nous disposons des 9 déciles de la distribution des revenus, pour d'autres nous disposons uniquement de la médiane et pour certaines nous n'avons aucune information (voir Tableau 14).

Informations Disponibles	Nombre de communes
Tous les déciles	529
Médiane	726
Aucune information	45

Tableau 14 : disponibilité d'information pour les communes

Cependant, étant donnée la distribution des ménages du RGP au sein des communes, le manque d'information se propage à un faible nombre de ménages puisque les communes pour lesquelles on dispose de tous les déciles sont de grandes communes contenant beaucoup de ménages (Tableau 15).

Informations Disponibles	Nombre de ménages
Tous les déciles	183 341
Médiane	4 325 537
Aucune information	1 491

Tableau 15 : disponibilité d'information pour les ménages

La génération d'un revenu, pour un ménage dans une zone donnée se fait maintenant de la façon suivante :

Pour les ménages pour lesquels on dispose de l'ensemble de la distribution du revenu (D1 à D9), on fait un tirage aléatoire pour affecter chaque ménage dans une tranche de revenu :

- $T1=[0,D1[ \rightarrow p1=1/10$
- $T2=[D1,D2[ \rightarrow p2=1/10$
- ...
- jusqu'à  $T10=[D9,\infty[ \rightarrow p10=1/10$

Puis, au sein de chaque tranche de revenu, on fait un tirage uniforme pour avoir non seulement, pour chaque ménage, une tranche de revenu mais également un niveau de revenu.

Pour les ménages pour lesquels on ne dispose que de la médiane, on calcule les moyennes des quartiles Q1 et Q3 dans le département et on procède comme précédemment, avec quatre tranches de revenu.

Pour les ménages pour lesquels on ne dispose d'aucune information, on calcule les moyennes des quartiles Q1, Q2 et Q3 dans le département et on procède comme précédemment, avec quatre tranches de revenu.

Il reste à localiser chaque ménage sur notre maillage en 49271 cellules.

C'est l'attribut code îlot qui est utilisé pour positionner chaque ménage dans une cellule du maillage.

Le fichier contient 4,5 millions de ménages (4 510 369 exactement) en Ile-de-France en 1999.

## 7.2.4 Table « jobs »

Les données disponibles au niveau des emplois sont beaucoup moins précises que celles des ménages. Nous disposons d'une synthèse des données du RGP qui nous indique le nombre

d'emplois dans chaque commune en quatre secteurs : agriculture, industrie, construction et tertiaire. Cette base est assez fiable mais est peut-être insuffisante au niveau de la sectorisation des activités car 75% des emplois en Ile-de-France sont en secteur tertiaire. Il est difficile d'obtenir auprès de l'INSEE un éclatement en plus de quatre secteurs.

Une autre source disponible est la base de données issue de l'Enquête Régionale Emploi (ERE). Le fichier détaillé ERE donne la liste des établissements en 15 secteurs (nomenclature NAP 15) présents en Ile-de-France, en indiquant l'activité principale de cet établissement, son adresse et le nombre de salariés.

Nous rencontrons deux difficultés majeures dans l'utilisation de cette dernière. D'abord le fait qu'elle ne concerne que les emplois salariés des entreprises ayant un effectif supérieur à 100 employés. Ensuite, cette enquête étant basée sur la déclaration des entreprises subit un biais dû à la déclaration de tous les emplois aux sièges sociaux des grandes entreprises. Par contre, l'IAURIF dispose de l'Enquête régionale sur l'emploi 2000 géolocalisée, acquise auprès de l'APUR qui a réalisé un important travail de relocalisation de certains effectifs parisiens, par recoupement avec d'autres fichiers. Les données n'ont été géolocalisées que pour les **emplois salariés du secteur privé**. Par contre, la base de données alpha-numérique comporte bien tous les emplois salariés, publics et privés mais ne comptabilise pas les emplois de type libéral.

Dans ce contexte, nous avons préféré utiliser les données issues du RGP. Les emplois sur les cellules de chaque commune ont été répartis en fonction de leur surface d'activité. On a aussi distingué les actifs qui travaillent à domicile.

Le principe de répartition est le suivant : on affecte par cellule les emplois tertiaires sur les postes du MOS de type commerces et bureaux, et les autres emplois (agriculture, industrie, construction) sur les postes activités secondaires du MOS. On a ajouté à la fin les emplois à domicile.

Au final, la table « jobs » contient 5 075 974 enregistrements en 1990 et 5 041 995 en 1999. Chaque enregistrement correspond à un emploi, renseigné par son secteur d'activité et sa localisation dans le maillage.

## 7.2.5 Autres tables

Les tables suivantes sont décrites de manière détaillée dans la documentation en ligne du site [urbansim.org](http://urbansim.org).

### 1) Les tables de base

#### - Table « **BASE\_YEAR** »

Il comporte 3 lignes :

YEAR
INTEGER
1990

#### - Table « **CITIES** »

C'est la table de correspondance entre code INSEE et nom des communes. Il comporte 1301 lignes.

CITY_ID ;	CITY_NAME
INTEGER ;	VARCHAR
75101	Paris 1 <sup>ème</sup>
75102	Paris 2 <sup>ème</sup>
75103	Paris 3 <sup>ème</sup>
...	

- **Table « COUNTIES »**

C'est la table de correspondance entre numéro de département et nom du département. Il comporte 9 lignes.

COUNTY_ID ;	COUNTY_NAME
INTEGER ;	VARCHAR
75	Paris
77	Seine-et-Marne
78	Yvelines
91	Essonne
92	Hauts-de-Seine
93	Seine-Saint-Denis
94	Val-de-Marne
95	Val-d'Oise

- **Table « MODELS »**

Il liste les différents modules d'UrbanSim.

RUN_ORDER ;	JAVA_CLASS
INTEGER ;	VARCHAR
1	org.urbansim.models.AccessibilityModel
2	org.urbansim.models.HouseholdTransitionModel
3	org.urbansim.models.EmploymentTransitionModel
4	org.urbansim.models.HouseholdRelocationChoiceModel
5	org.urbansim.models.EmploymentRelocationChoiceModel
6	org.urbansim.models.HouseholdLocationChoiceModel
7	org.urbansim.models.EmploymentLocationChoiceModel
8	org.urbansim.models.LandPriceModel
9	org.urbansim.models.DeveloperModel
10	org.urbansim.models.DeveloperModel_VacancyAdjustment

- **Table « ZONES »**

Il donne les temps de parcours depuis chacune des 606 zones du modèle de trafic vers des zones à définir par l'utilisateur : les aéroports, les gares parisiennes, les pôles d'emplois tels que la Défense, les villes nouvelles, ... Ces temps sont à faire calculer soit par le modèle de trafic METROPOLIS soit par Davisum.

- **Table « TRAVEL\_DATA »**

Cette table donne les logsommes de zone à zone (dans le découpage du modèle de trafic) par taux de motorisation (0,1, 2 ou 3 et plus). On envisage de considérer seulement deux taux de motorisation : 0 ou 1 et plus.

Exemple :

FROM_ZONE_ID ; INTEGER	TO_ZONE_ID ; INTEGER	LOGSUM0 ; FLOAT	LOGSUM1 FLOAT
1	1	-0.1	-0.2
1	2	-0.11	-0.21
1	3	-0.12	-0.22
...			

Les logsommes sont calculées par METROPOLIS.

## 2) Les chiffres de cadrages globaux population/emplois : données exogènes

Il sont obtenus d'une manière ou d'une autre, par exemple par un modèle macro-économique externe.

- **Table « ANNUAL\_EMPLOYMENT\_CONTROL\_TOTALS » :**  
Nombre d'emplois total sur la Région par secteur pour chaque année de 1990 à 1999
- **Table « ANNUAL\_HOUSEHOLDS\_CONTROL\_TOTALS » :**  
Nombre de ménages total sur la Région pour chaque année de 1990 à 1999.
- **Table « ANNUAL\_RELOCATION\_RATES\_FOR\_JOBS » :**  
Part des emplois totaux sur la Région par secteur qui sont relocalisés par année de 1990 à 1999 (UrbanSim fait l'hypothèse que cette part est constante).
- **Table « ANNUAL\_RELOCATION\_RATES\_FOR\_HOUSEHOLDS » :**  
Part des ménages totaux sur la Région par secteur qui déménagent par année de 1990 à 1999 (UrbanSim fait l'hypothèse que cette part est constante).

## 3) Les tables concernant les ménages

- **Table « ANNUAL\_RELOCATION\_RATES\_FOR\_HOUSEHOLDS » :**  
Taux de déménagement par type de ménage. A priori, il semble que l'utilisateur peut définir sa propre segmentation des ménages, ce qui est pratique car on segmente en fonction des données disponibles ! (par exemple, par tranche de revenu, par taille du ménage et par âge du chef de ménage). Par contre, c'est le même taux pour toutes les années. Il faudra voir s'il est possible d'entrer des paramètres par année (ou au moins par période).
- **Table « HOUSEHOLD\_CHARACTERISTICS\_FOR\_HLC » :**  
Cette table définit les caractéristiques d'un ménage à prendre en compte dans les modules de choix de localisation et de relocalisation résidentielle. Par exemple, la donnée de revenu des ménages est donnée par tranche. On a repris les tranches définies dans l'Enquête Globale Transports qui étaient en Francs et ramenés en Euros .

CHARACTERISTIC ; VARCHAR	MIN ; INTEGER	MAX INTEGER
Income	0	9146
Income	9147	18293
Income	18294	27440
Income	27441	36587
Income	36588	54881
Income	54882	91463

Income	91464	-1
Persons	1	-1
Workers	0	-1
...		

(-1 signifie qu'il n'y a pas de maximum)

- **Table « HOUSEHOLD\_CHARACTERISTICS\_FOR\_HT » :**

Cette table définit les caractéristiques d'un ménage à prendre en compte dans le module de transition démographique d'UrbanSim.

CHARACTERISTIC	MIN	MAX
VARCHAR	INTEGER	INTEGER
Income	0	9146
Income	9147	18293
Income	18294	27440
Income	27441	36587
Income	36588	54881
Income	54882	91463
Income	91464	-1
Children	0	-1
Age of Head	0	18
Age of Head	18	25
Age of Head	25	35
Age of Head	35	45
Age of Head	35	55
Age of Head	55	65
Age of Head	65	75
Age of Head	75	-1
Workers	0	-1
Cars	0	-1
...		

(-1 signifie qu'il n'y a pas de maximum)

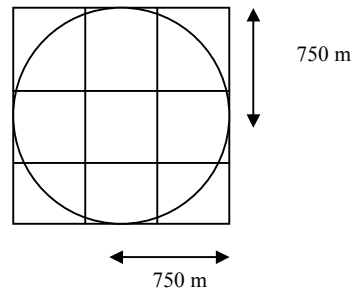
Il est intéressant d'un point de vue culturel de noter qu'il existe une table appelée « RACE\_NAMES ». Cette table décrit l'« origine ethnique » (non white = black, asian or hispanic americans, white americans) d'un ménage. Elle existe dans les statistiques américaines et utilisées dans les modèles aux Etats-Unis alors que son utilisation en France est strictement interdite. Il est évident que cette table ne sera pas utilisée dans notre projet. On peut penser que la variable qui pourrait la remplacer dans notre modèle est le revenu des ménages, même si l'on sait pertinemment qu'elle ne recouvre pas les mêmes notions.

- **Table « HOUSEHOLD\_LOCATION\_CHOICE\_MODEL\_SPECIFICATION » :**

Cette table comporte 80 lignes, chacune définissant des variables d'accessibilité locale qui seront calculées automatiquement par UrbanSim à partir du quadrillage en mailles carrées, par exemple, la surface commerciale accessible à pied depuis une cellule. On considère que la distance de marche est au maximum de 750 mètres depuis le centroïde d'une cellule.



UrbanSim va donc calculer pour chaque cellule la surface commerciale totale comprise dans les huit carrés voisins (avec une approximation sur la diagonale des quatre carrés aux coins assimilée à 750 mètres).<sup>5</sup> On comprend maintenant mieux la raison du maillage carré !



Ce calcul automatique par UrbanSim permet de caractériser chaque cellule par une sorte de valeur d'accessibilité locale. Plus cette valeur est élevée, plus elle permet d'atteindre des opportunités dans son voisinage.

SUB_MODEL_ID	EQUATION_ID	VARIABLE_NAME	COEFFICIENT_NAME
INTEGER	INTEGER	VARCHAR	VARCHAR
-2	-2	ln_commercial_sqft_within_walking_distance	In_commercial_sqft_within_walking_distance
-2	-2	ln_home_access_to_employment_1_car	In_home_access_to_employment_1_car
...			

#### 4) Les tables concernant les emplois

La structure des données d'emplois permet de définir des secteurs d'emplois et à l'intérieur des groupes d'emplois. Par exemple, dans le secteur tertiaire, les emplois de type banque et les emplois de type administratifs. Mais à l'heure actuelle nous ne disposons pas de données plus fines que les 4 secteurs d'emplois. Pour cette raison, les différentes tables concernant les emplois sont assez simples.

##### - Table « EMPLOYMENT\_ADHOC\_SECTOR\_GROUP\_DEFINITIONS » :

SECTOR_ID	GROUP_ID
INTEGER	INTEGER
1	1
2	2
3	3
4	4

<sup>5</sup> Un calcul simple de surfaces montre que la surface des 9 carrés est supérieure de  $4/\pi - 1$  soit 28% par rapport au disque de rayon 750 m. On peut imaginer un maillage en hexagones de 250 m de côté. Topologiquement parlant, un maillage en carrés et un maillage en hexagones sont équivalents, sauf que le second est une meilleure approximation de disques. Un calcul de surfaces montre que l'hexagone central plus les six qui l'entourent ont une surface inférieure de  $(1 - \frac{3\sqrt{3}}{2\pi})$  soit 17% seulement par rapport au disque qui circonscrit les 7 hexagones.

- **Table « EMPLOYMENT\_ADHOC\_SECTOR\_GROUP » :**

GROUP_ID	NAME
1	agriculture
2	industrie
3	construction
4	tertiaire

- **Table « EMPLOYMENT\_SECTORS » :**

SECTOR_ID	GROUP_ID
1	1
2	2
3	3
4	4

- **Table « EMPLOYMENT\_ADHOC\_SECTOR\_GROUP » :**

GROUP_ID	NAME
1	agriculture
2	industrie
3	construction
4	tertiaire

- **Table « EMPLOYMENT\_EVENTS » :**

Cette table est très utile. Elle permet de définir des événements importants qui de toutes façons ne sont pas prévisibles. Par exemple, l'implantation en 1998 du nouveau siège social de Renault à Saint-Quentin-les-Yvelines (Guyancourt) qui regroupe en 2004 près de 9000 emplois, l'implantation d'une nouvelle université dans le XIIIème arrondissement, ...

Ces événements sont caractérisés par le nombre d'emplois créés par secteur, l'année de déroulement et la localisation dans le maillage.

Par exemple (données fictives) :

GRID_ID	JOBS	SCHEDULED_YEAR	SECTOR_ID
515555	2000	1994	4
515556	1000	1995	4
675855	500	1997	4
...			

Evidemment, cette table est assez fastidieuse à constituer. Mais il peut être utile pour des calages locaux et du moins sur l'aire d'étude de la Tangentielle nord.

- **Table**

« **EMPLOYMENT\_LOCATION\_CHOICE\_MODEL\_SPECIFICATION** » :

Cette table comporte une centaine de lignes, chacune définissant des variables d'accessibilité locale qui seront calculées automatiquement par UrbanSim à partir du quadrillage en mailles carrées.

SUB_MODEL_ID	EQUATION_ID	VARIABLE_NAME	COEFFICIENT_NAME
INTEGER	INTEGER	VARCHAR	VARCHAR
3	-2	ln_work_access_to_population_1	ln_work_access_to_population_1
3	-2	ln_total_job_locations	ln_total_job_locations
...			

