



HAL
open science

Prévision de la demande de transport

Patrick Bonnel

► **To cite this version:**

Patrick Bonnel. Prévision de la demande de transport. Economies et finances. Université Lumière - Lyon II, 2002. tel-00268919

HAL Id: tel-00268919

<https://theses.hal.science/tel-00268919v1>

Submitted on 1 Apr 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prévision de la demande de transport

**Rapport présenté en vue de l'obtention du diplôme d'habilitation à
diriger les recherches**

**Université Lumière Lyon 2, Faculté de Sciences Economiques et de
Gestion**

**Patrick BONNEL
Laboratoire d'Economie des Transports
ENTPE, Université Lumière Lyon 2, CNRS
Décembre 2001**

Sommaire

Sommaire	3
Remerciements	5
Introduction	7
Chapitre 1 – Modélisation et planification	11
1. POURQUOI DOIT-ON PREVOIR ?	11
2. HISTORIQUE DE LA MODELISATION	15
3. QUELS MODELES ? POUR QUELS DOMAINES D’UTILISATION ?	25
Chapitre 2 - Cadrage méthodologique	33
1. SCHEMA DE LA PROCEDURE DE PREVISION	33
2. LE MODELE : RESULTAT D’UN DOUBLE PROCESSUS DE SIMPLIFICATION	42
3. OPERATIONNALITE ET PROBLEMATIQUE DU MODELE	45
4. EXEMPLE : CALAGE DU MODELE PRIX-TEMPS	51
5. CONCLUSION	64
Chapitre 3 - Analyse de la mobilité	67
1. MOBILITE : UNE DEFINITION EN EVOLUTION	67
2. DESCRIPTION DE LA MOBILITE ET DE SON EVOLUTION	69
3. TENDANCES STRUCTURELLES D’EVOLUTION DE LA MOBILITE : EXEMPLE D’ANALYSE, EXEMPLE DE QUANTIFICATION	75
4. CONCLUSION : LES PIEGES DE L’ANALYSE DE LA MOBILITE	96
Chapitre 4 - Production de données	99
1. POURQUOI A-T-ON BESOIN DE DONNEES ?	99
2. METHODOLOGIE DE PRODUCTION DES DONNEES	101
3. PRINCIPALES DONNEES DISPONIBLES EN MILIEU URBAIN EN FRANCE	116
4. LIMITES DU DISPOSITIF FRANÇAIS D’ENQUETES SUR LES DEPLACEMENTS URBAINS POUR LA MODELISATION	125
5. LES ENQUETES DE PREFERENCES DECLAREES	134
6. EXEMPLE D’APPLICATION DU SCHEMA METHODOLOGIQUE DE PRODUCTION DES DONNEES : EXEMPLE DE LA CONSTRUCTION DU CADRE CONCEPTUEL	139
7. CONCLUSION	142
Chapitre 5 - Fondements théoriques des modèles de prévision de la demande de déplacements	145
1. QUELQUES ELEMENTS DE THEORIE MICRO-ECONOMIQUE NEO-CLASSIQUE DU CHOIX DU CONSOMMATEUR	145
2. APPLICATION A LA SPHERE DES TRANSPORTS	149
3. CLASSIFICATION DES « MODELES A QUATRE ETAPES »	160
4. CONCLUSION	164

Chapitre 6 – Génération de la demande de déplacements	165
1. DEFINITIONS ET CHOIX DE L'UNITE DE MESURE	165
2. PRINCIPALES FAMILLES DE MODELES DE GENERATION	173
3. PRINCIPALES LIMITES DE L'ETAPE DE GENERATION	183
Chapitre 7 – Distribution des déplacements	191
1. DEFINITIONS ET CONCEPTS DE BASE	191
2. PRINCIPALES FAMILLES DE MODELES DE DISTRIBUTION	194
3. PRINCIPALES LIMITES DE L'ETAPE DE DISTRIBUTION	213
Chapitre 8 – Répartition modale – approche agrégée	219
1. DEFINITION ET VARIABLES EXPLICATIVES	219
2. PRINCIPALES FAMILLES DE MODELES DE REPARTITION MODALE	226
3. PRINCIPALES LIMITES DE L'ETAPE DE REPARTITION MODALE	240
Chapitre 9 - Modèles désagrégés de choix discrets	245
1. FONDEMENTS THEORIQUES	246
2. PRINCIPALES FAMILLES DE MODELES DESAGREGES	248
3. ESTIMATION DU MODELE DESAGREGE	270
4. SPECIFICATION DES MODELES DESAGREGES	286
5. PREVISION ET AGREGATION	293
6. PRINCIPALES LIMITES DE L'APPROCHE DESAGREGEE	297
Chapitre 10 – Affectation, zonage et codification des réseaux	303
1. POSITIONNEMENT DE L'AFFECTION DANS LA CONSTRUCTION ET L'UTILISATION D'UN MODELE A QUATRE ETAPES	303
2. PERIMETRE ET ZONAGE DE L'AIRE D'ETUDE	307
3. CODIFICATION DES RESEAUX	312
4. CONCEPTS DE BASE ET ETAPES PRELIMINAIRES	319
5. PRINCIPALES FAMILLES DE MODELES D'AFFECTION	328
6. PRINCIPALES LIMITES DE L'ETAPE D'AFFECTION	355
Annexe 1 - Données nécessaires au calage du modèle « Prix-temps » et corrigé	362
1. TRAFICS EN 1996	362
2. TEMPS DE PARCOURS EN MINUTES EN 1996	363
3. TARIFS DE BASE AUTOMNE-HIVER 1996 (CONVERTI EN EQUIVALENT EUROS 96)	363
4. CORRIGE DE L'EXERCICE DE CALAGE DU MODELE PRIX-TEMPS - ENSEMBLE DES DONNEES	364
Bibliographie	371
Liste des tableaux	393
Liste des illustrations	395
Liste des abréviations	399
Table des matières	401

Remerciements

A ma femme et mes filles à qui j'ai « volé » du temps.

A Eric Tabourin qui s'en est allé prématurément en nous laissant une partie de sa richesse que je me suis permis d'utiliser dans cet ouvrage.

A Alain Bonnafous qui a accepté de diriger cette habilitation et qui m'a prodigué des conseils et remarques toujours précieux.

Aux membres du jury d'habilitation Alain Bonnafous, Robert Chapleau, Yves Crozet, Emile Quinet, Philippe Toint, qui ont bien voulu prendre de leur temps pour lire ce document et se déplacer parfois de loin (et même de très loin !) pour participer au jury.

A ceux qui m'ont aidé dans cette rédaction par leurs remarques et critiques et notamment, Jimmy Armoogum, Alain Bonnafous, Cécile Clément, Lionel Clément, Thierry Du Crest, Hubert Julien, Jean-Loup Madre, Dominique Mignot, Pierre Odent, Charles Raux.

A tous mes étudiants qui m'ont fait évoluer et progresser dans la façon de présenter la modélisation.

A tous mes collègues au sein du Laboratoire d'Economie des Transports ou dans d'autres organismes en France ou à l'étranger avec qui j'ai eu le plaisir de travailler.

A l'ADEME (l'Agence De l'Environnement et de la Maîtrise de l'Energie) et plus généralement au PREDIT (Programme national de recherche et d'innovation dans les transports terrestres) qui grâce à leur financement m'ont permis de conclure cet ouvrage.

Comme il est d'usage, je reste toutefois seul responsable des erreurs, des omissions ou de certains positionnements exprimés dans cet ouvrage que je sais ne pas être partagés par tous.

Introduction

Les prémices de la modélisation de la demande de transport sont déjà anciens. On peut situer les premiers modèles à la fin du dix-neuvième siècle avec les modèles gravitaires appliqués à la modélisation de la demande de déplacements entre les agglomérations. En revanche, le développement des modèles urbains est beaucoup plus récent. Il se situe dans le courant des années cinquante aux Etats-Unis pour répondre aux besoins de développement de l'automobile. Le transfert en France s'est organisé à la fin de cette décennie et au début des années soixante. De nombreux développements ont ensuite vu le jour que ce soit sur le plan théorique ou sur le plan des outils mis à la disposition des modélisateurs. Le développement des modèles désagrégés fait partie de ces avancées majeures. Si les avis ne sont pas unanimes sur leur apport en termes d'opérationnalité des modèles (nous précisons cette notion chapitre 2, p. 45), leur contribution est indéniable sur le plan de la formalisation des fondements théoriques de la modélisation de la demande de déplacements.

La France a été présente lors des premiers développements, mais force est de constater que cette présence s'est fortement réduite par la suite mis à part quelques contributions (notamment, le modèle prix-temps dont l'usage est beaucoup plus répandu que dans les pays anglo-saxons). Les outils mis en place, en France, dans les années soixante-dix ont connu peu d'évolutions (sauf en ce qui concerne l'étape d'affectation de la demande) et sont encore assez largement utilisés aujourd'hui. Ces outils ont bien répondu aux problématiques de l'époque, centrées sur le dimensionnement et l'évaluation de nouvelles infrastructures de transport. En revanche, ils s'avèrent bien moins adaptés à celles d'aujourd'hui privilégiant une meilleure gestion des infrastructures existantes dans une perspective de développement durable. Il s'agit davantage d'évaluer des politiques de déplacements que des infrastructures de transport.

Cette évolution des problématiques nécessite également une évolution des outils de modélisation. L'ambition de cet ouvrage est d'y contribuer en organisant l'état des connaissances sur la prévision de la demande de déplacements urbains. Il a été initié en tant que support de cours pour l'enseignement de la modélisation de la demande à l'ENTPE (Ecole Nationale des Travaux Publics de l'Etat) et pour le DEA d'Economie des Transports et le DESS Transports Urbains et Régionaux de Personnes communs à l'Université Lumière Lyon 2 et à l'ENTPE. Sa finalisation a été réalisée grâce à un financement de l'ADEME dans le cadre du programme de recherche du PREDIT sur la modélisation de la demande de transport. Ce financement doit également déboucher sur la confection d'un didacticiel de formation à la modélisation, réalisé avec les sociétés MVA et IMTRANS, qui doit être achevé dans le courant de l'année 2002.

Avant de présenter le plan de l'ouvrage, il nous semble nécessaire de délimiter son champ d'application. Comme le titre l'indique, il est centré sur la modélisation de la demande de déplacements de personnes en milieu urbain. Nous n'abordons pas la modélisation des transports de marchandises, même si certains outils peuvent être aisément transposés à ce champ de l'économie des transports. De même, la modélisation de la demande interurbaine est peu abordée. Les liens entre l'urbain et l'interurbain étant plus étroits, nous développons certains modèles dont l'application est plus courante en non urbain qu'en urbain, comme la méthode d'Abraham pour l'affectation de la demande par exemple. Toutefois, la portée de cet ouvrage déborde largement ce seul champ de l'urbain et des déplacements de personnes. Les développements méthodologiques ou théoriques (notamment chapitres 2 et 5) ou la modélisation désagrégée (chapitre 9) sont abordés de manière suffisamment large pour que les enseignements puissent profiter à l'ensemble de la modélisation transport.

Une dernière précision s'impose. L'ouvrage couvre la modélisation de moyen ou long terme. En revanche, nous n'abordons pas la simulation de l'écoulement des flux qui relève généralement d'approches de très court terme, voire de temps réel. Ce champ de la modélisation s'appuie davantage sur l'ingénierie du trafic que sur l'économie des transports. Même si l'on peut trouver des outils communs, les fondements théoriques des modèles sont souvent différents.

Le premier chapitre est conçu comme une introduction. Nous donnons quelques éléments justifiant l'importance de disposer d'outils de prévision. Nous présentons ensuite un historique de la modélisation en parallèle avec un historique de la planification. Ces historiques ne se veulent pas exhaustifs, mais illustrent la nécessité d'une adéquation de la modélisation et de la production des données aux objectifs de planification. Nous délimitons enfin le domaine sur lequel est centré cet ouvrage.

Le plan de l'ouvrage s'appuie sur le schéma 2.1 de procédure de prévision (p. 34). Ce schéma permet d'insister sur la place du modèle dans le processus de prévision. Le modèle est un outil au service de la prévision. Il n'est donc pas la prévision à lui seul. Trop souvent le modèle a un rôle de boîte noire du fait de sa complexité apparente (qui tient en partie aux équations mathématiques dont la lecture n'est pas aisée pour tout le monde, aux techniques d'estimations parfois complexes, aux logiciels dont la maîtrise n'est pas immédiate...). Nous verrons dans les différents chapitres que le principe des modèles est souvent très simple, même si leur mise en œuvre ne l'est pas forcément. Nous avons donc voulu présenter tout d'abord ce cadrage méthodologique du processus de prévision (chapitre 2). Il nous permet d'insister sur la nécessité d'analyser l'objet d'étude. Cette analyse passe par une identification des objectifs de la prévision suivi de la construction d'un schéma explicatif de l'objet d'étude, car un modèle s'appuie forcément sur un schéma de causalité. Si ce schéma était plus souvent explicite, il faciliterait sûrement la compréhension de la modélisation. Un modèle nécessite également des données pour pouvoir être calé, validé, puis utilisé en prévision. Enfin, nous exposons la problématique des modèles qui résulte de la contradiction entre les trois conditions de l'opérationnalité des modèles : pertinence, cohérence et mesurabilité.