Les valeurs de ces variables sont calculées par UrbanSim qui utilise pour cela toutes les tables de données.

Les coefficients pondérant ces variables sont estimés par un logiciel externe d'économétrie puis sont entrés dans la table « **EMPLOYMENT\_LOCATION\_CHOICE\_MODEL\_COEFFICIENT** » .

- **Table « SQFT\_FOR\_NON\_HOME\_BASED\_JOBS » :**

Cette table donne la surface occupée par un emploi par type urbain.

Par exemple :

DEVELOPMENT_TYPE_ID	SQFT
INTEGER	FLOAT
1	12
2	18
...	

## 5) Les tables concernant le développement urbain

Le thème du développement urbain n'a pas encore été étudié au cours de cette première année de recherche. Apparemment, UrbanSim permet de prendre en compte des contraintes urbaines futures, des orientations politiques, des documents de planification locale. Ceci est très utile car depuis les lois de décentralisation en 1983, les communes se sont vues octroyé le pouvoir d'urbanisme local. Les documents de type POS/PLU sont donc importants et il est possible, dans une certaine mesure seulement, de prendre en compte ces documents dans le modèle. En effet, l'IAURIF peut avoir accès, dans le cadre d'un partenariat avec la DREIF, à la BD-POS qui est la base de données géolocalisées des 1300 Plans Locaux d'Urbanisme (PLU) de la Région. Cette base est centralisée et numérisée par la DREIF. L'IAURIF dispose de cette base de données dans un mode dégradé, c'est-à-dire que seuls les types suivants sont fournis : U, NA1, NA2, NB, NC, ND, ZAC, infrastructures (voir glossaire). Une demande de fourniture de la BD-POS sur l'ensemble de la Région ou au moins sur le territoire de la Tangentielle nord a été faite auprès de la DREIF/DUSD. L'IAURIF devrait réceptionner cette couche en novembre 2004. Il suffira ensuite de faire un croisement entre cette couche et le maillage.

La structure de données sur le développement urbain est constituée 13 tables. On décrit ci-après seulement les colonnes de chacune de ces tables et leur type alpha-numérique. Les approfondissements seront faits au cours de la deuxième année.

- **Table « DEVELOPMENT\_TYPE\_GROUP\_DEFINITIONS » :**

DEVELOPMENT_TYPE_ID ; GROUP_ID INTEGER ; INTEGER ...
--

- **Table « DEVELOPMENT\_TYPE\_GROUPS » :**

GROUP_ID ; NAME ; NON_OVERLAPPING_GROUPS INTEGER ; VARCHAR ; VARCHAR ...
--

- **Table « DEVELOPMENT\_TYPES » :**

DEVELOPMENT_TYPE_ID ; NAME ; MIN_UNITS ; MAX_UNITS ; MIN_SQFT ; MAX_SQFT ; PRIMARY_USE INTEGER ; VARCHAR ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; VARCHAR ...
--

- **Table « DEVELOPMENT\_CONSTRAINT » :**

CONSTRAINT_ID ; CITY_ID ; COUNTY_ID ; IS_IN_WETLAND ; IS_OUTSIDE_URBAN_GROWTH_BOUNDARY ; IS_IN_STREAM_BUFFER ; IS_ON_STEEP_SLOPE ; IS_IN_FLOODPLAIN ; PLANTYPE_X ; DEVTYPE_X  INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; VARCHAR ; VARCHAR ...
---

- **Table « DEVELOPMENT\_EVENTS » :**

GRID_ID ; SCHEDULED_YEAR ; DEVELOPMENT_TYPE_CHANGE_TYPE ; DEVELOPMENT_TYPE_ID RESIDENTIAL_UNITS_CHANGE_TYPE ; RESIDENTIAL_UNITS ; COMMERCIAL_SQFT_CHANGE_TYPE ; COMMERCIAL_SQFT ; INDUSTRIAL_SQFT_CHANGE_TYPE ; INDUSTRIAL_SQFT ; GOVERNMENTAL_SQFT_CHANGE_TYPE ; GOVERNMENTAL_SQFT ; RESIDENTIAL_IMPROVEMENT_VALUE_CHANGE_TYPE ; RESIDENTIAL_IMPROVEMENT_VALUE ; COMMERCIAL_IMPROVEMENT_VALUE_CHANGE_TYPE ; COMMERCIAL_IMPROVEMENT_VALUE ; INDUSTRIAL_IMPROVEMENT_VALUE_CHANGE_TYPE ; INDUSTRIAL_IMPROVEMENT_VALUE ; GOVERNMENTAL_IMPROVEMENT_VALUE_CHANGE_TYPE ; GOVERNMENTAL_IMPROVEMENT_VALUE ; FRACTION_RESIDENTIAL_LAND_VALUE_CHANGE_TYPE ; FRACTION_RESIDENTIAL_LAND_VALUE ;  INTEGER ; SHORT ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; FLOAT ...
---

- **Table « DEVELOPMENT\_CONSTRAINT\_EVENT » :**

SCHEDULED_YEAR ; CHANGE_TYPE ; CONSTRAINT_ID ; CITY_ID ; COUNTY_ID ; IS_IN_WETLAND ; IS_OUTSIDE_URBAN_GROWTH_BOUNDARY ; IS_IN_STREAM_BUFFER ; IS_ON_STEEP_SLOPE ; IS_IN_FLOODPLAIN ; PLANTYPE_X ; DEVTYPE_X  SHORT ; VARCHAR ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; INTEGER ; VARCHAR ; VARCHAR
---

- **Table « DEVELOPMENT\_HISTORY » :**

GRID_ID ;	STARTING_DEVELOPMENT_TYPE_ID ;	ENDING_DEVELOPMENT_TYPE_ID ;	YEAR
INTEGER ;	INTEGER	INTEGER ;	SHORT

- **Table « DEVELOPER\_MODEL\_SPECIFICATIONS » :**

SUB_MODEL_ID ;	EQUATION_ID ;	VARIABLE_NAME ;	COEFFICIENT_NAME
INTEGER ;	INTEGER ;	VARCHAR ;	VARCHAR
...			

- **Table « PLAN\_TYPES » :**

Cette table correspond à la description de la couche SDRIF 94.

PLAN_TYPE_ID ;	NAME
INTEGER ;	VARCHAR
0	non_renseigné
1	partiellement_urbanisable_SDRIF94
2	urbanisable_SDAU76
3	non_prioritaire_pour_Etat

- **Table « TRANSITION\_TYPES » :**

TRANSITION_ID ;	INCLUDE_IN_DEVELOPER_MODEL	STARTING_DEVELOPMENT_TYPE_ID ;	ENDING_DEVELOPMENT_TYPE_ID ;	HOUSING_UNITS_MEAN ;	HOUSING_UNITS_STANDARD_DEVIATION ;	HOUSING_UNITS_MIN ;	HOUSING_UNITS_MAX ;	COMMERCIAL_SQFT_MEAN ;	COMMERCIAL_SQFT_STANDARD_DEVIATION ;	COMMERCIAL_SQFT_MIN ;	COMMERCIAL_SQFT_MAX ;	INDUSTRIAL_SQFT_MEAN ;	INDUSTRIAL_SQFT_STANDARD_DEVIATION	INDUSTRIAL_SQFT_MIN ;	INDUSTRIAL_SQFT_MAX ;	GOVERNMENTAL_SQFT_MEAN ;	GOVERNMENTAL_SQFT_STANDARD_DEVIATION ;	GOVERNMENTAL_SQFT_MIN ;	GOVERNMENTAL_SQFT_MAX	HOUSING_IMPROVEMENT_VALUE_MEAN ;	HOUSING_IMPROVEMENT_VALUE_STANDARD_DEVIATION ;	HOUSING_IMPROVEMENT_VALUE_MIN ;	HOUSING_IMPROVEMENT_VALUE_MAX ;	COMMERCIAL_IMPROVEMENT_VALUE_MEAN ;	COMMERCIAL_IMPROVEMENT_VALUE_STANDARD_DEVIATION ;	COMMERCIAL_IMPROVEMENT_VALUE_MIN ;	COMMERCIAL_IMPROVEMENT_VALUE_MAX ;	INDUSTRIAL_IMPROVEMENT_VALUE_MEAN ;	INDUSTRIAL_IMPROVEMENT_VALUE_STANDARD_DEVIATION ;	INDUSTRIAL_IMPROVEMENT_VALUE_MIN ;	INDUSTRIAL_IMPROVEMENT_VALUE_MAX ;	GOVERNMENTAL_IMPROVEMENT_VALUE_MEAN ;	GOVERNMENTAL_IMPROVEMENT_VALUE_STANDARD_DEVIATION ;	GOVERNMENTAL_IMPROVEMENT_VALUE_MIN ;	GOVERNMENTAL_IMPROVEMENT_VALUE_MAX ;	YEARS_TO_BUILD ;		
INTEGER ;	BOOLEAN ;	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT ;	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT ;	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT ;	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT ;	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT ;	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT ;	INTEGER ;	INTEGER ;	FLOAT ;	FLOAT ;	INTEGER ;	INTEGER ;	INTEGER ;	INTEGER ;	INTEGER ;

- **Table « RESIDENTIAL\_LAND\_SHARE\_MODEL\_SPECIFIC\_VARIABLES » :**

Exemple :

VARIABLE_NAME
VARCHAR
devtype_9
non_residential_sqft
devtype_8
devtype_7
devtype_6
devtype_5
residential_units
devtype_4
devtype_3
devtype_2
devtype_1

- **Table « RESIDENTIAL\_LAND\_SHARE\_MODEL\_SPECIFICATION » :**

SUB_MODEL_ID ; EQUATION_ID ; VARIABLE_NAME ; COEFFICIENT_NAME
INTEGER ; INTEGER ; VARCHAR ; VARCHAR
...

- **Table « RESIDENTIAL\_UNITS\_FOR\_HOME\_BASED\_JOBS » :**

Cette table donne la correspondance entre nombre d'emplois à domicile (libéraux, indépendants, ...) et nombre de logements par type urbain.

Par exemple :

DEVELOPMENT_TYPE_ID ; RATIO
INTEGER ; FLOAT
1 0,3
...

Cet exemple signifie qu'un logement situé une cellule de type 1 peut accueillir  $1/0,3=3$  emplois.

6) Les tables concernant les prix de l'immobilier

- **Table « LAND\_PRICE\_MODEL\_SPECIFIC\_VARIABLES » :**

Cette table décrit les noms d'une cinquantaine de variables sur une colonne :

VARIABLE_NAME
VARCHAR
is_in_stream_buffer
ln_average_total_value_per_residential_unit_within_walking_distance
ln_residential_units
ln_total_employment_within_walking_distance
percent_slope
...

- **Table « LAND\_PRICE\_MODEL\_SPECIFICATION » :**

SUB_MODEL_ID ; EQUATION_ID ; VARIABLE_NAME ;	COEFFICIENT_NAME
INTEGER ; INTEGER ; VARCHAR ;	VARCHAR
...	

- **Table « LAND\_USE\_EVENTS » :**

GRID_ID ; SCHEDULED_YEAR ; PLAN_TYPE_CHANGE_TYPE ; PLAN_TYPE_ID ;
PERCENT_WATER_CHANGE_TYPE ; PERCENT_WATER ;
PERCENT_FLOODPLAIN_CHANGE_TYPE ; PERCENT_FLOODPLAIN ;
PERCENT_WETLAND_CHANGE_TYPE ; PERCENT_WETLAND ; PERCENT_SLOPE_CHANGE_TYPE ;
PERCENT_SLOPE ; PERCENT_OPEN_SPACE_CHANGE_TYPE ; PERCENT_OPEN_SPACE ;
PERCENT_PUBLIC_SPACE_CHANGE_TYPE ; PERCENT_PUBLIC_SPACE ;
PERCENT_ROADS_CHANGE_TYPE ; PERCENT_ROADS ;
PERCENT_STREAM_BUFFER_CHANGE_TYPE ; PERCENT_STREAM_BUFFER ;
IS_OUTSIDE_UGB_CHANGE_TYPE ; IS_OUTSIDE_UGB
INTEGER ; SHORT ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ;
VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ;
INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; INTEGER ; VARCHAR ; BOOLEAN
...

## 7.2.6 Table des constantes d'UrbanSim

- **Table « URBANSIM\_CONSTANTS » :**

Il s'agit d'une table qui regroupe toutes les constantes utilisées dans les calculs automatiques d'UrbanSim. Par exemple : WALKING\_DISTANCE\_CIRCLE\_RADIUS = 750 m.

Cette table sera étudiée au cours de la deuxième année.

HLC_ESTIMATION_SAMPLE_FRACTION ; LOGIT_CHOICE_THRESHOLD ;
LOGIT_CHOICE_SET_SIZE_FOR_ESTIMATIONCELL_SIZE ; UNITS ;
WALKING_DISTANCE_CIRCLE_RADIUS ; YOUNG_AGE ;
PROPERTY_VALUE_TO_ANNUAL_COST_RATIO ; LOW_INCOME_FRACTION ;
MID_INCOME_FRACTION ; NEAR_ARTERIAL_THRESHOLD ; NEAR_HIGHWAY_THRESHOLD ;
PERCENT_COVERAGE_THRESHOLD ; RECENT_YEARS ;
MAX_PERSONS_PER_HOUSEHOLD_FOR_CONTROL_TOTALS ;
DOUBLE ; INTEGER ; INTEGER ; DOUBLE ; VARCHAR ; FLOAT ; INTEGER ; FLOAT ; FLOAT ;
FLOAT ; FLOAT ; FLOAT ; INTEGER ; INTEGER ; INTEGER
...

- **Table « SAMPLING\_RATES » :**

Cette table donne la fraction de l'ensemble des alternatives (en l'occurrence 49 271 cellules, moins les cellules figées car contraintes ou ayant des « events » pré-définis) pour les

localisations possibles des emplois et des ménages. Le tirage de l'ensemble de choix se fait aléatoirement si cette fraction est inférieure à 100%. Cette méthode de tirage aléatoire apparaît un peu simpliste. Le Pr Waddell pense qu'il est possible de modifier la méthode pour mieux prendre en compte la réalité. Par exemple, des études ont montré que lorsqu'un ménage résidant en Ile-de-France déménage, la probabilité qu'il reste dans le même département est très forte.

L'avantage de prendre une fraction inférieure à 100% est que cela diminuera les temps de calculs.

ELC_SAMPLING_RATE ; HLC_SAMPLING_RATE
FLOAT ; FLOAT
1.0 ; 1.0

(1.0 signifie que les utilités sont calculées pour 100% des cellules ayant la possibilité d'accueillir un emploi ou un ménage).

### ***7.3 Base de données de calage***

Pour caler les différents modules d'UrbanSim, il est nécessaire de construire des bases de données de calage, c'est-à-dire d'observations réelles. Les sources sont des enquêtes ménages, le Recensement et le Mode d'occupation du Sol de l'IAURIF.

#### **7.3.1 Enquête Globale Transport (EGT)**

L'Enquête Globale de Transport (EGT) est une enquête de grande ampleur sur les déplacements des Franciliens. Depuis vingt-cinq ans, elle permet de suivre et d'interpréter les évolutions des pratiques des habitants de la région en matière de déplacements. Elle constitue une source précieuse d'informations pour élaborer les politiques régionales d'aménagement et pour prévoir les services et les infrastructures de transport.

L'objectif principal des EGT est d'arriver à mieux connaître la mobilité des Franciliens, c'est à dire leur manière de se déplacer, un jour moyen de semaine.

La dernière enquête date de 2001 et les précédentes éditions se sont déroulées en 1976, 1983, 1991 et 1997, cette dernière étant considérée comme une enquête intermédiaire.

Les EGT cherchent à mieux connaître les déplacements des individus, leurs motifs, leurs répartitions spatiale et temporelle, leurs modes, ... Les informations qu'elles apportent sont essentielles pour définir et prévoir les services et les infrastructures de transport, ainsi que pour planifier l'urbanisme de la région Ile-de-France.

Elles sont réalisées en partenariat entre l'Etat, la Région Ile-de-France, la Ville de Paris, le Syndicat des Transports d'Ile-de-France, la RATP, la SNCF et Cofiroute.

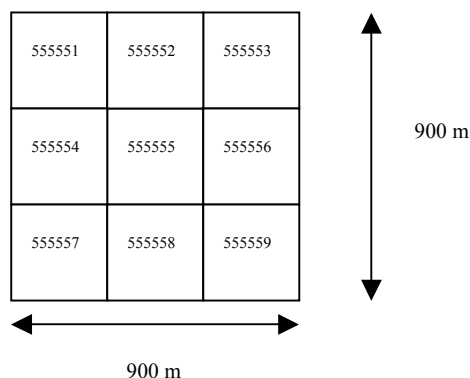
La dernière enquête a été menée entre octobre 2001 et avril 2002. Comme lors des précédentes enquêtes de 1976, 1983 et 1991, les ménages enquêtés sont issus d'un échantillon du recensement général de la population. Le principe de recueil est celui des enquêtes de déplacements menées classiquement dans les agglomérations françaises. Les ménages sont enquêtés en face à face. Chacun des membres du ménage de six ans ou plus est interrogé sur tous les déplacements qu'il a effectués la veille du jour de l'enquête. Pour chacun de ces



déplacements, on recueille un nombre important de caractéristiques : motif du déplacement, mode(s) de transport utilisé(s), origine, destination, heure de départ et d'arrivée, etc. Cette richesse, alliée au nombre important de ménages enquêtés (environ 10 500 ménages ont répondu à l'enquête 2001-2002) fait de cette enquête une source unique d'informations sur les déplacements quotidiens des Franciliens. On dispose en particulier de la donnée de revenu des ménages par tranche.

#### Localisation des ménages de l'EGT dans le maillage

On trouve dans l'EGT une localisation des ménages dans un quadrillage propre appelé carroyage. Il est constitué de mailles carrées de 300 mètres de côté, chaque maille est identifiée par un code en 6 chiffres. Le carreau de référence est centré sur la cathédrale Notre-Dame-de-Paris et est codé 555555. A partir de là, tous les autres carreaux se déduisent par une numérotation astucieuse :



Sur l'idée des fractales, un carré de 300 m de côté est identifié par un nombre à 6 chiffres (par exemple 555555), le carré de 900 m est lui-même identifié par le code en 5 chiffres 55555 et le carré de 2700 m par le code en 4 chiffres (5555).

La couche du carroyage EGT a donc été reconstruite dans le SIGR sur ce principe.

Ensuite il a suffi de faire un croisement entre le carroyage et notre maillage en 49 271 zones, ce qui a permis d'obtenir la table de correspondance entre les deux quadrillages. Puis on a créé une colonne supplémentaire dans le fichier EGT des ménages. Cette colonne est la localisation dans notre maillage du lieu de résidence du ménage.

Le fichier ménages de l'EGT 91-92 comporte au final 11 152 enregistrements (on a sorti préalablement du fichier ménages les résidents de l'Oise, numéro de département 60).

### **7.3.2 Enquête Nationale Logement (ENL)**

Cette enquête est réalisée par l'INSEE tous les 4-5 ans. Elle porte sur environ 5500 ménages. Elle devrait permettre de caler le modèle de localisation des ménages. Le modèle permettra de retracer les parcours résidentiels des ménages franciliens et constitue en quelque sorte en lui-même, c'est-à-dire indépendamment des autres modules d'UrbanSim, un outil de micro-simulation qui peut intéresser les démographes. En effet, selon eux, les techniques de micro-simulation sont en plein développement actuellement car elles répondent à une forte demande sociale (des prévisions administratives). Elles permettent de complexifier les catégories de projection. L'utilisation de projections multidimensionnelles (classiques) suppose le calcul de

probabilités de passage à un niveau de désagrégation qui remet en question leur fiabilité statistique et entraînent une forte rigidité des mécanismes. De plus, l'un des intérêts les plus fondamentaux de l'outil de micro-simulation - le suivi des trajectoires individuelles - est d'échapper à l'effet "d'écrasement" des phénomènes par la moyenne statistique et d'en approcher les composantes différentielles plus fines.

Par exemple, les mobilités résidentielles des ménages présents dans l'Enquête Nationale Logement de 1992 pourraient être micro-simulées jusqu'à un horizon lointain (2025) afin d'évaluer les taux d'effort de financement consacré au logement. Chaque ménage est caractérisé par sa situation dans une typologie spécifique, combinant variables d'habitat et niveau de revenu, qui permettait de définir son niveau d'effort financier à chaque période. Les probabilités de passage d'un état à l'autre sont issues des données de l'enquête et nourrissent un système de matrices de tirages aléatoires. La seule évolution démographique prise en compte pourrait être l'avancée en âge et la mortalité de la personne de référence des ménages.

Dans notre projet, on envisage d'utiliser une typologie au sein de laquelle chaque ménage évolue au fil du cycle de vie : d'une part, en prenant en compte la composition des ménages, et d'autre part, en enrichissant les catégories d'habitat considérées.

Les facteurs possibles concernent les « événements de la vie personnelle » et les « événements résidentiels ». Dans quelle mesure l'emménagement d'un ménage dans un nouveau logement, son choix, même sous contrainte, d'un nouveau domicile, correspondent-ils à une transformation - active ou subie - du groupe domestique ? De quelles permanences du groupe des co-résidents le maintien au même domicile est-t-il synonyme ? La « simultanéité » d'événements relatifs à la vie personnelle et à la situation résidentielle des membres du ménage, traduit l'adaptation du ménage à son logement.

Voilà tout ce que pourrait nous apporter l'ENL. Le seul inconvénient, et non des moindres, est qu'il n'y a pas de localisation plus précise que le département ! Ceci apparaît rédhibitoire dans notre tentative de calage du modèle dans notre maillage en 50 000 cellules. L'ENL peut cependant s'avérer utile pour caler des paramètres globaux sur la mobilité résidentielle et reconstituer des trajectoires résidentielles du point de vue temporel mais non spatial.

Les ENL sont disponibles aux années 1992, 1996 et 2002.

### **7.3.3 Base de données d'évolution du Mode d'Occupation du Sol (MOS)**

Les réflexions ont commencé. Cette base sera constituée au cours de la deuxième année. Elle servira au calage du module de développement urbain.

### **7.3.4 Recensement général de la population (RGP)**

Le RGP a déjà servi pour constituer la table « households ». Elle peut aussi servir à constituer un fichier de calage pour le modèle de localisation résidentielle. Il y a cependant une lacune : il n'y a pas de données sur le revenu. Cette donnée a été estimée grâce au fichier de la DGI.

## 7.4 Les données de réseaux

Présentons quelques données descriptives des graphes des réseaux VP et TC du modèle de trafic de l'IAURIF pour l'année 1999 qui seront utilisés par METROPOLIS.

### 1) Réseau routier

Le réseau routier de l'année 1999 comprend les autoroutes, les voies rapides, les routes nationales et départementales, ainsi que les voies communales les plus importantes.

Il comporte environ 6900 arcs bidirectionnels et 4600 noeuds. A chaque arc sont attachés les attributs suivants:

- le nom,
- le statut ou type,
- la longueur (en dam),
- le sens de circulation,
- le nombre de voies par chaussée,
- le nombre de chaussées,
- la classe de vitesse (1 à 46), fonction du type de voie, de nombre de files et de l'environnement,
- la classe de capacité (1 à 46), fonction du type de voie, de nombre de files et de l'environnement,
- le nombre de poids lourds à l'heure de pointe,
- l'existence éventuelle d'un péage.

Les mouvements tournants aux carrefours et aux échangeurs sont quant à eux décrits dans un fichier externe.

### 2) Réseau de transports collectifs

Le réseau ferré de référence de 1999 comprend le réseau lourd (lignes SNCF et RATP) ainsi que l'ensemble du réseau de bus RATP et les principales lignes de bus OPTILE (222 lignes). L'ensemble est constitué d'environ 5000 arcs bidirectionnels et 4000 noeuds, auxquels s'ajoutent 1350 arcs de correspondance. Il y a sur le réseau lourd 190 missions à l'heure de pointe, chaque mission étant caractérisée par une fréquence et un temps de parcours inter-gares.

Rappelons enfin qu'on a envisagé dans la partie 5 de faire tourner le modèle de trafic à deux années intermédiaires dans la période de calage 1990-1999. Ce sont les années 1993 et 1996. Cela suppose à chaque fois de construire les graphes des réseaux VP et TC pour ces deux années, en plus de l'année de base 1990 et l'année cible 1999.

Au delà de 1999, il faudra recenser tous les projets mis en service jusqu'à 2004 et les projets au delà de 2004. Il faudra aussi faire des hypothèses sur leurs années de mise en service. Par exemple, les opérations de transports en commun qui seront en service entre 2000 et 2004 puis entre 2005 et 2010 sont :

- 2000 à 2004 :
  - entre les gares de Saint-Germain-en-Laye et Noisy-le-Roi.
  - le prolongement de Météor à Saint-Lazare
  - prolongement du T1 à Noisy-le-Sec
  - la ligne de grande ceinture SNCF (fermée depuis 1936)
  
- 2005 à 2010 :
  - prolongement du T2 à Bezons
  - le tramway des Maréchaux sud (TMS)
  - le prolongement du TVM à la Croix de Berny
  - le prolongement de la ligne 13 du métro à Gennevilliers
  - le prolongement de Météor à Olympiades
  - le TCSP Plateau de Saclay-Massy (phase 1, côté Massy et la section côté Saint-Quentin)

On peut envisager de faire tourner le modèle de trafic tous les cinq ans : 2005, 2010, 2015, 2020, 2025.

## 7.5 Synthèse

La plupart des données d'entrée ont été recueillies, mises au format UrbanSim et converties en format MySQL. Elles sont en cours de test dans le module « Consistency check » d'UrbanSim. Ce module élabore automatiquement un diagnostic sur les données et met en lumière les incohérences de la base de données. Il vérifie l'intégrité de la base de données.

En ce qui concerne les données de calage, le tableau suivant fait la comparaison des 3 sources (EGT, ENL, RGP) sur des données indispensables au calage des différents modules d'UrbanSim. On constate qu'aucune source n'est parfaite. C'est une conclusion universelle.

	<b>EGT</b>	<b>ENL</b>	<b>RGP</b>
Exhaustivité	Non	Non	Oui
Localisation résidentielle actuelle dans le maillage en 49 271 cellules	Oui	Non (localisation au département seulement)	Oui (mais estimée)
Trajectoire résidentielle / Cycle de vie	Dernière commune habitée	Oui	Dernière commune habitée pour 1990 seulement et en 300 zones pour 1999
Donnée de revenu	Oui	Non	Reconstituée avec le fichier DGI
Caractéristiques du logement précédent	Non	Oui	Non
Caractéristiques du logement actuel	Les principales	Oui	Oui
Caractéristiques sociales du ménage au moment du déménagement	Non	Oui	Non
Caractéristiques sociales actuelles du ménage	Oui	Oui	Oui

Tableau 16 : disponibilité d'information pour les ménages

Sur une suggestion du Pr Waddell, on envisage, avant de nous lancer dans un travail ardu de calage sur le maillage (49271 cellules), de faire quelques essais de calage à un niveau plus agrégé : les 1300 communes. Toutes les données décrites dans cette partie 7 sont disponibles à la commune. Il est donc possible de travailler d'ors et déjà sur ce jeu de données.

Cette phase de calage est fondamentale. Selon le Pr Waddell, il faut y consacrer le temps nécessaire, sans doute plusieurs mois.

## 8 Conclusions

Le présent rapport est le rapport de la 1<sup>ère</sup> année. Il a permis d'abord de présenter l'objectif du projet SIMAURIF et les deux modèles UrbanSim et METROPOLIS. Le rapport a été rédigé de manière à ne laisser aucune opacité dans les aspects aussi bien théoriques que pratiques. En particulier, les modèles mathématiques sous-jacents ainsi que les structures de données sont présentés de manière détaillée et complètement transparente, ce qui n'est pas toujours le cas dans d'autres projets de recherches.

Le choix d'UrbanSim comme logiciel d'urbanisation repose sur les arguments principaux suivants : ses fondements théoriques sont solides (modèles mathématiques de choix discret) ; le logiciel est flexible et peut être adapté au cas de la Région Ile-de-France ; il existe une communauté croissante d'utilisateurs de par le monde ; et on peut trouver une équipe de développeurs à l'Université de Washington. Tout le problème consiste à construire des bases de données en entrée et de calage suffisamment complètes et fiables.

Le choix de METROPOLIS comme logiciel de trafic à coupler à UrbanSim apparaît moins évident. Il prend en compte pour le mode voiture particulière, contrairement aux modèles statiques classiques, une étape supplémentaire de choix de l'heure de départ, ce qui intervient dans le niveau de congestion donc sur les temps de déplacements et au final sur les choix de localisation des emplois et des ménages. Cet aspect dynamique est donc l'argument principal en faveur de METROPOLIS. Cependant, il reste à programmer un modèle de demande de déplacements qui sera celui de l'IAURIF et trouver un moyen de prendre en compte le mode transports en commun.

Puis nous avons présenté les tâches concrètes effectuées depuis juin 2003. Ainsi, nous nous sommes familiarisés avec le logiciel UrbanSim d'un point de vue pratique en faisant fonctionner l'étude de cas Eugène-Springfield. Nous avons construit une base de données complète et volumineuse d'occupation du sol sur l'Ile-de-France dans un quadrillage très fin compatible (près de 50 000 cellules) dans le format d'entrée d'UrbanSim. Parallèlement et tout au long de l'année, une revue de la littérature a été effectuée afin de mieux appréhender les « land use models » (modèles d'urbanisation) et leurs connexions possibles avec les modèles de trafic.

Cette recherche est réellement inédite en France, de par la dimension de l'aire d'étude et de la complexité des interactions qu'elle vise à modéliser. D'autres équipes de recherche s'approprient à se lancer dans un projet similaire. Il ne nous a pas échappé que l'on risquait de rencontrer de réelles difficultés au cours de la deuxième année mais loin de nous décourager, ces difficultés constituent un véritable défi à relever. Cela nécessitera sans doute la contribution de compétences multiples : économètres, urbanistes, démographes, économistes, modélisateurs.