La nécessité d'analyser l'objet d'étude pour construire le schéma de causalité nous conduit vers l'analyse de la mobilité urbaine (chapitre 3). Nous présentons les grandes tendances d'évolution de cette mobilité avant d'exposer deux exemples d'analyse qui permettent d'illustrer les difficultés de cette démarche.

La nécessité d'alimenter en données le modèle amène à examiner les méthodes de production des données (chapitre 4). Nous abordons quelques éléments méthodologiques relatifs à la production des données. Nous passons ensuite en revue les différentes productions de données fréquemment disponibles, en France, en milieu urbain. Ce chapitre nous offre l'occasion d'analyser cette production sous un regard critique dans une perspective de modélisation. Les données disponibles ne suffisent pas toujours à l'alimentation des modèles nécessitant parfois le recours aux enquêtes de préférences déclarées.

Nous pouvons alors entrer dans la modélisation proprement dite avec un premier chapitre qui expose les fondements théoriques des modèles de prévision de la demande (chapitre 5). Ces fondements reposent sur la théorie micro-économique du comportement du consommateur maximisant son utilité. Ce chapitre permet également de proposer une première classification générique des modèles entrant dans le champ de cet ouvrage. La plupart de ces modèles repose sur une chaîne à quatre étapes qui organise la suite des chapitres :

- génération de la demande de déplacements (chapitre 6) ;
- distribution des déplacements (chapitre 7) ;
- répartition modale – approche agrégée (chapitre 8) ;
- affectation des déplacements (chapitre 10).

Au milieu des chapitres 8 et 10, nous avons intercalé la présentation des modèles désagrégés (chapitre 9). Ce plan peut apparaître surprenant. Il se justifie d'un point de vue historique. Les modèles désagrégés de choix discrets ont été développés pour l'étape de répartition modale. Cette étape reste le domaine privilégié d'application de ces modèles, même si aujourd'hui l'approche désagrégée peut être développée pour l'ensemble des étapes.

Pour faciliter la lecture du manuel, le plan est identique pour chacun des chapitres 6 à 10. Il s'organise autour de trois sections :

- définitions et concepts de base. Cette section précise un certain nombre de définitions relatives aux données ou méthodes utilisées par les modèles de l'étape considérée ;
- principales familles de modèles. Nous présentons les principaux modèles disponibles pour traiter l'étape considérée. Cette présentation s'organise autour des grandes familles de modèles. Pour chacune d'elles, nous exposons les fondements de la méthode, la formalisation mathématique, parfois la démonstration et le plus souvent la procédure d'estimation ;
- principales limites. Nous avons cherché à regrouper les principales limites des modèles à l'issue de chaque étape. Certaines sont spécifiques à une étape, mais d'autres se retrouvent à chaque étape, ce qui conduit inévitablement à quelques redites. Il ne s'agit pas de décrire les modèles disponibles, mais plutôt d'attirer l'attention sur les hypothèses (trop souvent implicites) nécessaires à l'utilisation des modèles.

Pour les deux derniers chapitres, nous avons partiellement enfreint cette organisation en complétant la présentation avec de sections supplémentaires. Pour l'approche désagrégée, l'estimation des modèles nécessite des outils particuliers que nous analysons. De même, la spécification des modèles n'est pas aussi immédiate que pour les autres étapes justifiant de focaliser notre attention sur cette question. Enfin, l'agrégation en prévision est un problème spécifique aux modèles désagrégés. Pour l'affectation, avant de décrire les modèles, il nous a semblé utile d'analyser le choix du zonage et la codification des réseaux.

A la lecture de ce plan, on pourra être surpris de ne pas rentrer plus rapidement dans les modèles qui ne sont réellement abordés qu'à partir du chapitre 6. Ce choix est délibéré, car il nous semble important de rappeler que la prévision ne se limite pas à l'utilisation d'un outil « le modèle ». La modélisation s'inscrit dans un processus de prévision que nous détaillons dans le chapitre 2 et qui constitue probablement l'originalité de cet ouvrage comparativement aux ouvrages principalement anglo-saxons existant sur le marché. Ce processus comprend l'analyse des données (chapitre 3) et la production des données (chapitre 4). Enfin les modèles s'appuient sur des fondements théoriques qu'il est important de rappeler (chapitre 5). Toutefois, le lecteur s'estimant suffisamment familier avec ce cadre général pourra aborder directement cet ouvrage avec le chapitre 6.

Un tel ouvrage se lit assez rarement de bout en bout, mais plutôt par chapitre ou partie de chapitre. Nous avons cherché à répondre à ce mode de lecture en le concevant plutôt comme un manuel. Chaque chapitre peut se lire indépendamment des autres. Au sein de chaque chapitre, nous avons cherché à résumer l'essentiel des connaissances. Pour en faciliter la lecture, nous avons tenté d'organiser de la manière la plus lisible possible le plan des différents chapitres. La conception par chapitre « autonome » nous oblige toutefois à de fréquents renvois vers les autres chapitres. Parfois, cela pourra donner une impression de redites. Même si nous avons cherché à les réduire, elles sont inévitables.

Le manuel comprend un grand nombre d'équations. Elles facilitent la lecture et surtout la compréhension pour le lecteur familier avec ce langage. En revanche, elles peuvent rebuter le lecteur qui l'est moins. Nous avons donc cherché dans la mesure possible à permettre une lecture de ce manuel qui se limiterait au texte. Certains chapitres se prêtent assez facilement à l'exercice, mais d'autres moins (notamment le chapitre 9). Il nous semble toutefois, que la lecture du manuel reste

accessible au lecteur peu familier avec les mathématiques. Il lui suffira de survoler les équations et de ne pas aborder les démonstrations et la présentation des algorithmes au chapitre 10.

Enfin, pour le lecteur souhaitant aller au-delà des développements de cet ouvrage, nous proposons à la fin de chaque chapitre un ou plusieurs ouvrages permettant d'approfondir la thématique développée dans le chapitre.

Chapitre 1 – Modélisation et planification

Ce premier chapitre permet de situer notre sujet. Nous débutons par la présentation des multiples raisons qui justifient la nécessité de prévoir la demande de déplacements (section 1). Nous décrivons ensuite rapidement un historique de la modélisation afin d'illustrer l'évolution des outils en fonction de l'évolution des objectifs de la planification (section 2). Nous précisons alors la famille de modèles que nous prenons en considération et identifions son domaine principal de prédilection, à partir d'une approche systémique de l'urbain. Celle-ci, croisée avec les dimensions spatiales et temporelles, offre une grille de lecture de la nature des interactions prises en compte dans le modèle choisi (section 3).

1. POURQUOI DOIT-ON PREVOIR ?

Au delà de la nécessité de s'adapter le plus finement possible aux besoins, plusieurs exigences économiques et sociales rendent cette prévision indispensable, avec des enjeux liés aux caractéristiques des systèmes de transport :

- le coût très élevé des investissements, qui doivent être évalués avant d'être décidés ;
- les délais importants de construction et la durée de vie des investissements, qui inscrivent ces évaluations dans la durée ;
- l'anticipation des évolutions et des dysfonctionnements qu'il convient de prévenir ;
- l'accompagnement par les transports du développement économique et social ;
- l'évolution rapide du niveau de certains indicateurs de mobilité.

1.1. Coût des investissements

La comptabilité nationale nous permet de situer rapidement la place de l'investissement transport. Le coût des investissements dans la branche transport est très élevé. En 1977, ils représentent environ 7 % de l'investissement de l'ensemble des branches marchandes de la Nation, alors que la valeur ajoutée de la branche transport ne représente qu'un peu plus de 4 % du Produit Intérieur Brut Marchand (Bonnell et al., 2000, à partir des comptes transports en 1997 (INSEE, SES, 1998)). L'investissement représente en moyenne 17 % du Produit Intérieur Brut Marchand alors que dans la branche transport, il contribue pour 29 % à la valeur ajoutée de la branche (chiffres pour 1997, Bonnell et al., 2000, à partir des comptes transports en 1997 (INSEE, SES, 1998)). Par ailleurs, les investissements dans la filière transport représentent une part importante des investissements réalisés tant par l'Etat que par les collectivités locales. La puissance publique consacre près de 20 % de ses dépenses d'investissements au domaine des transports (14 % pour l'Etat et 20 % pour les collectivités locales en 1997 (Bonnell et al., 2000)).

De plus, les investissements en infrastructures ont le plus souvent des coûts unitaires très importants (les chiffres suivants représentent des ordres de grandeur en euros 2002 pour l'année 2002) :

- chacune des lignes de T.G.V. (Sud-Est, Sud-Ouest, Nord-Est) représente un investissement supérieur à 3 milliards d'euros ;
- le tunnel sous la Manche a coûté de l'ordre d'une quinzaine de milliards d'euros ;
- une autoroute coûte de 7 millions d'euros au kilomètre en rase campagne à plus de cent millions d'euros en milieu urbain dense. Le coût peut même monter à 150 millions d'euros lorsqu'il s'agit d'enterrer l'infrastructure ;
- une ligne de métro coûte de l'ordre de 50 à 80 millions d'euros du kilomètre.

Si la puissance publique se doit de choisir les investissements les plus rentables d'un point de vue économique et social, la question de la rentabilité est encore plus cruciale pour les investisseurs privés. Ceux-ci commencent à être davantage présents dans les investissements d'infrastructures de transports et sont sollicités par les collectivités pour assurer un partenariat dans le financement de certains ouvrages. Dans ce cas, la rentabilité repose essentiellement sur le trafic attendu. Les modèles de prévision de la demande de déplacements représentent alors un enjeu fondamental. La prévision nécessite de bien maîtriser l'outil utilisé (le modèle de prévision), mais également l'environnement du projet, pour éviter des déconvenues comme ORLYVAL ou le tunnel sous la Manche.

1.2. Délai de construction et durée de vie des investissements importants

La plupart des investissements de transport ont des durées de construction de plusieurs années :

- la décision de construire la ligne D du métro de Lyon, qui a fait suite à de nombreuses études techniques, a été prise en 1981, pour une mise en service en 1991 ;
- la décision de construire le T.G.V. Atlantique a été prise en 1982. Sa mise en service a eu lieu en 1989, pour la branche Ouest et en 1990 pour la branche Sud-Ouest.

Certes ces exemples ne représentent pas des records de rapidité de construction. Mais dans tous les projets importants d'infrastructures, plusieurs années sont nécessaires pour achever les constructions. De plus, pour d'autres projets, les délais sont encore plus longs en raison de la réalisation par étapes de l'infrastructure, comme pour les autoroutes ou certains contournements autoroutiers d'agglomération.

Enfin, ces durées doivent encore être allongées des délais de réalisation des études. Il faut donc pouvoir disposer de méthodes de prévision suffisamment précises pour que l'infrastructure projetée ne soit pas sous-utilisée lors de sa mise en service ou, à l'inverse, déjà saturée.

Par ailleurs, la durée d'amortissement de la plupart des investissements d'infrastructures est de l'ordre de 15 à 25 ans (ferroviaire, autoroute, aéroport, voie navigable). La rentabilité des investissements est également évaluée sur des durées comparables. Mais, au-delà de ces calculs économiques, la durée de vie de ces investissements est bien plus grande.

Le réseau ferroviaire actuel, mis à part les T.G.V., était déjà en place vers 1890. Son évolution a surtout constitué en ... des suppressions de lignes, des rectifications de tracé pour accroître les vitesses, des doublages de lignes, ou des électrifications de lignes. La même observation peut être faite pour les voies navigables, les réseaux de routes nationales et départementales, même si des accroissements de capacité ou de performance ont souvent été nécessaires.

Les investissements opérés dans le domaine des transports, et tout particulièrement pour les infrastructures, engagent donc le très long terme. Il est donc nécessaire de pouvoir établir des prévisions sur des termes lointains pour ces gros investissements.