## BIBLIOGRAPHIE

- [1] Abraham J. E. et Hunt J. D. (1999). Firm location in the MEPLAN model of Sacramento, Transportation Research Board Annual Meeting, Washington DC, USA.
- [2] Alonso W. (1964). Location and land use, toward a general theory of land rent. Harvard University press, Cambridge, MA, USA.
- [3] Arnott R. (2001), Urban economic aggregates in monocentric and non-monocentric cities, Boston College, MA, USA.
- [4] Arnott, R., de Palma, A. et Lindsey, R., 1990, Economics of a Bottleneck, *Journal of Urban Economics*, 27, 111-130.
- [5] Arnott R., K. Arrow, A. Atkinson et J. Drèze, eds. (1994). « Public Economics – William Vickrey ». Cambridge University Press, Cambridge, MA
- [6] Ben-Akiva, M., Cyna, M. et de Palma, A., 1984, Dynamic model of peak period congestion, *Transportation Research B*, 18, 339-355.
- [7] Ben-Akiva, M., A. de Palma et Kanaroglou, P., 1986, Dynamic Model of Peak-Period Traffic Congestion with Elastic Arrival Rates, *Transportation Science*, 20, 164-181.
- [8] Boyer, K.D. (1998), *Principles of Transportation Economics*, Reading, Mass: Addison Wesley Longman.
- [9] Briassoulis H. (200). Analysis of Land Use Change: Theoretical and Modeling Approaches, Regional Research Institute, West Virginia University, Morgantown (WV), USA.
- [10] CERTU (1996) Review of existing land/use transport models. CERTU, Lyon, France.
- [11] Christaller, W (1933). Die zentralen Orte in Süddeutschland. Jena: Gustav Fischer, (Translated (in part), by Charlisle W. Baskin, as Central Places in Southern Germany. Prentice Hall 1966).
- [12] Computational Science Education Project (1995), "Mathematical Optimization".
- [13] Consortium ISHTAR (2003). Plaquette de projet, ENEA, Italie.
- [14] de la Barra T., (1989). Integrated land use and transport modeling. Cambridge University Press, Cambridge, UK.
- [15] de Palma, A., 1998 - 2002 , QUATUOR : Outils dynamiques de simulation pour la gestion des déplacements en Ile-de-France. DRAST/ PREDIT, N°98MT30, N°99MT35, N°00MT66, Université de Cergy-Pontoise, THEMA.
- [16] de Palma, A., P. Hansen et Labbé, M (1990), « Departure Time and Route Choice for the Morning Commute », *Transportation Science*, 24, 276-286.
- [17] de Palma, A. et Marchal, F. (1993), « Integration of Dynamic Traffic Models and Land-Use Forecasting with an Application to Paris Area », document de recherche du THEMA, université de Cergy-Pontoise.
- [18] de Palma, A. et Marchal, F., 2002, Real cases applications of the fully Dynamic METROPOLIS Tool-box : an Advocacy for Global Large-scale Mesoscopic Transportation Systems, *Networks and Spatial Economics*, 2 (4), Kluwer Academic Publishers, 347-369.
- [19] de Palma, A., Marchal, F. et Nesterov, Y., 1997, METROPOLIS : A Modular System for Dynamic Traffic Simulation, *Transportation Research Record* 1607.
- [20] Donnelly T. G., Chapin F., Weiss S. F. (1964). A probabilistic model for residential growth. Institute for research in social science, Chapel Hill, USA.

- [21] Du Crest T. (1999). Modélisation interactive usage du sol - transport : présentation d'une application sur Lyon à l'aide du logiciel Tranus, Séminaire de modèles de trafic, INRETS, Arcueil, France.
- [22] Fontan, C., 2003, Choix de l'heure de départ et coûts des délais : enquête et estimations en Île-de-France, Thèse en sciences Economiques, Université de Cergy-Pontoise.
- [23] Fujita M., Krugman P. et Venables A.J (2001). « The Spatial Economy », MIT Press.
- [24] Fujita M., (1989) Urban economic theory : Land use and city size, Cambridge University Press, Cambridge (MA), USA.
- [25] Goodwin, P.B. (1992), « A review of demand elasticities with special reference to short and long run effects of price changes », Journal of Transport Economics and Policy 26(2), 155-167.
- [26] Gudmundur F. Ulfrasson (2002), "Calibrating UrbanSim", présenté à la journée d'UrbanSim, University of Washington
- [27] Hendrickson, C. et Kocur, G., 1981, Schedule delay and departure time decisions in a deterministic model, Transportation Science, 15, 62-77.
- [28] ISIS (2004). Deliverable 1.2 – Methods and Software for the Prediction of the effects on people behaviour of the implementation of selected actions and policies, Italie.
- [29] Larouche P. (1965). The simulation of residential growth in the Montreal region. Quebec department of roads, Quebec, Canada.
- [30] Lerman, S.R. (1975) A disaggregate behavioral model of urban mobility decisions, Ph.D. dissertation, Massachusetts Institute of Technology.
- [31] Losch, A.(1954). The Economics of Location. New Haven, CT: Yale University Press/Oxford: Oxford University Press (original German edition, 1940).
- [32] Lowry, I.S. (1963), Location parameters in the Pittsburgh model, Papers and Proceedings of the Regional Science Association, 11, 145-165.
- [33] McFadden D. (1981), Econometric Models of Probabilistic Choice in “structural Analysis of Discrete Data with Econometric Applications”, MIT Press, Cambridge, MA, 198-272
- [34] Miller, E., D. Kriger, et J. D. Hunt (1998), “Integrated Urban Models for Simulation of Transit and Land Use Policies,” TCRP Web Document 9, Final Report, TCRP Project H-12, National Academy of Sciences, USA.
- [35] Merlin P. (1964). Modèle d'Urbanisation spontanée. Modèle sommaire d'évolution géographique des emplois. Institut d'aménagement et d'urbanisme de la région parisienne, Paris, France.
- [36] Mohring, H. (1972), « Optimization and scale economies in urban bus transportation », American Economic Review 62, 591-604.
- [37] Oum, T.H., W.G. Waters II et J-S. Yong (1992), « Concepts of price elasticities of transport demand and recent empirical estimates: An interpretative survey », Journal of Transport Economics and Policy 26(2), 139-154, 164-169.
- [38] Putman, S. H., (1995). EMPAL and DRAM Location and LandUse Models: An Overview Paper distributed at the transportation model improvement programme's landuse modelling conference, Dallas Texas, February 19-21.
- [39] Quinet, E. (1998), « Principes d'Economie des Transports », Economica, Paris.
- [40] Southworth F., (1995). A technical review of urban landuse – transportation models as tools for evaluating vehivule travel reduction strategies, Center for transportation analysis, Energy division, Oak Ridge national laboratory, TN, USA.
- [41] Thunen JH, von (1826). « Der isolierte Staat in eziehung auf Landwirtschaft und Nationalökonomie ». Gustav Fisher, Stuttgart. Translation by C.M. Wartenburg (1966). The Isolate State. Oxford University Press.



- [42] Urban Simulation project (2003). "UrbanSim user manual", University of Washington
- [43] Urban Simulation project (2000). "UrbanSim application guide, Eugene-Springfield, Oregon", University of Washington
- [44] Urban Simulation project (1999). "Longitudinal calibration of UrbanSim for Eugene-Springfield", University of Washington
- [45] Vickrey, W.S. (1969). « Congestion Theory and Transport Investment ». American Economic Review (Papers and Proceedings) 59, 251-261.
- [46] Waddel P., Borning A., Noth M., Freier N. Becke M et Ulfarson G. (2002). Microsimulation of Urban Development and Location Choices: Design and Implementation of Urbansim. Forthcoming in Network and Spatial Economics, 2002.
- [47] Waddel P., Borning A., Noth M., Freier N. Becke M et Ulfarson G. (2002). Microsimulation of Urban Development and Location Choices: Design and Implementation of Urbansim. A paraître dans Network and Spatial Economics, 2002.

# ANNEXE 1 :

## ETUDE DE CAS EUGENE –SPRINGFIELD :

### PRESENTATION DES DONNEES

#### ▪ Occupation du sol

En attribuant un type urbain à chaque cellule, nous définissons en fait son occupation du sol principale. Pour le cas d'Eugène, 25 types ont été définis :

Devtype	Nom	Nombre de logements		Surface non résidentielle Sqft		Usage principal
		Min	Max	Min	Max	
1	R1	1	1	0	999	Résidentiel
2	R2	2	4	0	999	Résidentiel
3	R3	5	9	0	999	Résidentiel
4	R4	10	14	0	2499	Résidentiel
5	R5	15	21	0	2499	Résidentiel
6	R6	22	30	0	2499	Résidentiel
7	R7	31	75	0	4999	Résidentiel
8	R8	76	65000	0	4999	Résidentiel
9	M1	1	9	1000	4999	Mixte R/C
10	M2	10	34	2500	4999	Mixte R/C
11	M3	10	34	5000	24999	Mixte R/C
12	M4	10	34	25000	49999	Mixte R/C
13	M5	10	34	50000	9999999	Mixte R/C
14	M6	31	65000	5000	24999	Mixte R/C
15	M7	31	65000	25000	49999	Mixte R/C
16	M8	31	65000	50000	9999999	Mixte R/C
17	C1	0	9	1000	24999	Commercial
18	C2	0	9	25000	49999	Commercial
19	C3	0	9	50000	9999999	Commercial
20	I1	0	9	1000	24999	Industriel
21	I2	0	9	25000	49999	Industriel
22	I3	0	9	50000	9999999	Industriel
23	GV	0	9	1000	9999999	Bureau
24	Vacant urbanisable	0	0	0	0	Vacant urbanisable
25	Non urbanisable	0	0	0	0	Non urbanisable

*Tableau 17 : types de développement*

## ▪ Emplois

Les emplois sont classés en 15 secteurs, issus du système standard de classification industrielle SIC du Département du Travail américain :

Secteur	Codes de SIC	Secteur	Secteur agrégé	Effectif
0	99 – 999	Agriculture	Basique	1 033
1	2400 – 2499	Filières du Bois	Basique	5 634
2	2500 – 2599, 3200 – 3999	Autres matières durables	Basique	5 056
3	2000 – 2099,	Produits alimentaires	Basique	1 316
4	2100 – 2399, 2600 – 3199	Autre non durables	Basique	3 064
5	1500 – 1799	Construction	Basique	4 405
6	1000 – 1499	Industries extractives	Basique	49
7	4000 – 4999	Transports	Basique	5 009
8	5000 – 5199	Commerce de gros	Basique	5 761
9	5200 – 5999	Commerce de détail	Commerce de détail	22 231
10	6000 – 6999	Finance, Assurance, immobilier	Service	5 457
11	7000 – 8199	Services	Service	28 751
14	8200 – 8299	Education	Service	8 648
15	9000 - 9999	Gouvernemental	Service	5 217
Total	101631			

Tableau 18 : codes des emplois

## ▪ Ménages

Les ménages sont classés en fonction de 5 caractéristiques :

- le revenu
- la taille des ménages
- l'âge du chef de ménage
- le nombre d'enfants
- le nombre d'actifs

Revenu	Catégorie	Effectif en 1994
<5000\$	1	2 622
De 5000\$ à 9999\$	2	10 812
De 10000\$ à 14999\$	3	15 354
De 15000\$ à 24999\$	4	15 532
De 25000\$ à 34999\$	5	18 345
De 35000\$ à 49999\$	6	9 454
De 50000\$ à 74999\$	7	10 154
De 75000\$ à 99999\$	8	6 522
Plus de 100000\$	9	2 438
Total		91 233

Tableau 19 : Catégories de ménages en fonction du revenu

<b>Age du chef</b>	<b>Catégorie</b>	<b>effectif</b>
De 15 à 24	1	9 027
De 25 à 34	2	18 860
De 35 à 44	3	22 746
De 45 à 54	4	12 796
De 55 à 64	5	9 422
De 65 à 74	6	10 510
75 et plus	7	7 872
Total		91 233

*Tableau 20 : distribution de l'âge du chef de ménage*

<b>Taille du ménage</b>	<b>Catégorie</b>	<b>Effectif</b>
1	1	21 376
2	2	36 797
3	3	14 819
4	4	11 462
5 ou plus	5	6 779
Total		91 233

*Tableau 21 : distribution de la taille des ménages*

<b>Nombre d'enfants</b>	<b>Catégorie</b>	<b>Effectif</b>
0	1	63 069
1 ou plus	2	28 164
Total		91 233

*Tableau 22 : distribution du nombre d'enfants*

<b>Nombre d'actifs</b>	<b>Catégorie</b>	<b>Effectif</b>
De 0 à 1	1	14 845
De 1 à 2	2	33 091
Plus de 2	3	43 297
Total		91 233

*Tableau 23 : distribution du nombre d'actifs*

## ANNEXE 2 :

### ETUDE DE CAS EUGENE-SPRINGFIELD :

### PARAMETRES DES MODULES

- Module de choix de localisation des emplois

Variable	Définition	Remarques
DEVM2	Variable muette DevM2	1 où le devtype est celui concerné
DEVM3	Variable muette DevM3	
DEVM4	Variable muette DevM4	
DEV11	Variable muette DevI1	
DEV12	Variable muette DevI2	
DEV13	Variable muette DevI3	
DEVG	Variable muette DevG	
ACCPOP	Accessibilité à la population	
SAME600	Nombre d'emplois de même secteur dans un périmètre de 600 m	
BAS600	Nombre d'emplois de secteurs basiques dans un périmètre de 600 m	
RET600	De même pour les secteurs de commerce de détail	
SER600	Même pour les secteurs de service	
UNIT600	Nombre de logements dans un périmètre de 600 m	
LVAC600	Prix moyen de terrain dans le périmètre	
HWY	Variable muette de l'autoroute	
DISTHWY	Distance à l'autoroute	
LogSqft	Logarithme (surface non résidentielle dans la cellule)	
LogTotUnit	De même pour le nombre de logements	
LogTotVal	De même pour la valeur totale du bâti	
LogAccPop	De même pour l'accessibilité à la population	
LogAccEmp	De même pour l'accessibilité aux emplois	
LogDistHwy	De même pour la distance à l'autoroute	
cbdTime	Temps de parcours au centre ville	
airTime	Temps de parcours à l'aéroport	
LogUnit600	Logarithme (nombre de logements dans le périmètre de 600 m)	
LogLVAC600		
LogValUnit600	Logarithme (valeur des logements dans un périmètre de 600 m)	

devRes	Devtypes résidentiels (1-8)	1 si devtype est entre 1 à 8
devMix	devtypes mixtes (9-16)	1 si devtype est entre 9 à 16
dev17	Commercial 17	1 si devtype est entre 17
dev18	Commercial 18	1 si devtype est entre 18
dev19	Commercial 19	1 si devtype est entre 19
devIndGov	Devtypes industriel ou gouvernemental (20-23)	1 si devtype est entre 20 et 23
s00	Nombre d'emplois de secteur 00 dans le périmètre de 600 m	
...	...	
s15	Nombre d'emplois de secteur 00 dans le périmètre de 600 m	

Tableau 24 : Les variables de modèle de choix de localisation des emplois

▪ Module de choix de localisation des ménages

Variable	Définition	Remarque	Coefficient estimé	t-statistique
lowIncome	Limite supérieure de couche de bas revenu, 0-(cette valeur)			
highIncome	Limite supérieure de couche de haut revenu (lowincome – cette valeur)			
DEV1	Variable muette pour devtype 1	1 si devtype = 1		
....	...			
DEV16	Variable muette pour devtype 16			
DevResLow	0/1 variable muette pour devtypes résidentiels à densité basse (1, 2, 3)			
DevResMid	0/1 variable muette pour devtypes résidentiels à densité moyenne (4, 5, 6)		0.491776	3.043
DevResHigh	0/1 variable muette pour devtypes résidentiels denses (7,8)		0.673393	3.318
DevMixLow	0/1 variable muette pour devtypes mixtes à densité basse (9, 10, 11, 12, 13)		0.418055	2.474
DevMixHigh	0/1 variable muette pour devtypes mixtes denses (14,15,16)		0.418055	2.474
DevCIG	0/1 variable muette pour devtypes commercial, industriel, gouvernemental (17-23)			
LogTotUnit	Logarithme (nombre de logements)		-0.63749	-4.352
LogTotResVal	Logarithme (valeur de logements)		0.850695	5.84
LogTotImpVal	Logarithme (valeur totale du bâti)			
UnitAge	L'âge du quartier	1994 - année de construction, 0		

		pour AC>94		
LogAccEmp0	Logarithme (accessibilité aux emplois pour les ménages avec 0 voiture)		0.206453	2.314
...	...			
LogAccEmp3	Logarithme (accessibilité aux emplois pour les ménages avec 3 voitures)			
LogAccPop	Logarithme (accessibilité à la population)		-0.5007	-1.9
LogUnit600	Logarithme (nombre de logements dans périmètre de 600 m)		-0.61048	-3.46
LogValUnit600	Logarithme (valeur unitaire de logement dans le périmètre)		-0.43077	-3.381
LogSqftInd600	Logarithme (surface industrielle dans le périmètre)		7.01E-02	4.338
LogSqftCom600	Logarithme (surface commerciale dans le périmètre)		8.53E-02	4.015
LogSqftTot600	Logarithme (surface non résidentielle dans le périmètre)			
PctRes600	Pourcentage de surface résidentielle dans le périmètre			
PctMix600	De même pour mixte			
PctCom600	De même pour commerciale			
PctInd600	De même pour industrielle		-1.84E-02	-2.114
PctGov600	De même pour bureaux		-1.84E-02	-2.114
Emp0_600	Nombre d'emplois de secteur 0 dans un périmètre de 600 m		5.42E-03	2.922
Emp1_600	De même pour secteur 1			
Emp2_600	De même pour secteur 2		-3.23E-03	-1.99
Emp3_600	De même pour secteur 3		1.94E-03	1.796
Emp4_600	De même pour secteur 4		-2.03E-03	-2.503
...	...			
Emp8_600	De même pour secteur 8		2.48E-03	3.138
...	...			
Emp11_600	De même pour secteur 11		2.52E-04	3.157
...	...			
Emp14_600	De même pour secteur 14		-3.42E-04	-2.497
Emp15_600	De même pour secteur 15			
LogDistHwy	Logarithme (distance à l'autoroute)		0.111651	2.126
CbdTime	Temps de parcours au centre ville en minutes		-7.01E-02	-3.588
AirTime	Temps de parcours à l'aéroport		4.29E-02	5.467
LogLVAC600	Logarithme (valeur de terrain par acre dans un périmètre de 600 m)			

C0LogAccEmp0	Logarithme (accessibilité aux emplois pour le ménage à une voiture)	0 si le taux de motorisation est différent	-0.49957	-2.045
...	...			
C3LogAccEmp3	Logarithme (accessibilité aux emplois pour le ménage à 3 voitures)			
C0LogUnit600	Logarithme (nombre de logements dans le périmètre si le ménage à 0 voiture)	0 si le taux de motorisation est différent	1.114527	4.484
C1LogUnit600	De même avec 1 voiture		0.694446	4.881
C2LogUnit600	De même avec 2 voitures		0.203608	1.829
C3LogUnit600	De même avec 3 voitures			
IncLogSqftInd600	Revenu * Logarithme (surface industrielle dans le périmètre)		-1.68E-06	-4.118
IncLogSqftCom600	De même pour surface commerciale		-1.39E-06	-3.063
IncLogValUnit600	Revenu * Logarithme (valeur unitaire de logement dans le périmètre)			
IncLogTotUnit	Revenu * Logarithme (nombre de logements)			
IncCost	hc = (revenu – valeur totale des logements)/(10*nombre de logements)/1000	hc = ( Income(\$) - totresval/(10*tot units) ) / 1000		
IncCost2	HCT = Ln(IncCost + 1) if IncCost>0, sinon hct = IncCost	HCT = Ln(IncCost + 1) if IncCost>0, else hct = IncCost	0.315134	6.563
SizeLogTotUnit	Taille de ménage * Logarithme (nombre de logements)		-0.21446	-5.142
SizeLogUnit600	Taille du ménage * logarithme (nombre de logements dans le périmètre)		0.178222	3.519
R678NC	Variable muette pour les ménages sans enfant et pour les devtypes 6 à 8			
DEVMNC	Variable muette pour les ménages sans enfant et pour les devtypes 9 à 16			
CHILDN	Nombre d'enfants par cellule	# of children (0, 1, 2+) * units in cell		
R678A40	Variable muette pour l'âge du chef de ménage <40 et pour devtypes 6 à 8			



DEVMA40	Variable muette pour l'âge du chef de ménage <40 et pour devtypes 9 à 16			
IYB	Revenu * année de construction			
LOGRET	Logarithme (nombre d'emplois en commerce de détails dans le périmètre)			
CARLRET	Variable muette pour les ménages ayant moins de voitures que des actifs * LOGRET			
CARLMDN	Variable muette pour les ménages ayant moins de voitures que des actifs * nombre de logements dans le périmètre			
ACCEW0H	Variable muette pour le ménage sans actif * accessibilité en VP aux emplois			
ACCEW1H	Variable muette pour le ménage ayant actif * accessibilité en VP aux emplois			
ACCEC0	Variable muette pour le ménage sans voiture * accessibilité en TC aux emplois			
		Ln(L(0))	-2604.22	
		Ln(L(b))	-2386.82	
	Qualité d'estimation	R2	0.08348	
		Adj. R2	0.08068	
		Observations	1131	

Tableau 25 : Les variables de modèle de choix de localisation des ménages

- Module de développement immobilier

Ce module simule les transitions au niveau de l'occupation du sol des cellules. Pour chaque type urbain, un tableau indique les transitions possibles et les valeurs des paramètres pour ces transitions.

Variable	Définition
OldDev	Devtype de départ
NewDev	Devtype représenté par l'alternative (variable dépendant du logit multinomial)
Constant	Constante spécifique à l'alternative pour les devtypes 1 à 23
Develop	Variable muette pour toutes les alternatives qui représentent le devtype (1-23)
LogSqft	Logarithme (surface non résidentielle)
LogUnits	Logarithme (nombre de logements)
Prx_dev	Proximité au type urbain existant (1000 - distance au développement); 0 si distance > 1000 m
LagSame	Nombre de cellules urbanisées au même devtype dans le périmètre de 600

	m et pendant les 3 dernières années
LagRes	De même pour les devtypes résidentiels
LagMix	De même pour les devtypes mixtes
LagCom	De même pour les devtypes commerciaux
LagInd	De même pour les devtypes industriels
LagGov	De même pour les devtypes bureaux
LagDev	Nombre de cellules urbanisées dans le périmètre et pendant les 3 dernières années
LagUnits	Nombre de logements ajoutés aux cellules dans le périmètre et pendant les 3 dernières années
LagSqft	Surface non résidentielle ajoutée aux cellules dans le périmètre et pendant les 3 dernières années
PctSame	Pourcentage de terrain urbanisé dans le périmètre qui a le devtype d'alternative
PctRes	Pourcentage de terrain urbanisé dans le périmètre ayant devtypes résidentiels (R1-R8)
PctMix	Pourcentage de terrain urbanisé dans le périmètre ayant devtypes mixtes (M1-M8)
PctCom	De même pour devtypes commerciaux (C1-C3)
PctInd	De même pour devtypes industriels (I1-I3)
PctGov	De même pour devtype bureaux (GV)
PctDev	Pourcentage de terrain dans le périmètre en zone urbaine
LogLandval	Logarithme (valeur de terrain dans le périmètre)
Hwy	Distance à l'autoroute < 300 m
Art	Distance à une national < 300 m
Flood	Cellule en zone inondable * variable muette pour changement de devtype
Slope	Pente moyenne de cellule * variable muette pour changement de devtype
Stream	Cellule située dans les berges
UGB	Cellule située en dehors de la limite d'expansion de ville
Wetland	Cellule en terres marécageuses * variable muette pour changement de devtype
Commit	Terrain réservé dans la cellule * variable muette pour changement de devtype
Cost	Coût de développement pour l'alternative
Revenue	Revenu (prix actuel estimé pour l'alternative)
Vacancy	Taux de vacance pondéré local et régional pour devtype

Tableau 26 : Les variables de modèle de développement immobilier pour les quels les coefficients doivent être estimés par transitions possibles

▪ Module de fixation du prix de terrain

Variable	Définition	Variable	Coefficient	t-statistique
Constant	Intersection	INTERCEP	12.95617	46.439
R1	Variable muette pour le devtype concerné	1 si devtype = devtype concerné,	-4.46817	-22.067

		sinon 0		
R2			-4.46651	-20.528
...			...	...
R7			-4.64254	-19.229
R8			-4.95183	-18.233
M1			-4.31538	-19.34
M2			-4.3027	-17.629
...			...	...
M7			-4.44888	-16.22
M8			-4.62374	-15.293
C1			-4.53809	-19.604
C2			-4.67112	-17.966
C3			-4.16469	-15.085
I1			-4.61162	-19.357
I2			-4.73466	-17.751
I3	Variable muette pour le devtype concerné	1 si devtype = devtype concerné, sinon 0	-4.9369	-17.044
GV			-4.58708	-19.235
P1			-0.47965	-18.874
P2			-3.38358	-10.775
P3			0.484352	11.431
P5			-1.22695	-13.828
P6			0.028826	0.796
P9			0.198052	5.251
P12			-0.16427	-4.297
P13			0.647943	4.344
P15			0.425899	15.363
P17			0.701282	9.866
P19			-0.35068	-2.899
P21			1.559043	7.11
P23	0.877028	17.535		
P24	-0.88913	-7.931		
LGLVAL	Logarithme (prix de terrain)			
IMPVAL	Valeur totale de bâtis			
LGIVAL	Ln(impval+1) : logarithme (valeur de bâtis)		0.476122	25.407
LGUNIT	Ln(units+1) : logarithme (nombre de logements)		0.050865	1.999
LGSQFT	Ln(sqft+1) : logarithme (surface non résidentielle)			
LGUNIT6	Ln(units6+1) : logarithme (nombre de logements dans le périmètre de 600 m)		0.103744	9.156
LGSQFT6	Ln(sqft6+1) : logarithme (surface non résidentielle dans le périmètre)		0.018958	5.43
LGEMP6	Ln(totemp6+1) : logarithme (nombre d'emplois dans le périmètre)		0.057538	6.713

LGAVU6	Ln(avgvalu6+1) : logarithme (valeur moyenne de logement dans le périmètre)			
LGAVL6	Ln(avglval6+1) : logarithme (prix moyen de terrain dans le périmètre)			
LGACCP	Logarithme (accessibilité à la population)			
LGACCE	Logarithme (accessibilité aux emplois)	LGACCE	-0.27454	-14.105
LGPCTD	Logarithme (pourcentage de terrain urbanisé dans le périmètre)			
LGPCTR	Logarithme (pourcentage de résidentiel dans le périmètre)			
LGPCTM	De même pour mixte			
LGPCTC	De même pour commercial			
LGPCTI	De même pour industriel			
LGPCTG	De même pour bureaux			
HWY1	Variable muette pour la distance à l'autoroute	1 si dist<=300m sinon 0	-0.3391	-6.407
LGHWY	Log (distance à l'autoroute)	LGHWY	-0.05951	-7.53
WETLAND	Variable muette pour la présence de terres marécageuses	0 / 1		
UGB	Variable muette pour être située en dehors de limite d'expansion de ville	0 / 1		
STREAM	Variable muette pour la présence de berge	0 / 1		
SLOPE	Variable muette pour la présence de pente forte	0 / 1		
FLOOD	Variable muette pour la présence de zones inondables	0 / 1		
LRVACAN	Taux de vacance local et régional			
	R <sup>2</sup> ajusté			0.83

Tableau 27 : Les variables du modèle de fixation du prix de terrain

Liste des paramètres à estimer séparément par fichier :

Nom du paramètre
PROBABILITY_OF_RELOCATING

Tableau 28 : liste des paramètres à estimer dans le fichier *AnnualRelocationRatesForHouseholds.tab* à introduire sur une segmentation des ménages en fonction de l'âge et du revenu

Nom du paramètre
JOB_RELOCATION_PROBABILITY

Tableau 29 : liste des paramètres à estimer dans le fichier *AnnualRelocationRatesForJobs.tab* à introduire pour chaque secteur d'activité

Nom du paramètre
alternative_specific_constant
ln_commercial_sqft
ln_residential_units
Distance_to_development
proximity_to_development
n_recent_transitions_to_same_type_within_walking_distance
n_recent_transitions_to_residential_within_walking_distance
n_recent_transitions_to_mixed_use_within_walking_distance
n_recent_transitions_to_commercial_within_walking_distance
n_recent_transitions_to_industrial_within_walking_distance
n_recent_transitions_to_governmental_within_walking_distance
n_residential_units_recently_added_within_walking_distance
ln_n_residential_units_recently_added_within_walking_distance
commercial_sqft_recently_added_within_walking_distance
ln_commercial_sqft_recently_added_within_walking_distance
percentage_cells_of_the_same_type_within_walking_distance
percentage_of_residential_cells_within_walking_distance
percentage_of_mixed_use_cells_within_walking_distance
percentage_of_commercial_cells_within_walking_distance
percentage_of_industrial_cells_within_walking_distance
percentage_of_governmental_cells_within_walking_distance
ln_land_value
is_near_arterial
is_near_highway
is_in_floodplain
is_on_steep_slope
is_in_stream_buffer
is_outside_urban_growth_boundary
is_in_wetland

percent_water
percent_stream_buffer
percent_flood_plain
percent_wetland
percent_slope
percent_open_space
percent_public_space
percent_roads
plantype_1
plantype_2
plantype_3
plantype_4
plantype_5

Tableau 30 : liste des paramètres à estimer dans le fichier DeveloperModelCoefficients.tab à répéter pour chaque transition possible d'un Devtype à l'autre

Nom du paramètre
devtype_1
devtype_2
devtype_3
devtype_4
devtype_5
devtype_6
devtype_7
devtype_8
devtype_9
employment_in_sector_1_within_walking_distance
employment_in_sector_2_within_walking_distance
employment_in_sector_3_within_walking_distance
employment_in_sector_4_within_walking_distance
employment_in_sector_5_within_walking_distance
employment_in_sector_6_within_walking_distance
employment_in_sector_7_within_walking_distance
devtype_industrial_or_governmental
same_sector_employment_within_walking_distance
basic_sector_employment_within_walking_distance
retail_sector_employment_within_walking_distance
service_sector_employment_within_walking_distance
residential_units_within_walking_distance
average_land_value_per_acre_within_walking_distance
is_near_arterial
is_near_highway
ln_commercial_sqft
ln_residential_units
ln_total_job_locations
ln_available_job_locations
ln_total_value
ln_work_access_to_population_1

ln_work_access_to_employment_1
ln_distance_to_highway
travel_time_to_CBD
travel_time_to_airport
ln_residential_units_within_walking_distance
ln_average_land_value_per_acre_within_walking_distance
ln_improvement_value_per_residential_unit_within_walking_distance

*Tableau 31: liste des paramètres à estimer dans le fichier  
EmploymentLocationChoiceCoefficients.tab*

<b>Nom du paramètre</b>
devtype_1
devtype_2
devtype_3
devtype_4
devtype_5
devtype_6
devtype_7
devtype_8
devtype_9
devtype_category_residential
devtype_category_commercial
devtype_category_industrial
devtype_category_economic
devtype_category_unknown
percent_residential_within_walking_distance
percent_commercial_within_walking_distance
percent_industrial_within_walking_distance
percent_economic_within_walking_distance
percent_unknown_within_walking_distance
employment_within_walking_distance_1
employment_within_walking_distance_2
employment_within_walking_distance_3
employment_within_walking_distance_4
employment_within_walking_distance_5
employment_within_walking_distance_6
employment_within_walking_distance_7
building_age
cost_to_income_ratio
household_size_and_ln_residential_units
household_size_and_ln_residential_units_within_walking_distance
household_with_no_children_in_high_density_residential
household_with_no_children_in_mixed_use
income_and_ln_residential_units
income_and_ln_commercial_sqft_within_walking_distance
income_and_ln_industrial_sqft_within_walking_distance
income_and_ln_total_residential_value_per_residential_unit_within_walking_distance

income_and_year_built
ln_commercial_sqft_within_walking_distance
ln_distance_to_highway
ln_home_access_to_employment_0_cars
ln_home_access_to_employment_1
ln_home_access_to_employment_1_car
ln_home_access_to_employment_2
ln_home_access_to_employment_2_cars
ln_home_access_to_employment_3plus
ln_home_access_to_employment_3plus_cars
ln_home_access_to_employment_no_workers
ln_home_access_to_employment_transit
ln_home_access_to_employment_workers
ln_home_access_to_population_1
ln_industrial_sqft_within_walking_distance
ln_residential_improvement_value_per_residential_unit
ln_residential_units
ln_available_residential_units
ln_residential_units_within_walking_distance
ln_residential_units_within_walking_distance_0_cars
ln_residential_units_within_walking_distance_1_car
ln_residential_units_within_walking_distance_2_cars
ln_residential_units_within_walking_distance_3plus_cars
ln_residential_units_within_walking_distance_fewer_cars_than_workers
ln_retail_within_walking_distance
ln_retail_within_walking_distance_fewer_cars_than_workers
ln_total_land_value_per_acre_within_walking_distance
ln_total_nonresidential_sqft_within_walking_distance
ln_total_residential_value_per_residential_unit
ln_total_residential_value_per_residential_unit_within_walking_distance
percent_low_income_households_within_walking_distance
percent_low_income_households_within_walking_distance_if_high_income
percent_low_income_households_within_walking_distance_if_low_income
percent_low_income_households_within_walking_distance_if_mid_income
percent_minority_households_within_walking_distance
percent_minority_households_within_walking_distance_if_minority
percent_minority_households_within_walking_distance_if_not_minority
residential_units_when_household_has_children
travel_time_to_airport
travel_time_to_CBD
young_household_in_high_density_residential
young_household_in_mixed_use