1.3. Anticipation des évolutions et des dysfonctionnements

Les déplacements urbains sont de plus en plus considérés comme un système complexe interagissant avec d'autres systèmes, comme le système urbain, le système économique.... Certaines évolutions ou certains dysfonctionnements peuvent alors ne pas être apparents sans un effort de simulation parfois complexe. Le modèle QUINQUIN de simulation des besoins de financement en transports collectifs permet d'illustrer ce type de situation (Bouf, 1989 ; Tabourin, 1989). Au milieu des années 80, il a permis d'attirer l'attention des autorités organisatrices sur les risques de dérives financières des réseaux de transports collectifs, alors que ceux-ci étaient dans une situation de relative aisance financière avec la manne que constituait le versement transport. Il a de plus mis en évidence les leviers de commande du système de financement et leur incidence sur l'évolution des besoins.

C'est également le rôle des modèles de prévision de la demande de déplacements, qui doivent être en mesure de prévoir les effets de l'évolution des caractéristiques d'un maillon sur l'ensemble d'un réseau. Dans certains cas, des reports, même faibles, peuvent détériorer fortement une situation fluide, y compris sur des zones a priori non concernées par l'aménagement envisagé.

La montée en puissance des préoccupations environnementales en fournit un autre exemple. Les progrès technologiques font dire à certains que les émissions polluantes des véhicules thermiques ne sont plus un vrai problème. La situation est toutefois plus complexe car de multiples facteurs interviennent lors du calcul des émissions (progrès technologiques, croissance de la mobilité et surtout des distances parcourues, composition du parc, politiques de transport, conditions de circulation... (Nicolas, 1996)). Ces facteurs interagissent entre eux, rendant les simulations assez complexes. La prise en compte de ces dynamiques permet alors d'éclairer les décideurs sur les évolutions possibles des différents polluants et tout particulièrement de ceux contribuant à l'effet de serre. Ces outils sont indispensables si les pays européens veulent respecter les engagements internationaux qu'ils ont pris, notamment à Kyoto.

1.4. Accompagnement par les transports du développement économique et social

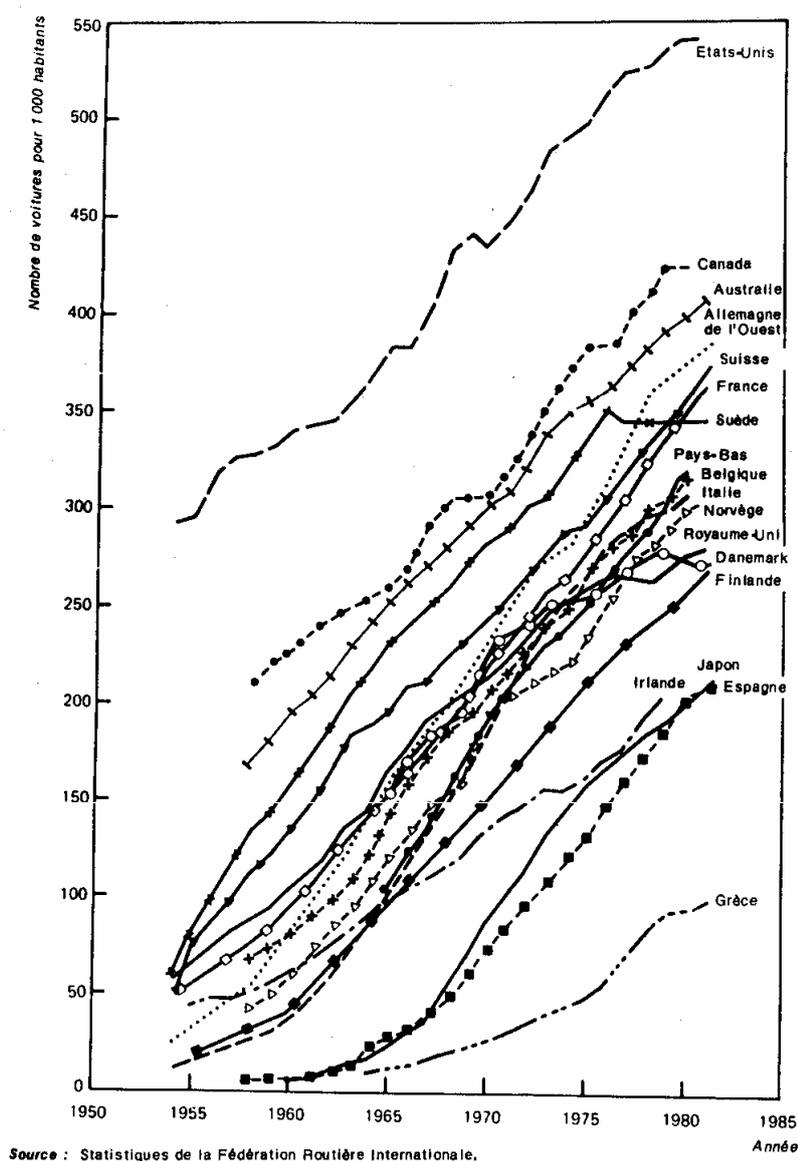
Tous les indicateurs macro-économiques mettent en évidence la forte contribution des transports à l'économie nationale et plus encore la forte interdépendance entre les transports et de nombreux secteurs économiques. Il est banal de dire que les transports ne sont pas une fin en soi, mais permettent de réaliser des activités ou de relier entre eux des lieux de production ou de distribution. On en mesure rapidement les conséquences lors des grèves dans les transports. Les grèves dans les transports collectifs urbains perturbent très fortement le fonctionnement de la ville, même lorsque la part de la voiture est très élevée. Dans les transports routiers de marchandises, elles peuvent mettre rapidement des entreprises en chômage technique avec le développement de la gestion en flux tendus.

Les infrastructures ont également contribué à façonner la répartition des activités sur le territoire national ou dans une agglomération. Il ne faut pas pour autant tomber dans le mythe des effets structurants des infrastructures de transport qui amèneraient seules le développement économique. En revanche, elles jouent un rôle ou elles accompagnent des mutations, ainsi que le montrent Offner et Plassard, dans leurs nombreux ouvrages sur la question (notamment Offner 1985 ; INRETS, 1982, 1985 ; Plassard, 1988 ; Bonnel, Thibaud, 1991).

1.5. Evolution rapide de certains indicateurs de mobilité

En milieu urbain, si le nombre de déplacements quotidiens moyen évolue peu, en revanche, les caractéristiques de la mobilité évoluent rapidement. Il est donc nécessaire de suivre ces évolutions pour anticiper sur les besoins en infrastructures ou sur la politique à mettre en œuvre afin de limiter certains changements non souhaités. Sans anticiper sur le chapitre portant sur l'analyse de la mobilité, nous évoquons ce phénomène à travers un exemple.

Le nombre de déplacements en voiture particulière progresse de façon continue sous l'effet conjugué de multiples facteurs. Parmi ceux-ci, on peut évoquer la motorisation. Le graphique 1.1 montre son évolution entre 1950 et 1985. On y lit une progression quasi linéaire pour la plupart des pays sur cette même période. Doit-on en conclure que cette croissance se poursuivra au même rythme ? On peut penser à des effets de seuils, mais l'exemple américain semble montrer que l'asymptote n'est pas encore atteinte. Par ailleurs, on sait qu'il y a un lien entre motorisation et mobilité voiture particulière, certains allant même jusqu'à lier mécaniquement les deux. La mobilité voiture particulière serait donc appelée à croître inexorablement ?



Graphique 1.1 : Evolution du taux de motorisation dans les pays de l'OCDE

Cet exemple est présenté de manière volontairement simpliste. Par contre, il illustre clairement le problème et l'enjeu de la modélisation. Une prolongation mécaniste de tendance amènerait à des besoins d'investissement considérables pour accompagner la demande. En revanche, une analyse des facteurs explicatifs du phénomène peut mettre en lumière des inflexions de tendances qui sont susceptibles de modifier la demande à plus ou moins long terme.

Pour toutes ces raisons, il est impératif de planifier l'organisation et le développement des transports. Cette planification doit pouvoir s'appuyer sur des prévisions aussi fines que possible, compte tenu des incertitudes pesant sur l'évolution de l'environnement économique et social. C'est ainsi que progressivement, des procédures de planification ont été mises en place, parallèlement à l'évolution des outils de modélisation.

2. HISTORIQUE DE LA MODELISATION

Nous empruntons largement à Dupuy (1975) pour notre présentation jusqu'aux années 1970 et à Margail (1989) au-delà. Nous essayons dans la mesure du possible de mettre en parallèle l'évolution de la planification avec celle de la modélisation. Cette mise en perspective permet de mieux comprendre la trajectoire suivie par les modèles de prévision de la demande de déplacements en France. Nous ne prétendons pas proposer un historique complet de la modélisation, encore moins de la planification, mais plutôt présenter les temps forts du développement de la modélisation en France. A ce titre, nous insistons surtout sur la période qui a conduit à l'introduction des modèles de prévision de la demande de déplacements en France, puis à sa diffusion.

2.1. *Années 50 : les premiers modèles*

En France, jusqu'à la seconde guerre mondiale, l'automobile s'est développée sur la base du réseau routier existant. Les investissements réalisés se limitent le plus souvent à des élargissements ou des travaux d'amélioration pour renforcer la sécurité et la qualité du service. L'objectif est alors de « résoudre les problèmes en cours par des réalisations à court terme » (Dupuy, 1975, p.17). On dispose alors de modèles donnant la largeur de la voie en fonction du trafic, avec des pondérations selon le type de véhicules en présence. Les recueils de données associés sont des comptages.

2.1.1. *Mise en place du processus de transfert des modèles américains*

Après la seconde guerre mondiale, tous les efforts sont mobilisés sur la reconstruction du patrimoine. L'objectif dans le domaine des transports est de permettre la circulation des marchandises et des personnes en résolvant les goulots d'étranglement. Rapidement, les préoccupations vont évoluer à la suite de la forte croissance de l'industrie automobile. Dans les années 50, elle occupe une place très importante dans l'économie française. A partir des années 60, elle commence à devenir très fortement exportatrice. Elle constitue un fer de lance de la croissance économique. Sa rapide démocratisation (la production a triplé entre 1953 et 1962) provoque rapidement des phénomènes de congestion. Il apparaît alors que sans investissement massif dans les infrastructures de transport, c'est le modèle de développement économique français qui risque d'être remis en cause.

La Direction des Routes opère alors, au cours des années 1950, une véritable mutation pour entrer dans l'ère moderne de la technicité, pour gérer le volume important de crédits d'investissements routiers lié à la création du Fonds Spécial d'Investissement Routier mis en place en 1951 :

- tout d'abord, au niveau de la modernisation de l'appareil de production en équipement routier ;
- ensuite, au niveau administratif, avec la mise en place des structures nécessaires à la gestion des investissements routiers.

Cette mutation se réalise de 1950 à 1960, par un véritable courant de transferts techniques venant des Etats-Unis Dupuy (1975) rappelle les principales étapes de ce transfert réalisées à la suite de nombreuses missions d'ingénieurs de l'Equipement aux Etats-Unis :

- 1953 mise en place d'un système de comptage à câble et de courbes horaires de trafic ;
- 1954 réalisation à Orléans, puis à Nantes de la première enquête origines-destinations selon la méthodologie américaine. Elle permet de déterminer les lignes de désir du trafic, ce qui était impossible avec les seuls comptages ;
- 1955-56 réalisation de la première enquête cordon à Nantes ;
- 1957 première étude sur la génération du trafic par les grands ensembles et l'incidence de la densité d'habitation sur le trafic ;
- 1960-61 réalisation de la première enquête origines-destinations avec interview à domicile à Rennes ;
- « en 1960, une prévision de trafic est effectuée pour le dimensionnement de l'autoroute du Sud près de Paris. Cette étude est encore envisagée comme un problème de rase campagne.

Pourtant moins d'un an plus tard, en 1961, l'OTU (Omnium Technique Urbain) qui vient de se créer et J. Elkouby (un des premiers ingénieurs envoyés aux USA), chacun de son côté travaillent à la mise au point d'un modèle de prévision de trafic spécifiquement urbain. » (Dupuy, 1975, p.87). A partir de cette année là, les problèmes urbains sont traités de manière spécifique.

Cette mutation répond au besoin de rationalisation des choix d'investissements. « *Seule la connaissance quantitative de ces grandeurs peut permettre non seulement de trancher dans les choix d'investissements mais de justifier de façon rationnelle les choix effectués. Puisqu'il s'agit de justifier l'emploi de fonds, la rationalité dominante sera de type économique. Elle permettra de mettre en regard des dépenses, des réalisations ayant une efficacité difficilement contestable.* » (Dupuy, 1975, p.87).