*Tableau 32 : liste des paramètres à estimer dans le fichier  
HouseholdLocationChoiceCoefficients.tab*



Nom du paramètre
devtype_1
devtype_2
devtype_3
devtype_4
devtype_5
devtype_6
devtype_7
devtype_8
devtype_9
ln_percent_commercial_within_walking_distance
ln_percent_residential_within_walking_distance
ln_percent_industrial_within_walking_distance
ln_percent_economic_within_walking_distance
ln_percent_unknown_within_walking_distance
plantype_1
plantype_2
plantype_3
plantype_4
plantype_5
is_in_flood_plain
is_in_stream_buffer
is_in_wetland
is_near_highway
is_on_steep_slope
is_outside_urban_growth_boundary
ln_average_land_value_per_acre_within_walking_distance
ln_average_total_value_per_residential_unit_within_walking_distance
ln_commercial_sqft
ln_commercial_sqft_within_walking_distance
ln_distance_to_highway
ln_home_access_to_employment_1
ln_home_access_to_population_1
ln_residential_units
ln_residential_units_within_walking_distance
ln_total_employment_within_walking_distance
ln_total_improvement_value
percent_water
percent_stream_buffer
percent_flood_plain
percent_wetland
percent_slope
percent_open_space
percent_public_space
percent_roads

Tableau 33 : liste des paramètres à estimer dans le fichier LandPriceCoefficients.tab

<b>Nom du paramètre</b>
devtype_1
devtype_2
devtype_3
devtype_4
devtype_5
devtype_6
devtype_7
devtype_8
devtype_9
residential_units
non_residential_sqft

*Tableau 34: liste des paramètres à estimer dans le fichier  
ResidentialLandShareCoefficients.tab*

## **ANNEXE 3 :**

### **MODELE DE DEMANDE DE L'IAURIF :**

### **PROCEDURE DE REDUCTION DES MOTIFS**

En pratique, on considère un plus petit nombre de motifs (la matrice dont les éléments sont les motifs comporte des cellules dont les effectifs peuvent être regroupés). En se basant sur les volumes de déplacements de la matrice des déplacements fournie par l'EGT (Enquête Globale des Transports) de 1991, l'IAURIF propose les 8 motifs suivants (légèrement différents de ceux considérés dans le corps du texte):

- m=1 Domicile : noté « DO »
- m=2 Lieu de travail : noté « TR »
- m=3 Affaires professionnelles : noté « APR »
- m=4 Affaires personnelles : noté « APE »
- m=5 Achats : noté « ACH »
- m=6 Enseignement supérieur : noté « ENS »
- m=7 Ecole : noté « ECO »
- m=8 Santé : noté « SA »

Les déplacements à partir du domicile sont sept : {DO → TR, DO → APR, DO → APE, DO → ACH, DO → ENS, DO → ECO et DO → SA}.

On considère aussi les motifs ayant comme destination le Domicile. En pratique on peut transposer les matrices du matin pour obtenir celle du soir. Cette méthode est rapide et économique, mais suppose une symétrie entre la pointe du matin et du soir. On peut aussi traiter la matrice du soir indépendamment de celle du matin, pour tenir compte du fait que les coûts de déplacements du matin (par ex. :  $c_{ij}$ ) sont différents de ceux du soir (par ex.  $c_{ji}$ ). Les motifs de déplacement, qui correspondent à 14 matrices O-D, une fois désagrégés dans l'espace, représentent environ 70% des déplacements en Ile-de-France.

On peut regrouper les autres motifs de déplacement en un seul motif. Les 7 autres matrices correspondantes sont dans ce cas égales aux 7 matrices initiales (pour les déplacements ayant le domicile comme origine) multipliées par un coefficient à extraire de l'EGT. Notons {TR, APR, APE, ACH, ENS, ECO et SA} par « Tous motif sauf DO », les autres déplacements sont décrits par les sept motifs supplémentaires suivants :

- Tous Motifs sauf DO → TR
- Tous Motifs sauf DO → APR
- Tous Motifs sauf DO → APE
- Tous Motifs sauf DO → ACH
- Tous Motifs sauf DO → ENS
- Tous Motifs sauf DO → ECO

Tous Motifs sauf DO → SA

Ces 7 motifs supplémentaires ajoutent 21% des déplacements. Par conséquent, avec 21 motifs (7x3), il est possible de reconstituer 91% des déplacements.

Les derniers 9% de déplacements sont ajoutés aux 21 motifs précédents par une simple règle de 3 sur le total des déplacements.

Il est possible de simplifier davantage cette approche, par exemple en regroupant « affaires personnelles et santé ». De plus, si on peut considérer que les déplacements liés à l'école primaire se font essentiellement à pied, il est possible de les omettre.

# ANNEXE 4 :

## FONDEMENTS THEORIQUES DE METROPOLIS

### Introduction

La complexité de l'analyse des déplacements en milieu urbain fait appel de plus en plus à des outils de simulation de comportements des usagers des systèmes de transport. Pour ces études de planification, deux approches sont utilisées : la première est dite statique et la seconde est dite dynamique. Dans le cas, de la première approche, les flux de déplacements sont simulés pour une courte période avec une matrice constante. Par conséquent, le résultat des flux sur les arcs est constant. Il représente le nombre moyen d'usagers par arc. Le choix analysé dans ces modèles se limite aux choix d'itinéraires. La seconde approche prend en compte la variation des flux au cours de la période simulée. Deux types de modèles existent, le premier simule la variation des flux du réseau au cours de la période d'analyse en prenant comme donnée d'entrée une matrice dynamique. De ce fait, seul le choix d'itinéraire est simulé même s'il y a prise en compte des délais à l'arrivée. Ces délais sont considérés comme une conséquence de la dynamique de la matrice des déplacements et du choix d'itinéraire (ou de la congestion sur le réseau). Le choix de l'heure de départ reste exogène. Le second type de modèle considère que le choix de l'heure de départ dépend du niveau de la congestion sur la route et les usagers adaptent leur choix de l'heure de départ en fonction de la variation de la congestion au cours de la période simulée. Dans ce cas, il n'est pas nécessaire de fixer une matrice de déplacements dynamique. C'est le cas du modèle de simulation utilisé dans cette étude : METROPOLIS. Cet outil procède en deux temps : dans un premier temps, il détermine les choix de l'heure de départ des usagers puis dans un second temps il effectue l'affectation des flux sur le réseau (ou le choix d'itinéraire).

La seconde section de ce chapitre est consacrée à la description des principes de base du modèle dynamique. Ce modèle a été proposé pour la première fois par Vickrey [45] en 1969. Par la suite, différentes extensions ont été explorées (voir Arnott et *al.*[3]). Etant données les contraintes d'une formulation mathématique et d'optimisation d'affectation des déplacements dans les grands réseaux, une approche heuristique se révéla nécessaire. C'est cette dernière qui est développée dans METROPOLIS, l'outil de simulation dynamique. Les différentes étapes développées dans cet outil sont également décrites dans cette section. Des travaux très approfondis ont été menés de 1999 à 2002 dans le but d'adapter METROPOLIS à la Région Ile-de-France (voir de Palma [15]). Ils ont pu être réalisés grâce à une subvention du PREDIT. Dans la troisième section, les données utilisées et les résultats pour l'adaptation de METROPOLIS à la région sont exposés. La dernière section est consacrée aux conclusions.

### Description théorique du modèle

Le modèle de simulation METROPOLIS peut être décomposé en trois principaux modules : module du choix de l'heure de départ, un module de choix d'itinéraire (ou de route) et un module de mise à jour. Le modèle de base a été proposé pour la première fois par Vickrey [45], puis analysé et développé par Hendrickson et Kocur [27]. Cependant, la majorité des

travaux théoriques développés à ce sujet a été effectuée par Arnott, de Palma et Lindsey [3]. Par ailleurs, la formulation mathématique et le calcul analytique d'un équilibre atteignent assez rapidement leur limite, même dans le cadre d'un réseau très simple. Par conséquent, le recours à un modèle de simulation comportemental se révèle nécessaire. L'approche de simulation a été développée par Ben-Akiva et *al.* [6] pour l'analyse d'un réseau avec une paire origine – destination. Dans Ben-Akiva et *al.*[7], ce modèle empirique a été étendu afin de prendre en compte la demande élastique. Dans cette section, dans un premier temps le modèle de base, suivant les développements d'Arnott et *al.* [3], est exposé et dans un second temps l'approche empirique utilisée dans METROPOLIS est décrite.

### **Modèle de base**

Le modèle de base suppose que les individus lors de leurs déplacements effectuent un choix de l'heure de départ en fonction du temps de parcours espéré et de l'heure désirée d'arrivée. Soit une population de  $N$  automobilistes se déplaçant dans un réseau d'une zone origine vers une zone destination reliée par une route de capacité fixe  $s$ . Le temps de trajet est variable au cours de la période d'analyse (par exemple la pointe du matin) et dépend du nombre d'usagers sur la route. En effet, lorsque le nombre d'usagers sur la route est supérieur à la capacité, une file d'attente commence à se développer augmentant ainsi le temps de parcours des usagers. Les individus doivent choisir une heure de départ qui minimise leur coût de déplacement. Contrairement au modèle statique, le coût de déplacement dans le modèle se compose d'un coût de trajet et un coût de délai à l'arrivée. Plus précisément si un individu arrive en avance par rapport à son heure désirée, il subit un coût de délai précoce et s'il arrive en retard il subit un coût de délai tardif en plus du coût du temps de trajet. Pour simplifier, soit une heure désirée d'arrivée,  $t^*$ , commune à l'ensemble des automobilistes. Le coût d'un usager de déplacement,  $C(t_d, t^*)$ , d'un usager s'écrit :

$$C(t_d, t^*) = \alpha t t(t_d) + \beta \max[0, t^* - t_a] + \gamma \max[0, t_a - t^*].$$

Les paramètres,  $\beta$  et  $\gamma$  représentent respectivement le coût unitaire du temps de trajet, le coût unitaire d'une arrivée précoce et le coût unitaire d'une arrivée tardive. La variable  $t_a$  représente l'heure d'arrivée d'usagers à destination telle que  $t_a = t_d + t t(t_d)$  avec  $t t(t_d)$  le temps de parcours associé à l'heure de départ  $t_d$ . Le coût est composé deux termes. Le premier représente le coût du temps de trajet  $t t(t_d)$  qui dépend de l'heure de départ  $t_d$ . Le second terme représente le coût du délai précoce et le troisième terme représente le coût du délai tardif. Si l'individu arrive en avance, le second terme est positif alors que le troisième est nul (soit le maximum entre '0' et une valeur négative). Et si l'individu arrive en retard, le premier terme est nul et le second est positif. Sans perte de généralité, le modèle suppose que les individus choisissent une heure de départ de manière à minimiser leur coût de déplacement. A l'équilibre les usagers subissent le même coût de déplacement. Dans ce modèle, le coût moyen relatif au temps de trajet passé dans la file d'attente est égal au coût moyen relatif aux délais à l'arrivée. Le coût moyen  $CM$  est tel que :

$$CM = \frac{\beta\gamma}{\beta + \gamma} \frac{N}{s}$$

De cette relation, il est possible de constater qu'une hausse du coût unitaire précoce  $\beta$  ou tardif  $\gamma$  accroît le coût moyen  $CM$  du fait que la période des départs (différence entre le premier et le dernier départ) devient plus importante. Cependant, une variation du coût unitaire du temps de parcours ne modifie pas le coût moyen mais uniquement la valeur du flux départ et donc l'heure du temps de parcours maximal.

### ***Choix de modélisation empirique***

Le modèle de simulation METROPOLIS peut être décomposé en quatre principaux modules : module de choix de l'heure de départ, un module de choix d'itinéraire, un module de mise à jour et un module de choix modal. Dans les paragraphes suivants, le principe du modèle est résumé en se basant sur les publications de Palma et al. [19],[18].

### ***Module de choix de l'heure de départ***

Le module de choix de l'heure de départ est probabiliste de type logit. La probabilité qu'un usager choisisse l'heure de départ  $t_i$  dans l'un des intervalles est donnée par :

$$P(t_i) = \text{Prob}[-C(t_i) + \mu_d \varepsilon_i > -C(t_j) + \mu_d \varepsilon_j] \text{ pour tout } i \neq j,$$

où les  $\varepsilon$  sont des variables aléatoires indépendantes et identiquement distribuées suivant une loi double exponentielle et où paramètre  $\mu_d$  est un paramètre d'échelle (positif).

La probabilité de choisir l'intervalle de départ  $[t_i, t_i + dt_i]$  est donnée par le logit continu qui s'écrit :

$$P(t_i)dt_i = \frac{\exp[-C(t_i)/\mu_d]}{\int_0^T \exp[-C(x)/\mu_d] dx} dt_i.$$

### ***Module de choix d'itinéraire***

Le processus de choix d'itinéraire dans METROPOLIS est stochastique et modélisé par un logit multinomial. Le modèle déterministe est un cas particulier du modèle stochastique ; il correspond à la situation où le paramètre d'hétérogénéité tend vers 0. Ce processus est modélisé de façon désagrégée où l'individu doit sélectionner son itinéraire à chaque nouvelle intersection du réseau. Le modèle de simulation suppose que les individus disposent d'informations sur les temps de parcours sur l'ensemble des itinéraires alternatifs. Pour

déterminer le choix d'un itinéraire, l'individu doit combiner les informations relatives aux temps de parcours simulé et historique.

### ***Module de processus itératif***

Le processus itératif du modèle permet d'atteindre une situation d'équilibre pour lequel aucun usager n'a intérêt à modifier son heure de départ et son itinéraire sur le réseau. Pour cela, lorsque l'individu décide de son heure de départ et de son choix d'itinéraire, il doit tenir compte des informations historiques afin de mettre à jour les temps effectués lors des itérations précédentes. Ce processus de mise à jour est dit processus d'apprentissage. Différents processus sont disponibles dans METROPOLIS. Cependant, seul le processus d'apprentissage exponentiel est exposé car il est souvent adopté (comme par exemple lors du calage de METROPOLIS au réseau de l'Ile-de-France) et cela pour différentes raisons. La première est que ce processus est souvent utilisé dans la littérature comme dans les modèles de lissage exponentiel dans l'analyse des séries temporelles ou encore les modèles d'adaptation et de négociation des prix et des salaires. La seconde raison est que, par expérience, c'est le seul qui permet d'atteindre la situation d'équilibre assez rapidement avec le moins de fluctuations possible autour du point d'équilibre.

En faisant appel au processus d'apprentissage le temps de trajet,  $tt^{H,w+1}(t)$ , pour le jour  $w+1$  pour chaque heure de départ s'écrit :

$$tt^{H,w+1}(t) = (1 - \lambda)tt^{H,w}(t) + \lambda tt^{S,w}(t),$$

où :

$tt^{H,w}(t)$  est le temps de trajet Historique le jour précédent  $w$  pour l'heure de départ  $t$ ,

$tt^{S,w}(t)$  est le temps de trajet Simulé le jour précédent  $w$  pour l'heure de départ  $t$ , et

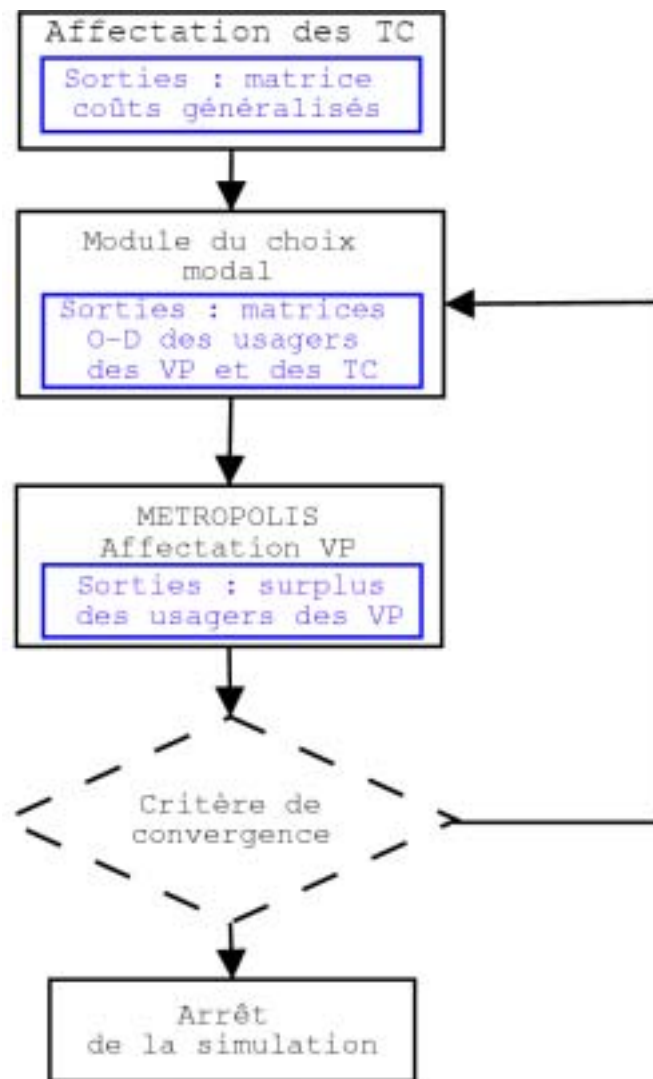
$\lambda$  est le paramètre d'apprentissage qui est strictement compris entre 0 et 0.5. Lorsque la valeur de  $\lambda$  tend vers 0, l'individu offre plus d'importance au temps de trajet historique et lorsqu'elle tend vers 0.5, l'individu offre plus d'importance au temps de trajet simulé.

L'équilibre dynamique du système est atteint lorsque aucun individu n'a intérêt à modifier son heure de départ et son itinéraire. Pour ce type de modèle, il n'existe pas, de nos jours, de formulation mathématique démontrant l'unicité d'équilibre du fait de la double convergence du modèle. En effet, l'outil de simulation suppose une convergence par rapport au choix de l'heure de départ et une convergence de choix d'itinéraire. Par conséquent, METROPOLIS fait référence à un état stationnaire expérimental. Un critère statistique STAC (Space-Time Aggregate Criterion) permet de vérifier la convergence du modèle. Il permet de comparer les temps historiques et simulés sur chaque arc. Théoriquement l'état stationnaire est atteint lorsque, pour l'ensemble des arcs, si, pour une heure donnée, le temps de parcours historique est égal au temps de parcours simulé.



### ***Module du choix modal***

Dans la version actuelle de METROPOLIS, les déplacements des usagers des transports en commun et de la voiture sont simulés séparément. Cependant, un module supplémentaire, intégré à METROPOLIS, prend en compte la simulation désagrégée du choix modal. Pour affecter un individu à un mode donné, il compare le surplus de l'individu qu'il en tire de l'usage des deux modes. Le surplus, dû à l'usage des transports en commun, est donné par le coût généralisé de ce mode et est fourni par un modèle externe, celui de l'IAURIF. Ce dernier modèle d'affectation des transports en commun fournit une matrice des coûts généralisés. Cette matrice de coûts généralisés reste constante au cours des itérations du modèle. La structure du modèle pour la prise en compte du choix modal a été représentée dans le schéma de la Figure 22.



*Figure 22 : processus de simulation globale*

Le calcul du choix modal s'effectue à chaque itération. Etant donnée que la matrice des coûts généralisés est constante, c'est uniquement la matrice des surplus relative à l'usage de la voiture qui varie. Au cours des itérations, la première étape consiste à calculer le choix modal

des usagers ce qui permet de reconstruire une nouvelle matrice origine – destination des usagers de la voiture particulière (VP). L'étape suivante simule l'affectation des déplacements en VP générant ainsi la nouvelle matrice des surplus des usagers des VP. Si le critère d'arrêt de convergence est vérifié, la simulation s'arrête sinon une nouvelle itération recommence avec un nouveau calcul du choix modal.

Le critère d'arrêt des simulations dans METROPOLIS avec choix modal suppose d'une part que la convergence est vérifiée à travers le critère STAC et d'autre part que la part de chaque mode de déplacement reste constante. En réalité, le second critère peut être ignoré puisque le critère STAC prend en compte le fait que les flux historiques et simulés sur les itinéraires doivent converger. Si la part des usagers de la voiture varie, le flux simulé ne correspond plus au flux historique et donc la convergence de la partie affectation des VP n'est pas atteinte et du coup l'ensemble du processus (affectation VP et choix modal).

La formulation adoptée, pour la modélisation du choix modal, est de type logit binaire de sorte que la probabilité de prendre la voiture s'écrit :

$$P_{VP} = \frac{\exp[V_{VP} / \mu_m]}{\exp[V_{VP} / \mu_m] + \exp[V_{TC} / \mu_m]}$$

où :

$V_{VP}$  représente l'utilité déterministe associée à l'usage de la voiture particulière,

$V_{TC}$  représente l'utilité déterministe associée à l'usage des transports en commun, et

$\mu_m$  est le paramètre d'hétérogénéité lié au choix modal.

### **Calcul de l'utilité des TC :**

L'utilité des TC,  $V_{TC}$ , dépend uniquement du coût associé au temps de parcours en TC,  $tt_{TC}$ , et d'une pénibilité,  $p_{TC}$ , associé à l'usage des TC. L'utilité des TC s'écrit dans ce cas :

$$V_{TC} = -\alpha_{TC} tt_{TC} - p_{TC},$$

où  $\alpha_{TC}$  représente le coût unitaire associé au temps de parcours en TC.

La majorité des modèles d'affectations des déplacements des usagers TC ne prennent pas en compte ni la congestion dans les transports en commun ni le choix de l'heure de départ. L'affectation se fait uniquement en fonction du temps de parcours (éventuellement généralisé). Il en résulte que le temps de parcours d'une origine à une destination correspond au temps de parcours sur le chemin le plus court. De ce fait, quelque soit le nombre de déplacements des usagers des TC, la matrice des temps de parcours ainsi que la matrice des coûts généralisés restent constantes d'une itération à une autre. Cette constance des coûts généralisés justifie le calcul du choix modal utilisé au cours des simulations.

### **Calcul de l'utilité des VP :**

A priori, le même raisonnement peut être appliqué pour les VP par la prise en compte du temps de parcours et d'une certaine pénibilité. Toutefois, dans le modèle d'affectation des VP, deux éléments diffèrent des TC :

- METROPOLIS prend explicitement en compte la dynamique de la congestion de sorte que le temps de parcours dépend de l'heure de départ,
- METROPOLIS prend en compte les coûts des délais à l'arrivée qui dépendent également de l'heure de départ.

Dans cette situation, la variation du temps de parcours et des délais est explicitement prise en considération. De ce fait, l'utilité d'un usager de la voiture  $V_{VP}$  est :

$$V_{VP} = -\alpha_{VP} tt_{VP}(t_d) - CDA_{VP}(t_d),$$

où  $tt_{VP}(t_d)$  et  $CDA_{VP}(t_d)$  représentent respectivement le temps de parcours et le coût de délai à l'arrivée associés à l'heure de départ  $t_d$ .

Le modèle utilisé dans METROPOLIS permet de prendre en compte de façon indirecte l'existence des captifs d'un mode ou d'un autre à travers la valeur de la pénibilité,  $p_{TC}$ , associée à l'utilité des TC. Les individus captifs de la voiture particulière se voient affecté d'une valeur de pénibilité  $p_{TC}$  très forte. Par contre pour, les captifs des TC se voient affectés d'une valeur de pénibilité très faible voir négative.

## Données utilisées

Pour la mise en place des simulations, l'outil dynamique METROPOLIS utilise des données dont la majorité est semblable à celles utilisées dans les modèles statiques. Elles concernent essentiellement les données relatives à l'offre et à la matrice origine - destination. Les données d'entrées et de sorties sont stockées dans une base de données MySQL. Ce qui permet d'extraire des informations agrégées comme par exemple le temps de parcours et le coût généralisé par paire origine – destination ou des informations spécifiques à chaque usager comme l'heure de départ, l'heure d'arrivée, son coût généralisé, etc. Cette approche permet aisément, à travers des requêtes SQL, la communication des résultats des simulations à d'autres outils de simulation comme Urbansim, tout comme il est possible de rapatrier des informations d'autres outils.

L'offre de transport est décrite par les données représentatives du réseau physique et des fonctions d'impédance ou de congestion gérant les flux sur les liens. Le réseau routier est fourni par l'IAURIF. Il est décrit par un ensemble de noeuds et un ensemble de liens (ou arcs). Les noeuds sont caractérisés par leur position géographique et leur numéro d'identification. Les liens sont caractérisés par un noeud d'origine et un noeud de destination, une longueur, le nombre de voies, la capacité du lien et son type. Dans l'outil de simulation METROPOLIS, à chaque lien est associée une fonction de congestion. La capacité du lien est égale à la somme des capacités des voies. Ces données sont tirées directement de la base de données du modèle statique de l'IAURIF.

Les données spécifiques à la demande se décomposent en deux groupes. Le premier groupe concerne les flux de déplacements décrits par une matrice origine – destination statique. Le second groupe concerne les paramètres comportementaux des usagers. L'un des avantages de METROPOLIS est qu'il est possible de créer plusieurs matrices origine - destination et plusieurs segments. En effet, à chaque matrice il est possible d'associer un ou plusieurs segments d'usagers.

Les sections suivantes permettent de décrire l'adaptation de METROPOLIS au réseau de l'Ile-de-France. Celui-ci couvre l'ensemble de la région d'Ile-de-France soit une surface de 12000 km<sup>2</sup>. Il décrit les déplacements entre 522 zones (ancien découpage) dont 34 zones à l'extérieur de la région Ile-de-France. Les 488 zones restantes regroupent les 1300 communes de la région francilienne. Plus une zone est proche de Paris, plus sa densité est élevée. En terme de voirie, le réseau est représenté par environ 13600 arcs monodirectionnels, l'équivalent de 29700 km de voies. La matrice des déplacements utilisés dans l'adaptation de METROPOLIS au réseau de l'IAURIF représente les flux de déplacements pendant l'heure de pointe du matin entre 7h30 et 8h30.

Les paramètres comportementaux dynamiques représentent les coûts unitaires relatifs au temps de trajet, au délai à l'arrivée précoce et le délai à l'arrivée tardive. Ils ont été estimés à partir de des données de l'enquête type MADDIF (voir Fontan [22]). Les études sur le sujet montrent généralement que la valeur des paramètres de la fonction de coût dépend du motif de déplacement. De plus, pour les motifs de déplacement autre que le travail, deux sous populations ont été constituées selon l'heure désirée d'arrivée. Les valeurs des coûts unitaires et ainsi que le paramètre d'hétérogénéité sont calculées pour quatre segments d'individus. Le tableau ci-dessous représente les valeurs des paramètres de ces quatre classes d'individus. Les valeurs entre parenthèses représentent la moyenne et l'écart-type d'une loi normale.

Groupe de motif	$\alpha$ (€/h)	$\beta$ (€/h)	$\gamma$ (€/h)	$\mu_d$ (€)	$T^*$
Travail vers Paris et PC	12,96	6,08	7,52	2,67	(08h29, 60)
Travail vers GC	12,96	8,37	17,44	1,66	(08h24, 50)
Autres motifs (groupe 1)	12,96	5,24	10,63	2,43	(08h54, 54)
Autres motifs (groupe 2)	12,96	5,24	10,63	2,43	(10h49, 53)

Tableau 35 : valeurs des paramètres utilisées dans METROPOLIS

Pour adapter l'outil dynamique au réseau de l'IAURIF, une matrice statique de 970 000 déplacements en VP est utilisée. Afin de tenir compte des différents types d'utilisateurs, la matrice statique a fait l'objet d'une segmentation. Le tableau suivant résume la matrice utilisée. Chaque ligne du tableau représente le nombre de déplacements ; par exemple le nombre de déplacements internes à Paris (hors intra-quartiers) est égale à 43 303 déplacements pendant la pointe horaire du matin. Les autres motifs regroupent l'ensemble des déplacements pour la région d'Ile-de-France mais pour des motifs autres que le travail. Cette représentation correspond à la saisie qui devrait être effectuée dans METROPOLIS.

Lors de simulations, les départs vont être dispersés au cours de la période simulée. De ce fait, les nombres de déplacements simulés et observés pendant la période de pointe horaire ne coïncident pas. En fait, pendant la période d'une heure de pointe du matin, il y aura moins de déplacements simulés qu'observés. Par conséquent, il faut augmenter le nombre total de déplacements pour l'ensemble des paires origine - destination. Ainsi le nombre de déplacements pendant l'heure de pointe correspondra au nombre de déplacements observés pendant la même période de pointe horaire.

Type de liaison	Nombre de déplacements
Paris – Paris	34 303
Paris – PC	20 319
Paris – GC	9 151
PC – Paris	23 574
PC – PC	143 738
PC – GC	29 293
GC – Paris	12 326
GC – PC	48 943
GC – GC	211 854
Autres motifs (groupe 1)	200 790
Autres motifs (groupe 2)	235 710
<b>TOTAL</b>	<b>970 000</b>

Tableau 36 : nombre de déplacements selon la répartition IAURIF et MADDIF

### Calage de METROPOLIS

Le processus du calage consiste à effectuer un nombre assez important de simulations et retenir uniquement la simulation qui satisfait le critère de calage défini à l'avance. Le critère de calage consiste à minimiser la différence entre les données observées et les données simulées. Les données observées sont tirées des observations de l'enquête globale transports (EGT) de l'année 1991. Elles représentent la distribution du nombre de déplacements en fonction de la durée du trajet. Etant donné que les paramètres comportementaux sont estimés à partir de l'enquête MADDIF, il ne reste que deux variables à calibrer : la demande et la capacité des voies de circulation.