2.1.2. Premiers modèles : émergence d'une nouvelle problématique

Si les enquêtes origines-destinations ont supplanté les méthodes de comptages, c'est que ces dernières ne répondaient plus au besoin de l'époque. Le comptage était utilisé pour répondre à des trafics existants sur des infrastructures existantes grâce à des actions de court terme. Dès lors qu'il s'est agi de réaliser de nouvelles infrastructures, la connaissance des lignes de désir du trafic s'imposait. Mais à son tour, l'enquête seule n'a plus suffi à résoudre les problèmes dès lors qu'il a fallu prévoir les besoins pour accompagner (ou plutôt devancer) le développement automobile. C'est alors que sont nés les modèles de prévision de trafic aux Etats-Unis au début des années 1950. Dupuy note d'ailleurs à propos de la prévision aux Etats-Unis « *Elle est simplement un instrument de gestion dans un système tout entier orienté vers la croissance et le maintien de cette croissance, de même que le comptage seul était un instrument de gestion dans un système stable.* » (Dupuy, 1975, p.22).

Les premiers modèles développés aux Etats-Unis reposent sur la méthode des facteurs de croissance. Le trafic futur T'_{ij} s'exprime à partir du trafic actuel T_{ij} par la formule suivante :

$$T'_{ij} = f(T_{ij}, F_i, F_j)$$

avec :

$$F_i = T'_i/T_i \quad \text{et} \quad F_j = T'_j/T_j$$

T_i = nombre de déplacements ayant une extrémité dans la zone i

T_j = nombre de déplacements ayant une extrémité dans la zone j

La fonction f prend différentes formes. Les variables les plus souvent prises en compte sont :

- « *le taux de motorisation des ménages de la zone ;*
- *le revenu moyen des ménages de la zone ;*
- *la distance au centre-ville ;*
- *la densité résidentielle ;*
- *le nombre d'emplois ;*
- *le nombre de logements ;*
- *la surface d'occupation du sol. » (Dupuy, 1975, p.25)*

La méthode des facteurs de croissance extrapole le trafic futur à partir du trafic existant. Elle est donc dépendante du contexte actuel. Elle ne permet pas de simuler correctement le trafic d'échange avec des zones dont l'urbanisation est en très forte croissance. Un trafic enquêté très faible voire nul, même multiplié par un facteur important, donnera toujours un trafic très faible ou nul. On tente alors d'extrapoler le trafic entre deux zones à partir des caractéristiques de ces deux zones sans passer par l'intermédiaire du trafic existant. On utilise des lois gravitaires ou des régressions linéaires. Cette évolution correspond à une problématique nouvelle qui se traduit par « *trois modifications principales :*

- *l'adoption systématique d'un horizon de long terme ;*

- *l'abandon de la référence contextuelle aux déplacements existants ;*
- *la prise en compte de l'ensemble d'un « réseau » au lieu d'un axe isolé. » (Dupuy, 1975, p.30).*

Sur le plan des modèles, la principale conséquence de cette évolution de problématique est l'émergence de deux phases : « génération » et « distribution ». Il y a dissociation entre une phase de génération du nombre de déplacements émis ou reçus par zone (génération) et la répartition de ces déplacements entre les différentes zones (distribution). Ces deux phases sont ensuite complétées par l'ajout de deux autres phases : la répartition modale et l'affectation du trafic. La phase de choix du mode a « surtout eu une fonction technique interne : le choix du mode permettant d'éliminer de façon rationnelle et avec la précision nécessaire des déplacements qui ne font pas ou peu appel aux dimensions de la route. La part de déplacements effectués autrement qu'en automobile (à pieds, en deux roues, en transport en commun) correspond dans l'ensemble du traitement à un résidu qui n'interviendra plus dans la suite du calcul. » (Dupuy, 1975, p.39). De plus, la part de ces modes tend à diminuer « et il ne serait venu à l'idée de personne qu'ils avaient quelque avenir ». Enfin, l'affectation renforce l'intérêt de la création de nouvelles infrastructures. En privilégiant le temps dans l'affectation et en utilisant le plus court chemin, on garantit le chargement de nouvelles infrastructures aux caractéristiques géométriques améliorant les conditions de circulation. Avec les « modèles à 4 étapes », c'est donc tout un édifice en faveur de l'investissement routier automobile qui se met en place, au moment où en France, on envisage le transfert de ces modèles.

Cette nouvelle problématique, introduite aux Etats-Unis, répond à un changement de préoccupation en matière d'investissements routiers. Il s'agissait de substituer un vaste programme de création de voies nouvelles à des aménagements ponctuels de capacité. Ce changement d'échelle a vu le jour à Détroit tout d'abord, puis Chicago au début des années 50. Les deux villes prévoyaient de construire un vaste réseau d'autoroutes urbaines. Mais « un rapide calcul fit apparaître que [les financements classiques] aurait permis d'inaugurer la deuxième autoroute du programme au bout de... 17 ans » (Dupuy, 1975, p.32). Les investissements nécessaires imposaient de recourir à des emprunts. Les modèles de prévision de trafic à long terme devenaient alors indispensables pour garantir les remboursements de ces emprunts à partir de taxes prélevées sur le trafic automobile. La mécanique de justification et de financement du développement du réseau routier était mise en marche. Elle est instituée, en 1956, par le vote du « *Federal Aid Highway Act* ».

2.2. Années 60 : l'acclimatation des modèles de trafic

Dès le début des années 1960, le dispositif est en place pour assurer le transfert des modèles de prévision de trafic américains en France. Les premières enquêtes origines-destinations ont déjà été réalisées dans plusieurs agglomérations, permettant d'alimenter les modèles avec des données françaises. D'emblée, la nouvelle problématique de la prévision américaine, à savoir l'adoption du long terme et la décontextualisation de la prévision, va être privilégiée en France. On y retrouve les mêmes objectifs qu'aux Etats-Unis. Il s'agit de rationaliser les choix d'investissements routiers nécessaires à la croissance de la motorisation garante de la croissance économique.

La première étude de prévision de trafic a été réalisée en 1960 pour le dimensionnement de l'autoroute Sud de Paris. Cette étude a été traitée comme un problème de rase campagne. Mais dès l'année suivante, des modèles spécifiquement urbains sont en cours d'élaboration. « *A partir de cette date, il y aura une appréhension spécifique des problèmes de transports urbains considérés comme distincts des problèmes de rase campagne* » (Dupuy, 1975, p.91).

En 1961, l'IAURP (Institut d'Aménagement et d'Urbanisme de la Région Parisienne) crée un département « Transports » qui aura pour première tâche la réalisation d'un inventaire des modèles de prévision de trafic urbain développés aux Etats-Unis. Il met ensuite au point son premier modèle en 1963 pour la prévision du trafic du RER (Réseau Express Régional) parisien. C'est également à cette

date qu'apparaissent les premières agences locales du SERC qui deviendront plus tard les CETE (Centre d'Etudes Techniques de l'Équipement).

Un questionnaire spécial sur les déplacements domicile-travail est introduit pour le recensement de 1962 dans quinze agglomérations. Il servira de base à l'estimation de paramètres adaptés au contexte français.

Une note rédigée par Gérondeau (1966), du SERC (Dupuy, 1975, p.93) illustre bien la doctrine officielle : « *La solution des problèmes de transport et d'urbanisation dans les agglomérations ne saurait reposer sur des transports en commun coûteux et peu souples. Seuls des réseaux routiers et autoroutiers fournis peuvent à la fois écouler le trafic à venir et ouvrir à l'urbanisation de nouveaux terrains* », et à propos des tests de cohérence des réseaux routiers « *Ces tests de cohérence sont heureusement rendus possibles aujourd'hui grâce aux enquêtes réalisées et aux méthodes d'études élaborées par le Ministère de l'Équipement, au cours des années récentes. Dans un proche avenir, les schémas de structure et de transport seront systématiquement, dès leur conception, soumis à ces tests, de façon à ce que les solutions retenues soient valables aussi bien fonctionnellement qu'économiquement.* » « *Une autre note du SERC intitulée « Le test des Schémas de structure et des Transports » explique l'utilisation de l'outil de test. Il s'agit de la célèbre chaîne américaine des 4 modèles (ou modules) : génération, distribution, répartition modale, affectation* ».

	moyen terme	long terme		très long terme
petite agglomération		Cherbourg Alençon Beauvais Hâguenau-Bischwiller	Mers le Tréport Dieppe Douai Eu le Tréport	
agglomération moyenne	St Nazaire Nancy Rouen Vallée de la Sambre	Angoulême Avignon Caen Chartres Dijon Le Havre Montpellier Mulhouse Nancy	Nantes Pau Rouen Toulon Valence Dunkerque Evreux Orléans Valenciennes	Strasbourg Le Havre Grenoble
grande agglomération	Marseille 75 Métro Lyon Métro Marseille Bordeaux 75	Métropole Nord Métro Lyon Métro Marseille Bordeaux 85		

Tableau 1.1 : Tests de cohérence dans les études de transport au cours des années 1960
(Source : Averous, 1971)

Les années qui suivent, se traduisent par une diffusion de cette doctrine officielle au niveau local. Le modèle du SERC est utilisé dans de nombreuses agglomérations qui dans le même temps se dotent de plans d'urbanisme, baptisés schémas de structure. Ces schémas de structure font l'objet de tests de cohérence selon la doctrine du SERC pour vérifier la cohérence à long terme entre les hypothèses d'urbanisme et le réseau routier. Le tableau 1.1 (Dupuy, 1975, p.98), « *établi en septembre 1970, donne une idée des tests réalisés ou commencés au cours des années 1960* ».

En 1967-68, le modèle DAVIS (Distribution, Affectation et visualisation du trafic) voit le jour. A sa création, c'est un modèle à 4 étapes. Mais rapidement seule la phase d'affectation sera développée. De nombreux concurrents voient le jour, développés par des bureaux d'études.

La fin des années 1960 marque sinon un tournant, au moins un répit dans le développement de la modélisation. Sur le plan politique, Mai 1968 a suscité des réflexions dans les milieux avancés de l'urbanisme. Les grandes villes de province commencent à connaître des problèmes importants de circulation. La situation financière des transports collectifs, « *longtemps masquée par la politique de « rationalisation » (suppression des lignes déficitaires, suppression des dessertes nocturnes...)* » (Dupuy, 1975, p.102) est catastrophique. Le stationnement est devenu un problème. Ce contexte amène à des réflexions sur les modèles de trafic. Cependant, dans le même temps, la machine administrative continue de fonctionner et fait tourner les modèles.

2.3. Début des années 1970 : la mise en place d'une procédure de planification urbaine

Les années 1970 marquent une généralisation de l'utilisation des modèles, dont la place se trouve codifiée par la mise en place de nouvelles procédures de planification urbaine (tableau 1.2). La Loi d'Orientation Foncière (LOF) est votée le 30 décembre 1967 (Lois et décrets). « *Elle est destinée à mettre en place les nouvelles procédures de planification urbaine. Elle institue les Schémas Directeurs d'Aménagement et d'Urbanisme (SDAU), les Plans d'Occupation des Sols (POS) et les Zones d'Aménagement Concertées (ZAC)* » (Margail, 1989, p.29). Il faudra toutefois attendre plusieurs Circulaires entre 1971 et 1973, et notamment la Circulaire N°72-67 du 3 mai 1972 pour que les principales règles dans lesquelles s'inscrivent les études de transport, voient le jour. Cette dernière Circulaire (tableau 1.2) définit 5 niveaux d'étude selon le terme envisagé.

Cette codification s'accompagne de la constitution de « *dossiers pilotes diffusés dans toutes les villes concernées. Il s'agit de véritables manuels prévoyant en détail les conditions de lancement des études, les outputs attendus et les instruments à utiliser... Du point de vue des modèles de trafic, il s'agit probablement d'une date significative dans l'histoire de leur développement. En effet, les dossiers pilotes font tous référence à l'utilisation de modèles.* » (Dupuy, 1975, p.106-107). Les modèles de prévision de trafic se trouvent ainsi légitimés tant pour le long terme (EPIT : Etude Préliminaire d'Infrastructure de Transport) que pour le court terme (plan de circulation, institué par la Circulaire du 16 avril 1971).

2.4. 1975 - 1980 : le tournant

Cette période est marquée par la crise énergétique et le renversement de la croissance économique. Elle l'est aussi « *par l'attention portée aux problèmes d'ordre social, psychologique, économique, écologique, etc. A la recherche de solutions techniques à apporter aux problèmes « techniques », ont succédé des préoccupations en termes de « qualité de vie »* ». (Margail, 1989, p.40).

Les EPIT commencent à être critiquées. Des voies s'élèvent pour dénoncer l'impasse dans laquelle se trouve la politique du « tout automobile ». La lecture des prévisions de trafic conduit à des besoins d'infrastructures difficilement réalisables, d'une part en raison des montants financiers à mobiliser, d'autre part par la défiguration des villes qu'ils imposent. Parallèlement, l'Etat commence à se désengager financièrement. Ainsi, les pouvoirs publics abandonnent la planification à long terme au profit de la gestion du présent.

Objectifs à atteindre	Horizon de l'étude	Nature des études à entreprendre	Applications
Niveau I Plan d'organisation générale des transports incluant le schéma des infrastructures de transport de l'agglomération.	Horizon du SDAU 30 à 40 ans	Organisation générale à long terme des transports et, en première étape, études préliminaires d'infrastructure de transport.	Incorporation du schéma d'infrastructure des transports dans le SDAU.
Niveau II Plan des emprises à réserver. Règles d'urbanisme au voisinage des voies existantes ou à créer, liste des opérations à déclarer d'utilité publique à l'occasion des POS. Possibilité d'acquisition foncière d'opportunité.	15 à 20 ans	Etudes approfondies de transports et de tracés. Etablissement d'avant-projets géométriques. Détermination du plan des emprises à réserver.	Inscription au POS des terrains réservés et des règles d'urbanisme relatives aux infrastructures de transport et de voirie.
Niveau III Programmation quinquennale d'opérations individualisées.	Deux plans quinquennaux (10 ans)	Etudes de programmation et de financement à dix ans. Etablissement des dossiers d'inscription au prochain plan quinquennal.	Acquisition foncière d'opportunité.
Niveau IV Avant-projet sommaire.	1 à 5 ans	Mise à jour des avant-projets géométriques.	Enquête d'utilité publique. Acquisitions foncières.
Niveau V Exécution des travaux	-	Avant-projets détaillés.	Réalisation des infrastructures ou mise en œuvre des plans de circulation.

Tableau 1.2 : Mise en place d'une procédure de planification urbaine
(Source : Magazine Equipement-Transport, Moniteur du 10 juin 1972, p.133)

C'est également à cette période qu'est lancé le programme de recherche « ATP socio-économie des transports » qui donnera le jour à un foisonnement d'approches et d'analyses sur la mobilité urbaine. Elles amèneront un flot de critiques sur la pertinence des modèles. La principale étant que les relations utilisées dans les modèles entre un certain nombre de variables d'entrée et la mobilité, ne sont pas explicatives des comportements individuels. C'est notamment pour répondre à cette critique que sont développés les modèles désagrégés aux Etats-Unis. Toutefois, leur pénétration en France restera très faible, y compris aujourd'hui (Les travaux les plus significatifs en milieu urbain sont peu nombreux : Abraham et al., 1961 sur les choix d'itinéraires ; CETUR, 1985 ; Daly, 1985 ; CETE de Lyon et al., 1986 (travaux sur Grenoble et Nantes) ; Bouyau, 1988 (travaux sur Rennes) ; Hivert et al., 1988 (modélisation sur ORLYVAL) ; RATP, Cambridge Systematics, 1982 ; Rousseau, Saut, 1997 (modèle IMPACT de la RATP) ; CERTU, 1998b (modèle de choix combinés destination, mode sur Lyon), et des travaux précurseurs mentionnés par Géraud (1977) : le modèle Paris – Rouen – Le Havre (CERAU, SETEC, 1968) et le modèle interurbain TRIP de la SETEC et du SAEI (Marche, Dumontier, 1974) qui sont présentés par Géraud comme des modèles désagrégés directs). Les

recherches menées soulignent la multiplicité des facteurs influençant la mobilité, nécessitant l'adoption d'un cadre conceptuel plus riche que celui sous-tendant les modèles classiques de « prévision à 4 étapes ». Ces recherches seront poursuivies tout au long des années 1980. Mais force est de constater que si elles ont effectivement mis en évidence le manque de pertinence de ces modèles, elles n'ont par contre pas réussi à proposer des outils opérationnels de prévision. Ainsi, les modèles continueront à être utilisés, accroissant le gap entre les chercheurs et les praticiens.

2.5. Années 1980 à nos jours : vers une approche globale des problèmes de déplacements

2.5.1. Instruments de planification

La décennie 1980 s'inscrit dans un contexte de changement social, avec la décentralisation qui renforce le pouvoir des élus locaux. Fin 1982, la Loi d'Orientation des Transports Intérieurs (LOTI Loi N° 82-1153 du 30 décembre 1982 publiée au Journal Officiel du 31 décembre 1982 (Lois et décrets)) est votée. Elle institue le « droit aux transports », c'est-à-dire la possibilité pour les usagers de se déplacer dans des conditions raisonnables de prix, de temps d'accès et de confort, ainsi que de coût pour la collectivité. Pour être effectif, ce droit nécessite une bonne connaissance des besoins de déplacements en liaison avec l'organisation de la ville pour ce qui est des déplacements urbains. De la LOTI, découlent les Plans de Déplacements Urbains (PDU). Il s'agit d'une procédure de concertation sur l'organisation des déplacements, de la circulation et du stationnement dans les agglomérations, en lien avec les projets d'urbanisme. La démarche sera utilisée pour de nombreuses d'études de déplacements urbains. De ce fait, les problèmes sont abordés de manière plus globale, intégrant tous les modes de transports. De plus, « *l'élaboration du PDU est l'occasion d'une concertation très large, où s'articulent la planification urbaine et la gestion des divers moyens de transport.* » (Margail, 1989, p.49). Toutefois, 13 ans après, les décrets d'application n'étaient toujours pas sortis, dans la mesure où l'institutionnalisation de la démarche soulève de nombreuses questions de pouvoirs et des problèmes complexes d'ordre juridique. L'absence de cadre juridique précis (entre 1983 et 1996) n'a pas empêché cependant (et peut-être même a encouragé) les villes à se lancer dans la démarche. Vers le milieu des années 90, les PDU ont même été réactivés et de nombreuses agglomérations se sont lancées dans la procédure. L'expérience acquise au cours de ces années a conduit le CERTU (Centre d'Etudes sur les Réseaux, les Transports, l'Urbanisme et les constructions publiques) à publier un guide méthodologique et pratique pour la conduite de la réflexion (CERTU, 1996).

La longévité de la procédure et son actualité méritent que l'on s'attarde quelque peu sur les plans de déplacements urbains d'autant plus que la loi sur l'air et l'utilisation rationnelle de l'énergie les a relancés. « *Le concept des « Plans de Déplacements Urbains » (PDU) a été formalisé pour la première fois par la Loi d'orientation des transports intérieurs (LOTI) en 1982. Elle a initié les PDU dans son article 28, en leur donnant comme enjeu principal de définir « les principes généraux de l'organisation des transports, de la circulation et du stationnement » pour permettre « une utilisation plus rationnelle de la voiture et assurer la bonne insertion des piétons, des véhicules à deux roues et des transports en commun.* » » (CERTU, 1996, p.14).

Le PDU marque une volonté d'approche globale des problèmes de déplacement dans la ville et non plus seulement de circulation automobile ou de transports collectifs. L'expérience de nombreuses villes, qui ont mené des politiques trop parcellaires sans coordination forte entre les modes, montre que la maîtrise des déplacements passe par une réflexion globale. Les efforts de promotion ou les investissements lourds opérés dans les transports collectifs n'ont pas toujours apporté les fruits que l'on pouvait en attendre, avec même parfois, si ce n'est des baisses de fréquentation, des baisses de parts de marché vis-à-vis de l'automobile (l'exemple de Lyon est significatif à cet égard. D'après les enquêtes ménages déplacements de 1986 et 1995, la part de marché des transports collectifs (au sein des déplacements motorisés) a diminué de 23,5 à 20,6 % et celle de l'automobile s'est accrue de 73,6 à 77,6 %, malgré le doublement de l'offre métro et la création de 20 lignes de bus au cours de la même période). C'est que dans le même temps, la politique d'urbanisme ou la politique d'investissement

routier ou de gestion de la circulation a largement accompagné les tendances lourdes de développement de l'usage de la voiture.

Cette situation conduit à la triple crise des déplacements que connaissent nombre d'agglomérations, qui se nomme déficit public, congestion, pollution (Raux, Tabourin, 1991 ; Bonnel et al., 2001). La crise des finances publiques limite fortement les possibilités d'investissements nouveaux pour répondre à la croissance de la demande par davantage d'infrastructures routières ou de transports collectifs. Celle de la congestion s'accompagne d'un accroissement des heures perdues dans les embouteillages et d'une perte d'efficacité des systèmes de transport. Enfin, celle de l'environnement se traduit par des nuisances accrues qui conduisent à des réflexions sur la réduction de la pollution automobile et sur la place des différents modes de transport, dans le cadre d'un développement durable (Banister, 1998).

La démarche adoptée dans les PDU doit donc s'efforcer de faire face à de nombreux défis :

- coordonner l'ensemble des acteurs et notamment les différentes compétences territoriales au niveau des pouvoirs de police, de circulation, de stationnement, d'urbanisme ou de l'autorité organisatrice des transports collectifs ;
- travailler à l'échelle de l'agglomération et pas seulement de la ville centre ou du centre-ville. Du fait de l'élargissement du bassin d'emploi ou de vie, ce périmètre dépasse bien souvent celui de l'autorité organisatrice à qui est confiée la maîtrise d'ouvrage des PDU ;
- prendre en compte les interactions entre les différents modes de transports : voiture particulière, transports collectifs, marche à pied et deux roues, tout en prenant en compte les problématiques du stationnement et du transport de marchandises ;
- appréhender les interactions entre l'urbanisme et les déplacements, notamment au niveau des facteurs contribuant à l'étalement urbain tant des activités que de l'habitat et à l'impact de la spécialisation des espaces ;
- améliorer la qualité de l'environnement urbain, tant au niveau de la réduction des externalités négatives des déplacements que de la qualité de la vie.

Jusqu'à la fin 1996, « *Le PDU est un document de référence pour toutes décisions des pouvoirs publics ayant des implications sur les déplacements urbains* ». En revanche, il n'est pas de nature contractuelle. De ce fait, « *il n'édicte pas des prescriptions contraignantes et ne subordonne pas les décisions des différentes autorités (en particulier, il n'y a pas de mise en conformité obligatoire des autres documents de planification)* » (CERTU, 1996, p.23). De ce fait, le PDU nécessite une forte volonté de coopération entre les différents acteurs. La concertation est une nécessité dans la démarche pour que les différents acteurs concernés acceptent et mettent en œuvre les orientations adoptées dans le PDU. C'est ce qui fait à la fois la force et la faiblesse des PDU.

Depuis l'adoption de la Loi sur l'air et l'utilisation rationnelle de l'énergie (Loi n°96-1236 du 30 décembre 1996), des PDU nouvelle version ont vu le jour. L'article 28-1 de la LOTI a été modifié en élargissant les orientations des PDU. Toutefois, la principale nouveauté de ces PDU nouveaux concerne leur portée. « *En effet, « les décisions prises par les autorités chargées de la voirie et de la police de circulation (c'est-à-dire les Maires) ayant des effets sur les déplacements urbains doivent être compatibles ou rendues compatibles avec le PDU* ». De plus, les POS et les plans d'aménagement des ZAC (Zones d'Aménagement Concertées) devront prendre en compte les orientations des PDU quand ils existent. Enfin, le projet de plan devra être soumis à enquête publique, chaînon manquant dans la LOTI » (Lettre du GART, N°125, p.1). Le vote de la Loi SRU (Loi N°2000-1208 du 13 décembre 2000 relative à la Solidarité et au Renouvellement Urbain (Lois et décrets)) a encore renforcé l'article 28.1 relatif aux PDU qui sont rendus obligatoires dans les agglomérations de plus de 100 000 habitants :

« Art. 28-1. – *Les plans de déplacements urbains portent sur :*

« 1°A *L'amélioration de la sécurité de tous les déplacements, notamment en définissant un partage modal équilibré de la voirie ...*

- « 1° La diminution du trafic automobile ;
- « 2° Le développement des transports collectifs et des moyens de déplacements économes et les moins polluants, notamment l'usage de la bicyclette et la marche à pied ;
- « 3° L'aménagement et l'exploitation du réseau principal de voirie d'agglomération ... , afin de rendre plus efficace son usage notamment en l'affectant aux différents modes de transport... ;
- « 4° L'organisation du stationnement... ;
- « 5° Le transport et la livraison des marchandises ... ;
- « 6° L'encouragement pour les entreprises et les collectivités publiques à établir un plan de mobilité et à favoriser le transport de leur personnel, notamment par l'utilisation des transports en commun et du co-voiturage.
- « 7 La mise en place d'une tarification et d'une billetterie intégrées pour l'ensemble des déplacements... ».