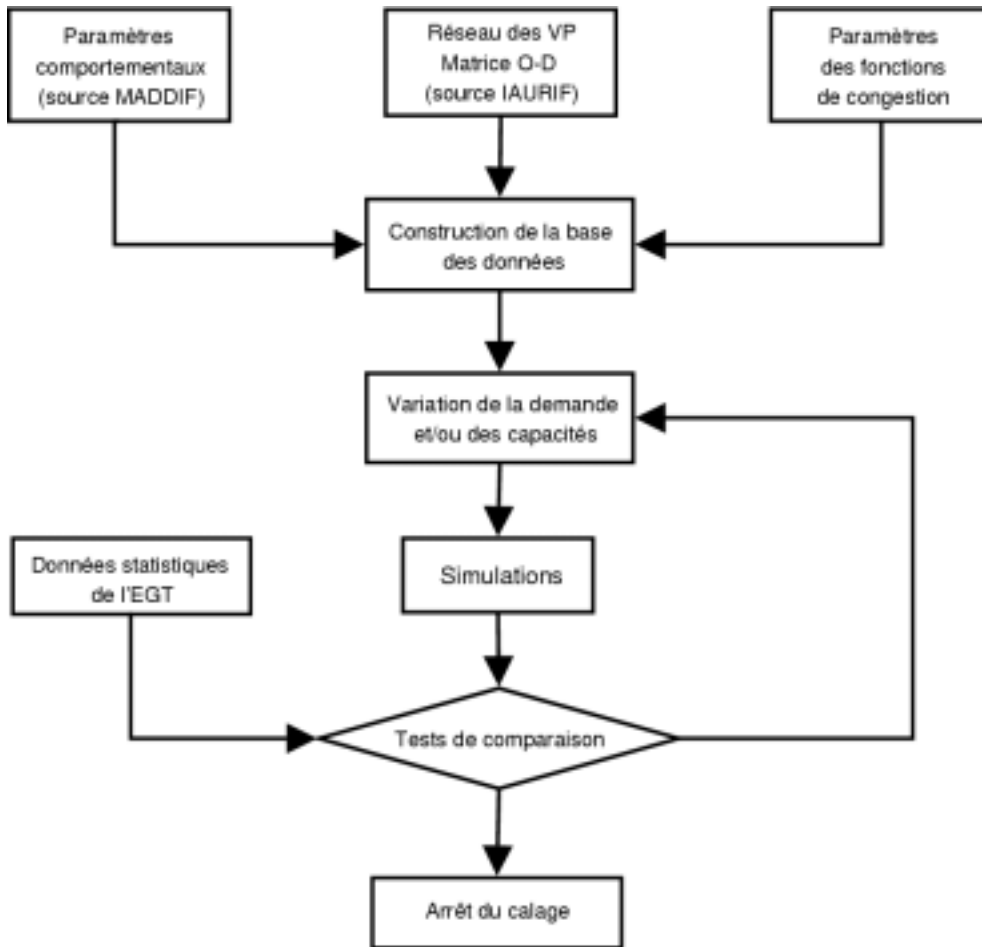


Figure 23 : processus de calage

La Figure 23 représente le processus de calage appliqué à l'adaptation de METROPOLIS à la Région d'Ile-de-France. L'objectif du calage consiste à reproduire d'une part le nombre de déplacements donné par l'IAURIF entre 7h30 et 8h30 et d'autre part les données de l'EGT 91. Pour atteindre cet objectif, deux variables sont modifiées : les capacités des routes et la demande totale. Pour ce faire, une batterie de simulations est effectuée qui fait évoluer d'une part la demande totale et d'autre part les capacités des routes de façon uniforme c'est-à-dire multiplier les capacités des routes par le même facteur. La demande est multipliée par des coefficients allant de 3 à 5 avec un pas de 0,25. Les demandes totales simulées varient de 2910000 (=3x970000) usagers à 4850000 (5x970000) usagers. Les sous-matrices varient également uniformément puisqu'elles sont multipliées par le même coefficient. Quant à l'offre, les capacités varient également de façon uniforme. Elles sont multipliées par un coefficient variant de 0,80 à 1,10 avec un pas de 0,05. En résumé, 63 simulations sont effectuées pour déterminer le scénario de référence. Le scénario final ou calibré nécessite de multiplier la demande par un coefficient égal à 3,75 (soit une matrice totale de 3 637 500 usagers) et la capacité initiale par un coefficient égal à 0,90.

## Résultats du scénario de référence

L'outil de simulation METROPOLIS permet de calculer différentes mesures d'efficacité. Le tableau ci-dessous décrit les résultats du scénario calibré. En moyenne, sur l'ensemble de la matinée, le temps de trajet est égal à 32,50 minutes soit un temps de trajet supplémentaire de 8,16 minutes par rapport au temps de trajet à vide. Le coût du délai (ou de « deshorage ») moyen est égal à 2,45€ soit presque le double du temps de congestion qui est de 1,76€.

<b>Mesures d'efficacité</b>	<b>Valeurs</b>
Temps de parcours [min]	32,50
Temps de parcours à vide [min]	24,34
Temps de parcours de congestion [min]	8,16
Délai d'arrivée précoce [min]	28,28
Délai d'arrivée tardif [min]	19,07
Vitesse moyenne [km/h]	28,90
Part des arrivées précoces [%]	47,25
Part des arrivées tardives [%]	33,15
Coût moyen [€]	9,47
Coût du temps de parcours à vide [€]	5,26
Coût du temps d'attente [€]	1,76
Coût moyen de deshorage [€]	2,45
Volume total des déplacements [million de véh.km]	63,70
Distance moyenne parcourue [km]	17,51
Nombre d'arcs utilisés par individu	17,61
Indice de congestion	28,85

*Tableau 37 : mesures agrégées de l'efficacité du trafic*

## Conclusions

Dans le cadre de ce projet et afin de prendre en compte l'interaction transport – urbanisation, le modèle de simulation METROPOLIS est utilisé pour compléter le modèle d'urbanisation UrbanSim. L'avantage de ce modèle est qu'il rend mieux compte des coûts de déplacement des usagers du réseau. Par conséquent, les décisions de choix de localisation des agents économiques seront plus précis. En effet, les modèles statiques de trafic considèrent uniquement une période restreinte, d'une heure ou deux, d'analyse puis par un jeu de facteurs, les résultats de simulation sont transposés à l'ensemble de la matinée. Cependant, les résultats obtenus à travers les modèles statiques représentent uniquement une moyenne des flux sur les routes et se limitent à la modélisation du choix d'itinéraire dans l'étape d'affectation. Cependant, les travaux d'adaptation de METROPOLIS à la Région d'Ile-de-France, menés au cours de ces dernières années par le groupe TT&R, montre qu'une fraction importante du coût de déplacement est souvent omise. Elle est essentiellement due à la non prise en compte du choix de l'heure de départ dans les modèles statiques. Cette composante du coût total concerne les coûts des délais. Selon les résultats obtenus, ces coûts représentent le double des

coûts de congestion (hors temps de parcours à vide). Ce qui peut influencer de façon significative les choix de localisation des individus et des emplois.

Une partie importante de calage du modèle a été réalisée. Les résultats des simulations du modèle peuvent être considérés comme représentatifs de l'état du trafic en Ile-de-France. Un des avantages de ce modèle est qu'il permet de rendre compte des comportements des usagers de façon désagrégée, autrement dit individu par individu. Les résultats du modèle décrivent, pour chaque usager, l'heure de départ, l'heure d'arrivée, le délai à l'arrivée, le coût total de déplacement et surtout le surplus de chaque usager. En effet, le modèle UrbanSim, en dehors d'autres données sur le comportement des agents économiques, nécessite pour fonctionner d'avoir le surplus des usagers comme donnée d'entrée. METROPOLIS fournit ce surplus des usagers à UrbanSim. Ensuite UrbanSim calcule les nouvelles localisations des habitations et des activités qui les communique à un troisième modèle (nom du module spécifique : Reality). Ce dernier modèle recalcule les nouvelles matrices des flux de déplacements pour les différents paires origine – destination. Puis ces flux de déplacements sont réutilisés par METROPOLIS afin de recalculer les nouvelles valeurs des surplus des usagers. Cette interaction entre les trois modèles se fait chaque année (période de simulation d'UrbanSim) jusqu'à la date de prévision fixée à l'avance par le planificateur. L'interaction entre ces trois est rendue possible grâce à la simplicité de communication des informations entre ces trois modèles. En effet, la raison principale est que les trois modèles utilisent le même langage de gestion des données MySQL. Et grâce à des requêtes SQL et la structure des bases de données MySQL, il est possible de transférer des informations de la base d'un modèle à un autre.



## ANNEXE 5 : GLOSSAIRE

Certains mots et expressions sont en anglais pour faciliter la lecture de la documentation d'UrbanSim.

### *Adaptateur des « Logsommes »*

Ce processus se charge de l'adaptation du format des résultats de METROPOLIS au format UrbanSim selon l'équation:

$$SC_{ij} = \mu^d \log \int_0^t \exp(-C_{ij}(u) / \mu^d) du$$

$C_{ij}(u)$  étant le coût généralisé pour le trajet de  $i$  à  $j$  partant en  $u$  (l'heure de départ).  
 $\mu^d$  est le paramètre d'hétérogénéité dans le modèle logit de choix de l'heure de départ.  
 $t$  est la durée de simulation (24 heures par exemple).

### *Baseyear Data*

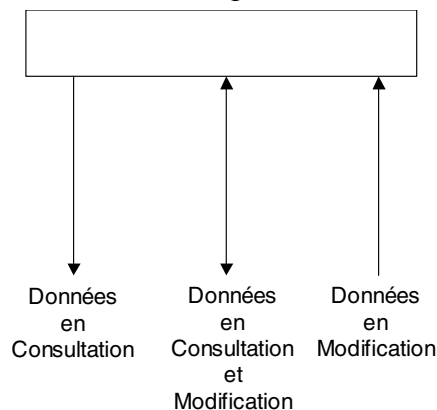
Données socio-économiques décrivant l'ensemble à « l'année zéro »

### *Déplacement*

Suite de trajets, un trajet est défini par un mode unique. Par exemple, pour aller de Paris 15<sup>ème</sup> à Versailles, il y a un déplacement composé de 2 trajets (un en métro puis un en train).

### *Données*

Ensemble structuré de tables et paramètres, elles sont représentées par



### *METROPOLIS*

Modèle dynamique de transport permettant entre autres le calcul des temps de parcours sur un réseau. Les données d'entrée seront fournies par les résultats du modèle de demande à 3 étapes, lui-même alimenté par les résultats d'UrbanSim. Les résultats sont ajustés par le processus « Logsomme » et réintégrés à UrbanSim par le « ScenarioRecreator »

### ***Modèle à 3 étapes***

Modèle mathématique permettant sur la base de données socio-économiques de reconstituer la demande de déplacements et de simuler cette demande dans le futur. Il est composé des trois premières étapes du modèle à 5 étapes décrit au chapitre 5, soit :

- Calcul des émissions et des attractions de déplacements par zone et par motif de déplacement
- Distribution des déplacements entre zones par motif de déplacement
- Répartition de ces déplacements par mode de transport

### ***MySQL***

Base de données utilisée par UrbanSim et METROPOLIS, robuste et gratuite pour des applications non commerciales.

### ***Processus***

Ensemble des algorithmes réalisant une tâche bien définie; ils sont représentés par



### **Scénario**

Ensemble de données et paramètres nécessaires à une exécution du moteur UrbanSim : années de début et de fin, données socio-économiques, paramètres de mobilité, période concernée, réseau de transports, matrice des temps de parcours en transports en commun, etc

### **Scenario Recreator**

Module d'UrbanSim servant à combiner les résultats d'une simulation, les résultats du modèle de trafic et les scénarios, pour générer un nouveau scénario.

### **UrbanSim**

Modèle d'urbanisation, composé d'une base de données MySQL, d'un moteur de simulation, d'un module de vérification de la cohérence des données et du module ScenarioRecreator

### **Windev**

Outil de développement de programmes Windows, caractérisé par une interface graphique de qualité, un déploiement facile et des temps de calcul compatibles avec les caractéristiques du projet.

## **Définition des outils de l'urbanisme local :**

### **COS**

Coefficient d'occupation du sol. Il correspond au rapport entre la surface habitable hors œuvre nette et la surface du terrain. Un COS de 1 permet de construire un bâtiment d'une superficie équivalente à celle du terrain : sa valeur est indiquée dans le POS. Pour comparaison, le COS de Paris est de 3.

### **OPAH**

Une Opération Programmée d'Amélioration de l'Habitat est un programme concerté entre l'Etat, l'Agence Nationale pour l'Amélioration de l'Habitat (ANAH) et une Ville. Son objectif est de soutenir les travaux de réhabilitation du parc de logements privés anciens par des aides financières.

Ces aides sont attribuées pour la réalisation de travaux permettant d'améliorer le confort ou l'esthétisme des bâtiments, dans les parties privatives comme dans les parties communes.

L'intérêt pour une Ville est de sauvegarder son patrimoine et une certaine qualité architecturale et pour les habitants, d'améliorer leurs conditions de vie à moindres frais.

### **PADD**

Le Projet d'Aménagement et de Développement Durable s'ajoute désormais aux pièces qui constituaient le POS et c'est un élément central du PLU.

Document prospectif, le PADD expose en effet les intentions d'une commune pour les années à venir : orientations générales d'urbanisme, conditions d'aménagement et d'équipement ... Il se traduit par un règlement obligatoire pour l'ensemble de la commune (dont la lettre doit être respectée) et des orientations d'aménagement facultatives concernant certains quartiers, l'aménagement d'une ZAC par exemple, ou des politiques sectorielles, traitement des entrées de ville par exemple (dont l'esprit doit être respecté).

### **PLU**

Dans le cadre de la loi SRU, le Plan Local d'Urbanisme se substitue au POS.

Comme le POS, le PLU définit les règles de constructibilité et précise le droit des sols. Il reste donc un document essentiel à consulter avant d'acheter un bien immobilier ou de lancer des travaux importants.

Mais le PLU est un document de planification urbaine plus ambitieux et plus complet que le POS : plus qu'un simple instrument de réglementation des sols, c'est un outil au service d'un projet de territoire.

Elaboré en fonction d'un diagnostic d'ensemble et à la suite d'une réflexion sur le devenir d'une commune, il exprime un projet global pour les années à venir, définit les grandes orientations d'aménagement urbain et les actions qui en découleront.

Il permet ainsi à une commune de mieux exprimer son projet de développement, avec une vision globale, à long terme et pour l'ensemble de son territoire. Le PLU se compose principalement :

- d'un rapport qui expose le diagnostic, analyse l'environnement de la commune et précise les principaux besoins actuels et futurs.
- du PADD, qui présente le projet de la Ville pour les années à venir.
- des orientations facultatives d'aménagement.
- du règlement et des documents graphiques fixant les règles d'utilisation du sol.

## **POS**

Le Plan d'Occupation des Sols était, avant la loi SRU, le document d'urbanisme qui réglementait l'utilisation des sols. Il permettait à une commune de maîtriser son urbanisation en servant de base pour la délivrance des permis de construire et déclarations de travaux.

Le POS a aujourd'hui cédé la place au PLU ; il ne s'agit pas d'un simple changement de nom, mais bien d'une nouvelle approche de l'urbanisme local.

## **SHON**

Surface hors œuvre nette. Elle comprend toute la superficie du bâtiment hormis les combles, sous-sols, balcons, loggias, etc. La surface totale est appelée « surface hors œuvre brute ».

## **SRU**

La loi Solidarité et Renouvellement Urbain du 13 décembre 2000 a refondu en profondeur le droit de l'urbanisme et sa pratique par les communes. Son article 1 (article L-121-1 du Code de l'Urbanisme énonce les principes que doivent désormais respecter les documents d'urbanisme : équilibre, diversité des fonctions urbaines et de mixité sociale mais aussi respect de l'environnement.

Cette loi induit donc une nouvelle conception de la planification urbaine : il s'agit désormais d'exprimer un projet global intégrant la notion de développement durable à côté de celle d'aménagement urbain.

## **ZAC**

Les Zones d'Aménagement Concerté, créées le plus souvent à la suite du départ d'entreprises, permettent d'aménager de nouveaux quartiers mariant harmonieusement activité économique (pour préserver l'emploi et les ressources de taxe professionnelle), habitat (social ou en accession à la propriété), équipements publics et espaces verts. Contrairement au POS, le PLU s'applique également aux ZAC. Les communes qui mettent en place les ZAC jouent ainsi un rôle d'accompagnement fort dans l'urbanisation.

## **Zone NC ou ND**

Terrains agricoles ou à protéger et donc inconstructibles

## **Zone NA**

Première étape vers l'urbanisation ; elle peut bénéficier d'un COS et passer ensuite en U ou en ZAC, c'est-à-dire en zone constructible.

## **Zone UR**

Terrains où le droit à construire est temporairement limité, et placés par la commune en réserve foncière.

## **Zone U**

Zone urbaine