La maîtrise d'ouvrage du PDU est confiée à l'autorité organisatrice des transports collectifs. A ce titre, il fournit l'occasion à l'autorité organisatrice de tenter d'influer sur la priorité accordée aux transports collectifs dans la politique d'agglomération. Toutefois, le champ d'action limité de l'autorité organisatrice (notamment sur le plan spatial) pourrait réduire la portée des PDU nouvelle version du fait de leur nouveau caractère contraignant. La loi est cependant trop récente pour porter un premier diagnostic.

L'horizon d'étude du PDU est relativement court, puisqu'il concerne généralement une planification à 10 ans. De ce fait, cet instrument est complété par les Schémas Directeurs et les DVA (Dossier de Voiries d'Agglomération) institués par la circulaire du 15 novembre 1991 (CETUR 1993a,b), qui concernent la planification de long voire très long terme.

2.5.2. Evolution des objectifs des politiques de déplacements urbains : conséquences pour la modélisation

Cette période marque une lente affirmation de l'évolution des politiques de déplacements urbains. Il ne s'agit plus seulement d'assurer l'écoulement des flux de véhicules, ni même de déplacements, mais beaucoup plus d'assurer la gestion des déplacements urbains dans une perspective de court/moyen terme (PDU), qui tienne compte des orientations de long terme (schéma directeur, DVA) et des contraintes de développement durable.

La LOTI a introduit le droit au transport. Cela signifie qu'une attention plus grande doit être portée aux besoins de déplacement des individus. Il ne suffit plus d'assurer la circulation des véhicules (particuliers ou collectifs), il faut apporter de bonnes conditions de déplacement aux individus, quel que soit le moyen de transport utilisé (voiture, bus, métro, mais aussi deux roues ou marche à pied). Mais parallèlement, la triple crise des déplacements urbains (financement, congestion, environnement) s'est progressivement affirmée comme une réalité. Les collectivités n'ont plus les moyens et/ou la possibilité de répondre à la demande uniquement par l'accroissement de la capacité des infrastructures. Enfin, plus récemment les réflexions sur le développement durable amènent à poser différemment la question des déplacements urbains. La croissance de la mobilité (et tout particulièrement de l'automobile) n'est plus forcément synonyme de développement et de progrès. De ce fait, d'autres politiques doivent être mises en œuvre qui permettent d'assurer une meilleure complémentarité et gestion des modes de transport et une régulation de la demande de déplacements.

L'analyse des politiques mises en œuvre dans différents pays européens (Bonnel et al., 1994a) montre que les actions développées sont assez variées. Elle souligne surtout la nécessité d'une approche globale des problèmes qui intègre tous les modes de transport, y compris le stationnement, dans un cadre spatial étendu à l'ensemble du bassin d'attraction de l'agglomération. De nombreuses pistes sont envisagées. Elles traduisent le foisonnement des réflexions, mais également la difficulté à résoudre cette triple crise. Sans être exhaustif, on peut citer les stratégies suivantes : le développement

de l'utilisation des transports collectifs ou de l'usage des parcs relais en liaison avec des axes lourds de transports collectifs, la promotion des modes de proximité par la création et la protection de réseaux piétonniers ou de pistes cyclables, l'encouragement du car pool, la maîtrise de l'usage de la voiture par diverses mesures de « *traffic calming* », la maîtrise de la production et de l'usage du stationnement public et privé, la tarification différenciée dans le temps et l'espace de l'usage des infrastructures de transport, la modulation des horaires de travail ou le développement du télétravail, l'introduction de nouveaux véhicules (petits véhicules ou véhicules électriques), des autoroutes intelligentes permettant d'optimiser l'usage des infrastructures... (CETUR, 1994 ; Brög et al., 1996).

Cette liste pourrait encore être allongée. Elle illustre clairement l'évolution des objectifs de politiques de déplacements. Elle met surtout en évidence l'évolution de la problématique de la prévision de la demande de déplacements. Il ne s'agit plus seulement d'avoir des modèles qui déroulent successivement les quatre étapes : décision de se déplacer, choix de la destination, choix du mode, choix de l'itinéraire. Cette évolution de la problématique « *implique par exemple l'élaboration de modèles qui ne s'intéressent pas uniquement aux déplacements automobiles (affectation de trafic), qui soient sensibles aux politiques d'orientation de la demande (choix multimodal) ou de réduction de la mobilité motorisée (pratiques piétonnières, impacts des équipements de proximité, etc.) ou encore qui permettent de simuler de nouvelles actions d'orientation de la demande inédites jusqu'ici (régulation tarifaire, réglementation d'accès, etc.)* » (Raux, Chapleau in Bonnel et al., 1997).

La construction de ces nouveaux modèles nécessite également des données nouvelles. Trois exemples permettent d'illustrer ces évolutions tant au niveau de la production des données que des modèles construits à partir de ces données.

Les caractéristiques des déplacements doivent être connues de manière plus fine pour évaluer l'incidence de modifications de l'offre de transport sur les choix de mode, sur les itinéraires empruntés... La représentation simplifiée de l'agglomération par un découpage en zone limite les possibilités de repérage fin de la proximité aux transports (distances entre les lieux d'activités terminales ou initiales et les arrêts de transports collectifs ou le lieu de stationnement, durée et pénibilité des ruptures de charges...). Ces caractéristiques influent pourtant sur le choix du mode de transport. Des expériences de géocodification en coordonnées x-y des localisations mettent en évidence l'importance des approximations et des erreurs que peut entraîner une codification zonale pour l'analyse et la modélisation micro-économique de la mobilité (Bonnel et al., 1994b, 1997 ; Chapleau, De Cea, 1983 ; Chapleau, 1991, 1993). Le modèle MADITUC construit à partir des données des enquêtes régionales origines-destinations réalisées à Montréal et Toronto est une illustration de la faisabilité et de l'apport du repérage fin des données spatiales (Chapleau, 1997 ; Dorland et al., 1997).

La prise en compte de l'environnement social et spatial des individus vise à mettre en évidence le jeu des contraintes et opportunités qui structurent la mobilité quotidienne. C'est pour les appréhender que se sont développées des approches en termes d'activités. Elles « *ont permis de faire évoluer le paradigme de la mobilité : on est passé progressivement de la notion de déplacement à celle d'activité et, par le biais de l'auteur de cette activité, à la notion d'individu situé socialement dans son entourage (familial, professionnel, etc.), ce dernier générant des obligations interpersonnelles* » (Raux, Chapleau in Bonnel et al., 1997). Cette évolution s'est traduite par un glissement du recueil des déplacements vers le recueil des activités (Bonnel, 1985, 1995b ; Axhausen, 1995 ; Stopher, 1998). Cette focalisation sur les activités est une reconnaissance explicite du fait que la demande de déplacements est une demande dérivée du besoin de réaliser des activités. Elle permet de plus de limiter l'oubli des petits déplacements dans les enquêtes auto-administrées par carnet de bord, car elle semble mieux correspondre à la représentation que les enquêtés ont de leur propre mobilité (Stopher, 1998). « *Ces approches en termes d'activités ont donné lieu au développement de modèles fondés sur les activités, dont HATS est l'exemple le plus connu* (Jones et al., 1980) » (Raux, Chapleau in Bonnel et al., 1997). Ces travaux se poursuivent actuellement et donnent lieu à une production

scientifique assez abondante (nombreux papiers à la « *Conference on Activity Based Approaches: Activity Scheduling and the Analysis of Activity Patterns* » tenus à Eindhoven aux Pays-Bas en mai 95 et publié sous Ettema, Timmermans, 1997 ; Timmermans et al., 1997).

Parmi l'ensemble des mesures envisagées de gestion des déplacements urbains, certaines d'entre elles ne sont pas encore appliquées ou le sont dans des contextes très différents de ceux qui peuvent se présenter dans certaines agglomérations. Les méthodes traditionnelles s'appuyant sur les comportements observés (encore appelés « préférences révélées » [*revealed preferences*]) sont parfois défailtantes pour simuler les conséquences de ces mesures. C'est pour répondre à ces situations que se sont développées diverses méthodes de recueil des données dites de « réponses déclarées » [*stated responses*], même si sous ce terme générique, on regroupe des méthodes de nature très différente (Lee-Gosselin, 1997, chapitre 4, p. 134). Elles permettent de simuler les réponses face à des situations hypothétiques non expérimentées par l'enquêté, ou encore d'explorer expérimentalement les mécanismes du changement de comportement. Ces méthodes permettent de mieux comprendre ce processus de changement et de mieux le piloter pour atteindre les objectifs de la politique de déplacements urbains (Andan, Raux, 1997; Jones, 1997).

Toutefois, si ces évolutions sont relativement présentes dans le monde anglo-saxon ou dans les milieux de la recherche, elles n'ont que faiblement pénétré le monde des études en France, mis à part quelques expériences anciennes ou plus récentes (par exemple, CERTU, 1998b). De ce fait, en France, la pratique courante des modèles de prévision de la demande de déplacements (et des enquêtes : l'examen des dossiers méthodologiques du CERTU est significatif à cet égard (CETE, CETUR, 1989 ; CERTU, 1998a)) n'a pas connu d'évolutions importantes. Certes, les modèles ont bénéficié d'améliorations, notamment dans les algorithmes d'affectation qui prennent en compte la congestion ou encore dans leur plus grande convivialité tant au niveau de l'utilisateur que pour les sorties graphiques. Des modèles complémentaires permettent bien souvent de calculer des volumes d'émission de polluants ou des niveaux de bruit. Mais leurs fondements théoriques n'ont pas évolué. Nous centrons donc notre attention sur ces modèles que l'on peut qualifier de « classiques » et que l'on englobe généralement sous le terme générique de modèles à quatre étapes, même s'ils ne mettent pas tous en œuvre ces quatre étapes ou bien lui en adjoignent d'autres.

3. QUELS MODELES ? POUR QUELS DOMAINES D'UTILISATION ?

Le choix de centrer notre attention sur les modèles « classiques » à quatre étapes s'explique principalement par les considérations suivantes. Ce manuel vise à apporter les éléments théoriques et méthodologiques de base permettant d'aborder la pratique professionnelle avec un minimum de recul et de connaissance sur la modélisation que ce soit en tant que commanditaire ou destinataire d'études de prévision de la demande ou en tant que producteur de telles études. Il convient alors d'être au fait des modèles couramment utilisés par les bureaux d'études ou les CETE.

Par ailleurs, il nous semble indispensable de connaître les outils de la modélisation « classique » avant d'aborder les nouvelles approches en cours de développement. En effet, nombre de celles-ci utilisent pour partie les outils classiques, que ce soit en conservant certaines étapes, en venant se greffer sur celles-ci ou encore en reprenant certains des algorithmes développés pour ces modèles. De plus, la connaissance des apports et limites des modèles à quatre étapes permet le plus souvent de mieux apprécier la portée des nouveaux modèles.

Enfin, comme les outils plus classiques, les nouvelles approches ne permettent pas de répondre à l'ensemble des questions en matière de prévision de la demande. Les modèles que l'on regroupe sous le terme générique de modèles à quatre étapes conservent donc des domaines de pertinence. C'est le cas notamment de la prévision de trafic sur de nouvelles infrastructures. Par exemple, un modèle relativement « simple », comme TERESE (SEMALY, 2000), a eu plusieurs fois l'occasion de montrer son opérationnalité (nous revenons sur cette notion dans le chapitre suivant) en prévoyant des trafics

peu éloignés des trafics effectivement observés pour de nouvelles infrastructures de transports collectifs en site propre (SEMALY, 1996 ; GESMAD, 2000).

Si ces modèles ne permettent pas de répondre à l'ensemble des questions, il convient de préciser leurs domaines principaux de pertinence. Notre présentation de l'historique de la modélisation a mis en évidence le glissement progressif de la problématique de la planification : d'un axe isolé à un réseau, des véhicules vers les individus, puis plus récemment de la charge de nouvelles infrastructures à la régulation des déplacements. La réponse à la question « que doit-on prévoir ? » nous renvoie en fait à la question « quelles sont les politiques de déplacements urbains ? ». La déclinaison de ces politiques met en évidence l'élargissement de la problématique de la prévision. Il en découle une multiplicité des dimensions à prendre en compte et une complexité de l'exercice de prévision.

Nous proposons une grille de lecture de cette complexité qui permet de l'organiser. Cette grille de lecture se décline selon deux dimensions. La première s'appuie sur une conception systémique de l'urbain qui permet de le décomposer en trois sous-systèmes (Bonnafous, Puel, 1983). La seconde définit les dimensions de l'exercice dans l'espace et le temps. Cette double entrée permet de définir la nature des interactions devant être prises en compte par le modèle et donc le choix de la famille de modèles à retenir. Sur cette base, nous pouvons définir le domaine de prédilection des modèles à quatre étapes.

3.1. Une décomposition du système urbain en sous-systèmes

Nous proposons une lecture de la ville maintenant classique mais toujours aussi structurante pour l'analyse (Bonnafous, Puel, 1983, nous développons toutefois une interprétation légèrement différente des trois sous-systèmes, plus adapté à notre objectif). Elle s'organise sur une décomposition et une articulation du système urbain en trois sous-systèmes :

- le sous-système de transport ;
- le sous-système de localisation des activités ;
- le sous-système des relations sociales.

Le sous-système de transport comprend l'ensemble des infrastructures de transport quel qu'en soit le mode de transport sur le territoire considéré. Ces infrastructures sont considérées avec leurs caractéristiques techniques : capacité, vitesse, courbe débit-vitesse, longueur, fréquence pour les transports collectifs.... Ces données sont enrichies par les conditions d'utilisation définies à travers la politique de déplacements urbains mise en place. C'est à l'intérieur de ce sous-système que se met en œuvre l'action sur l'offre de transport, que ce soit à travers la création ou la transformation d'infrastructures, ou à travers la politique de déplacements urbains qui peut en modifier l'usage ou le fonctionnement. C'est également à partir de ces données que l'on pourra déterminer le coût généralisé des déplacements.

Le sous-système de localisation des activités correspond à l'inscription dans l'espace des équipements où sont réalisées les activités urbaines. Dans une problématique de modélisation, les équipements les plus souvent considérés sont le domicile, les emplois, les établissements d'enseignement, les équipements commerciaux et de loisirs. Ceux-ci sont considérés à travers leurs caractéristiques physiques : localisation, taille, nature détaillée d'activités..., économiques : prix du terrain ou du bâti, chiffres d'affaires des activités..., de fonctionnement : horaire d'ouverture... ou encore d'image : composition sociologique des quartiers, type d'activités.... C'est à l'intérieur de ce sous-système que se mettent en œuvre les logiques de production-transformation de l'espace par les différents acteurs.

Le sous-système des relations sociales correspond à l'ensemble des relations et activités des acteurs du système urbain. Son inscription dans l'espace traduit le fonctionnement socio-spatial de la ville. Pour les individus, il transparaît au niveau des pratiques urbaines, c'est-à-dire des

comportements d'activités et de mobilité. Ceux-ci sont repérés à travers diverses enquêtes sur les déplacements (enquête ménages déplacements, cordon, d'utilisation des réseaux, comptages routiers...). C'est à l'intérieur de ce sous-système que l'individu définit son programme d'activités (nous utilisons à dessein le terme de « programme d'activités » au sens d'un ensemble d'activités que l'individu souhaite réaliser par distinction avec le « schéma d'activités » qui correspond à l'ensemble des activités effectivement réalisées).

Nous aurons l'occasion de revenir à de nombreuses occasions sur cette décomposition du système urbain qui se révèle particulièrement structurante et stimulante.

C'est de la « confrontation » entre les trois sous-systèmes que se concrétise le schéma des activités réalisées par les individus qui est inscrit dans l'espace et se traduit par l'utilisation du sous-système de transport et donc par la demande de déplacements (tout au moins lorsqu'il y a déplacement dans l'espace). Le modèle doit donc traduire ces interactions entre les trois sous-systèmes. Toutefois, ces interactions sont multiples et le plus souvent les relations ne sont pas univoques. L'analyse du système reste donc encore complexe. L'adjonction des dimensions spatiales et temporelles permet de réduire cette complexité en organisant l'analyse de ces interactions.

3.2. Dimensions de l'exercice : espace-temps

Par rapport à la question de la prévision de la mobilité quotidienne qui est le problème qui nous concerne, les interactions entre les trois sous-systèmes s'organisent selon des temporalités variables. Quelques exemples permettent de l'illustrer. Si l'on considère les interactions entre le sous-système des localisations et le sous-système des relations sociales, dans le court terme, on peut considérer le sous-système des localisations comme figé. La dynamique des relations entre les deux sous-systèmes sera donc principalement univoque, l'inscription dans l'espace du programme d'activités découlera de l'état du sous-système des localisations. En revanche, dans le long terme, les projets des individus peuvent se traduire par des changements de localisation résidentielle et donc par une évolution du sous-système des localisations qui en retour se traduira par une nouvelle inscription spatiale du programme d'activités des individus. Dans cet exemple, l'état ou l'évolution du sous-système des transports peut bien évidemment influencer sur ces évolutions.

On pourrait donner d'autres exemples de l'influence de la dimension temporelle dans la nature des interactions entre ces trois sous-systèmes. Cette situation résulte en fait des rythmes d'évolution des trois sous-systèmes. C'est le sous-système des relations sociales qui bénéficie des évolutions les plus rapides tout au moins dans la traduction qui nous préoccupe ici, c'est-à-dire la mobilité quotidienne. Inversement, les évolutions du sous-système des localisations sont beaucoup plus lentes en raison de la forte inertie de ce sous-système. Le sous-système des transports occupe le plus souvent une situation intermédiaire, tout au moins dans les exercices de modélisation les plus fréquents. En effet, un des objectifs de la modélisation est généralement de prévoir la fréquentation d'une nouvelle infrastructure ou de simuler diverses actions sur ce sous-système.

L'espace constitue la seconde dimension importante de l'exercice. La prévision sur un espace très restreint, par exemple le fonctionnement d'un carrefour, ne mettra pas en jeu la même nature d'interactions que sur l'ensemble d'une agglomération. Dans le premier cas, le sous-système des localisations et pour une large part celui des transports seront considérés comme figés, ou bien leurs évolutions seront considérées comme exogènes au modèle. Dans le second, la même hypothèse serait plus hardie.

On peut ainsi croiser les deux dimensions et déterminer des domaines de modélisation différents (tableau 1.3, adapté de CETUR, DDE du Rhône, 1994).

Cette décomposition est très schématique et plus qu'une frontière, c'est plutôt un continuum que l'on a sur les dimensions spatiale et temporelle. Les frontières ne doivent donc être considérées que de manière indicative.

	court terme < 5 ans	moyen terme 5 à 15 ans	long terme > 15 ans
quartier	-A- interactions négligeables	-B-	
zone	-A-B-	-B- prise en compte partielle des interactions entre les sous-systèmes urbains	-B-C-
agglomération	-B-	-B-C-	-C- les interactions sont prépondérantes pour l'avenir lointain

Tableau 1.3 : Echelles spatio-temporelles et modélisation

Sur le plan spatial, nous avons tout d'abord le quartier qui correspond à un périmètre restreint par exemple pour étudier le fonctionnement d'un carrefour ou la création d'une zone 30. Cet espace étant réduit, il est le plus souvent considéré presque comme un isolat par rapport au reste du territoire urbain. A l'opposé, on trouve l'agglomération dans sa globalité. Entre les deux, nous avons défini la zone, c'est-à-dire un espace suffisamment grand pour ne pouvoir être considéré comme un isolat dans son fonctionnement, mais de taille inférieure à l'agglomération. Nous ne chercherons pas à mieux le définir car les frontières sont largement arbitraires.

Sur le plan temporel, nous avons considéré également trois niveaux. Le court terme (inférieur à 5 ans) permet généralement de considérer (par simplification) que le système urbain est figé en dehors de la modification du sous-système des transports que l'on introduit et de ses conséquences sur les déplacements. Dans le moyen terme (5 à 15 ans), les sous-systèmes ne peuvent plus être considérés comme figés du fait de l'évolution de la ville. En revanche, les interactions entre les trois sous-systèmes restent limitées ou plus précisément, on se limite à des relations univoques sans considérer les rétro-actions. Enfin dans le long terme (plus de 15 ans), les trois sous-systèmes évoluent avec des interactions fortes entre eux (actions et rétro-actions entre les trois sous-systèmes). De nouveau, les frontières temporelles ne sont qu'approximatives.

Le croisement des deux dimensions conduit théoriquement à 9 situations. On peut en fait se ramener à 3 groupes :

A - interactions négligeables entre les trois sous-systèmes : les sous-systèmes des localisations et des transports sont considérés comme figés (à l'exception des modifications sur ces sous-systèmes que l'on modélise), seule l'action de la modification sur ces sous-systèmes est considérée. Ce premier cas correspond à une situation de court terme et à un espace urbain restreint. Concrètement, on trouvera, dans ce cas, des études de carrefour, des modifications de lignes de bus faiblement connectées à l'ensemble du réseau, des créations de zone 30... c'est-à-dire des études de portée limitée dans le temps et l'espace (schéma 1.1) ;

- B - les sous-systèmes des localisations et des transports ne peuvent plus être considérés comme figés. Ils interagissent entre eux, mais ces interactions sont considérées comme réduites et ne sont donc pas prises en compte. Les modifications de ces deux sous-systèmes qui ne sont plus limitées à l'objet d'analyse ont des incidences sur le sous-système des relations sociales. Mais seules les actions de ces sous-systèmes sont prises en compte, en considérant que les rétroactions du sous-système des relations sociales sur les deux autres sous-systèmes peuvent être négligées. Ce second cas, correspond principalement à des situations de moyen terme ou lorsque l'espace considéré ne représente qu'une partie de l'agglomération. Ce groupe correspond généralement à des études de prévision d'usage d'infrastructures nouvelles (ligne de transports collectifs en site propre, voie routière d'agglomération) ou à des études de plan de circulation ou de restructuration de réseaux de transports collectifs (schéma 1.2) ;
- C - les trois sous-systèmes évoluent et sont en interactions fortes. Ces interactions ne sont plus considérées seulement de manière univoque entre les trois sous-systèmes. Ce dernier cas correspond principalement à des études de long terme ou à des études portant sur l'ensemble du fonctionnement de l'agglomération. On devrait trouver dans ce groupe principalement des études de scénarios de politiques de déplacements ou des études de schéma de transport (les études de PDU ou de DVA devraient se trouver dans ce groupe, même si elles sont souvent positionnées dans le groupe B, car les outils ne sont pas très nombreux pour répondre aux problématiques du groupe C). Il ne s'agit généralement pas d'étudier une infrastructure particulière, mais plutôt d'analyser le fonctionnement de la ville sous différents scénarios de politique de déplacements ou différents schémas de voirie ou de transports collectifs (schéma 1.3).

Il doit être clair que cette présentation est en fait une représentation de la réalité. En effet quels que soient le terme ou le champ spatial, on peut donner des exemples d'interactions non univoques entre les trois systèmes. Le choix du groupe pour une étude est donc une réduction opérée par l'analyste en fonction de l'importance qu'il accorde à ces interactions (nous reviendrons sur la question de la réduction-simplification dans le prochain chapitre). L'intérêt de cette classification est de conduire à des outils le plus souvent différents pour chacun des groupes. Nous illustrons cette situation à travers la représentation des interactions entre les trois sous-systèmes.

3.3. Domaine d'utilisation des modèles à quatre étapes

Il est possible de représenter schématiquement (schémas 1.1 à 1.3) la nature des interactions correspondant à chacun des groupes.

3.3.1. Interactions faibles : modèle analogique, élasticité...

Dans le premier groupe (schéma 1.1), les interactions considérées sont très faibles. Les sous-systèmes des localisations et des transports sont considérés comme des isolats, non sujets à des évolutions tant sur le périmètre d'étude qu'au cours de la durée d'étude, à l'exception des modifications des sous-systèmes que l'on modélise. Aucune interaction entre ces deux sous-systèmes n'est donc prise en compte. Les interactions avec le sous-système des relations sociales sont uniquement univoques : la modification d'un des deux sous-systèmes (localisation ou transport) entraînant la modification du sous-système des relations sociales ou plus exactement pour ce qui concerne de sa traduction à travers le schéma d'activités réalisées.

Dans les études qui concerne ce groupe, on trouve presque uniquement des prévisions relatives à des actions sur le sous-système des transports (le court terme ne permet généralement pas d'envisager des actions non marginales sur le sous-système des localisations à l'exception de l'étude de générateurs de trafic comme un nouveau centre commercial). On est donc généralement en présence d'une interaction du type : la modification du sous-système des transports entraîne la modification du sous-système des relations sociales (en l'occurrence des déplacements puisque le modèle ne prend

généralement en compte que l'inscription spatio-temporelle de ce sous-système qui découle de sa confrontation avec les deux autres sous-systèmes).

On considère donc que l'on est en présence de réactions du type stimulus-réponse relativement simples pour lesquels un modèle analogique peut être utilisé, voire dans sa forme la plus simple des coefficients d'élasticité.

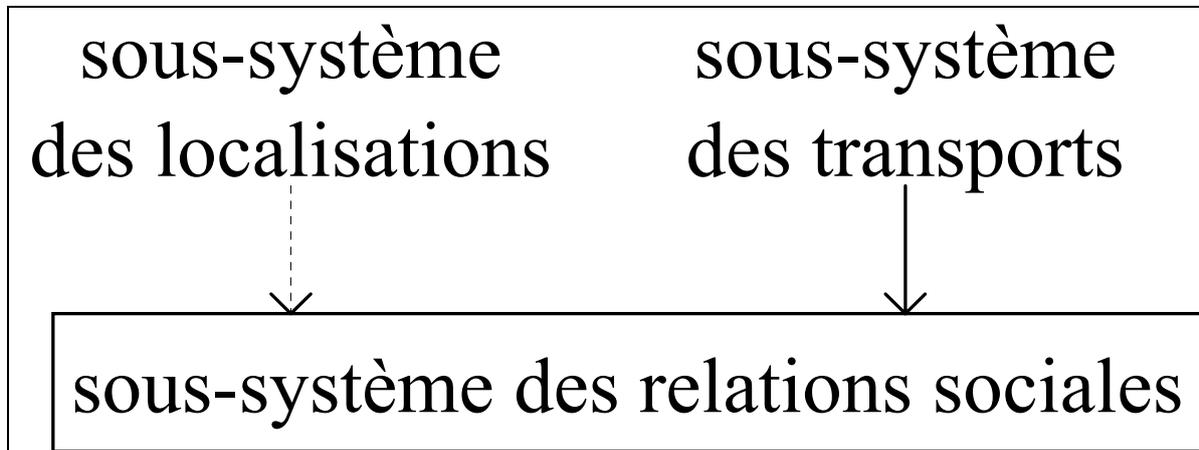


Schéma 1.1 : Groupe A - faible interaction

Pour l'illustrer, nous prendrons l'exemple de la modification du tracé d'une ligne de bus faiblement connectée au reste du réseau. Les correspondances de cette ligne ne sont donc pas trop nombreuses avec le reste du réseau, on la considérera le plus souvent comme un isolat. Ainsi par hypothèse, seule l'offre de cette ligne évolue au sein du sous-système des transports. De même, le sous-système des localisations sera considéré comme figé. En revanche, pour évaluer la fréquentation de la ligne, on procédera par analogie avec d'autres lignes s'il s'agit par exemple d'une modification de tracé. On retiendra alors une fréquentation équivalente si les caractéristiques sont très proches, ou bien, on retiendra des ratios d'usage équivalents. Dans ce cas, on peut identifier les générateurs concernés par la ligne et leur attribuer des ratios d'usage par personne en fonction de la nature du générateur, de sa position par rapport aux arrêts, des caractéristiques d'exploitation de la ligne de bus... (signalons qu'il s'agit de la méthode la plus couramment utilisée par les réseaux de transports collectifs dans le cas de modifications peu importantes concernant des lignes de bus ayant des échanges pas trop importants avec le reste du réseau). La méthode peut bien évidemment être améliorée pour tenir compte d'autres paramètres.

L'élasticité pourra être utilisée pour répondre à des questions telles que : que ce passe-t-il si la fréquence d'une ligne de bus augmente de 10 % ?, ou si les tarifs augmentent de 3 % ? On applique alors la formule mathématique suivante $\frac{\Delta A}{A} = e * \frac{\Delta B}{B}$ où e est le coefficient d'élasticité.

Dans ce groupe, on pourra également trouver d'autres modèles, comme les modèles de simulation microscopique : dimensionnement ou fonctionnement de carrefour, calcul de feux...

3.3.2. Interactions univoques : modèle à quatre étapes

Dans le second groupe (schéma 1.2) les trois sous-systèmes ne sont plus considérés comme des isolats. Ils interagissent entre eux, mais ces interactions sont considérées comme réduites et sont en grande partie univoques. L'évolution du sous-système des localisations est ici considérée comme exogène. Ses modifications sont des données d'entrée du modèle qui agissent principalement sur le sous-système de relations sociales. Une flèche (en pointillé) a été indiquée sur le schéma du sous-

système des localisations vers le sous-système des transports. Elle traduit non pas une prise en compte directe par le modèle de cette relation, mais plutôt une influence du premier sous-système sur la définition de l'offre de transport ou de la politique de déplacements mise en œuvre. Le sous-système des transports agit lui aussi sur le sous-système des relations sociales. Une flèche (pointillée) a été indiquée en retour. De nouveau, il ne s'agit pas d'une prise en compte directe par le modèle, mais plutôt d'une rétroaction au vu des conséquences sur le sous-système des relations sociales des évolutions des deux autres sous-systèmes. Cette rétro-action est traitée hors du modèle, mais peut donner lieu à de nouveaux scénarios qui seront modélisés.

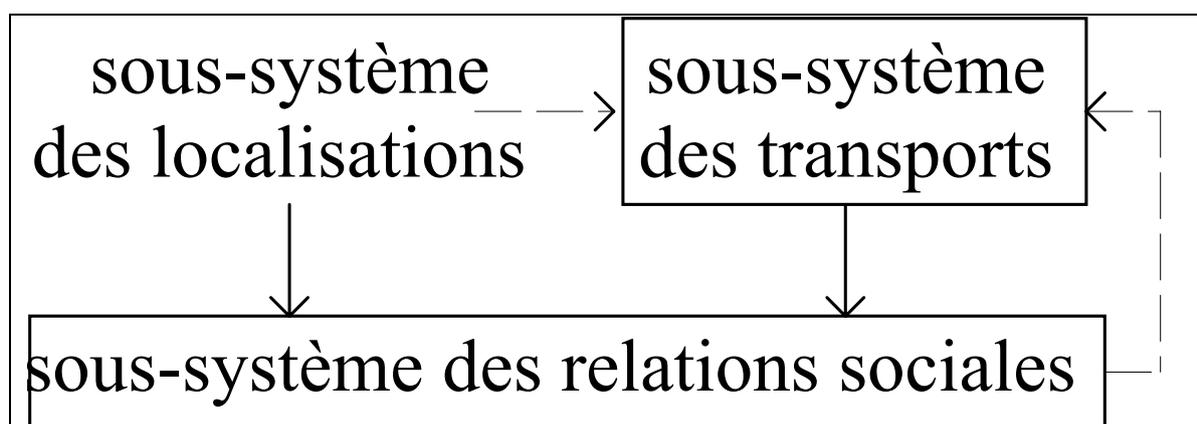


Schéma 1.2 : Groupe B - prise en compte partielle des interactions entre les sous-systèmes urbains

Ce second groupe est le domaine d'application privilégié de la famille de modèles englobés sous le terme générique de modèles à quatre étapes (ces modèles ne présentent pas tous une structure en 4 étapes, certaines étapes ne sont pas traitées par certains modèles (modèles mis en œuvre dans le logiciel TERESE de la SEMALY ou le logiciel DAVIS développé par l'INRETS (Institut National de Recherche sur les Transports et leur Sécurité) par exemple) ou à l'inverse certaines étapes peuvent être ajoutées comme le passage à la période de pointe.... Le traitement des étapes peut être séquentiel comme le propose la plupart des logiciels du marché ou non (comme POLYDROM par exemple qui regroupe les 3 premières étapes en une seule)). Nous illustrerons la nature des relations prises en compte dans chacune des étapes des modèles dans les chapitres 6 à 10.

3.3.3. Actions-rétroactions : modèle stratégique, transport-urbanisme...

Dans le troisième groupe (schéma 1.3) les trois sous-systèmes évoluent et sont en interactions fortes. De plus, ces interactions ne sont plus considérées de manière univoque entre les trois sous-systèmes. Dans cette configuration, aucun des trois sous-systèmes ne peut être considéré comme exogène à la modélisation. Il faut donc modéliser l'influence de chacun des trois sous-systèmes sur les deux autres et les rétro-actions provoquées en retour. C'est évidemment le cas le plus complexe. On trouve dans cette configuration les modèles stratégiques (Lichère, Raux, 1997a) ou encore les modèles que l'on regroupe sous le terme générique transport-urbanisme ou « *land-use* » (Clément, 1996).

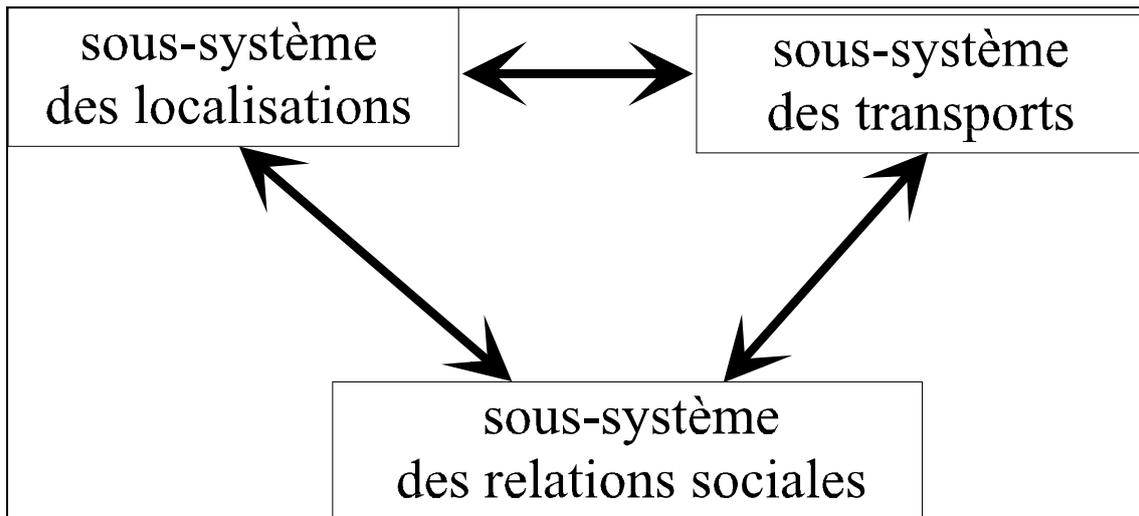


Schéma 1.3 : Groupe C - interactions prépondérantes entre les trois sous-systèmes

Ce manuel se situe dans le domaine de la famille des modèles englobés sous le terme générique de modèles à quatre étapes, domaine de prédilection du groupe B (toutefois une bonne partie de cet ouvrage est pertinent également pour les deux autres groupes A et C). Avant de présenter ces modèles, nous exposons quelques éléments de méthodologie.

Chapitre 2 - Cadrage méthodologique

Nous présentons tout d'abord notre conception de la procédure de prévision (section 1). Celle-ci isole volontairement le modèle mathématique. Nous insistons ainsi sur la place du modèle mathématique en temps qu'outil de la prévision, certes important, mais qui ne doit pas faire oublier l'importance des autres phases, sans lesquelles l'outil modèle n'est qu'un leurre ne pouvant conduire qu'à de piètres résultats. Cette séparation nous permet également de distinguer la phase de construction du modèle au sens de la représentation théorique d'un objet, de sa formalisation mathématique. Nous pouvons ainsi plus aisément rappeler « *qu'il n'y a pas de modèle sans objectifs... [car] on ne construit pas d'outil dont on se saurait à quoi il pourrait servir* » (Legay, 1990).

Dans une seconde section, nous rappelons que le modèle n'est qu'une « *représentation simplifiée de la réalité, destinée à mieux la comprendre ou à agir sur elle* » (Guitton, 1964). Nous illustrons ce double processus de simplification. Il consiste tout d'abord à privilégier certaines entrées, certaines sorties et certaines relations causales que l'on peut représenter à travers le cadre conceptuel. Ces relations sont ensuite formalisées mathématiquement (section 2).

Le modèle étant une représentation simplifiée, il convient de s'assurer de son opérationnalité (section 3). Nous présentons les trois conditions nécessaires et suffisantes de cette opérationnalité décrites par Bonnafous (1989). Ces conditions forment également la problématique du modèle, car il est aisé de montrer qu'elles sont contradictoires.

Enfin pour illustrer ces notions quelques peu théoriques, nous présentons un exercice : le calage d'un modèle prix-temps (section 4). Cet exercice nous permettra de passer en revue les différentes notions développées dans les trois sections précédentes.

1. SCHEMA DE LA PROCEDURE DE PREVISION

Le schéma (schéma 2.1) que nous présentons est évidemment quelque peu... schématique. La séparation des phases telles que nous l'illustrons n'est pas aussi nette, tout comme leur enchaînement. De même, la « mise à l'écart » de l'outil qui occupe ainsi d'une certaine manière un rang subalterne est critiquable. Nous développons plus loin les raisons de ces choix.

Avant de décrire chacune des phases, nous présentons l'architecture générale de ce schéma. Axhausen (1998) dans un article récent fournit une illustration de ce processus. Nous isolons tout d'abord dans une première phase, la définition des objectifs et des enjeux, phase qui devrait théoriquement être un préalable à toute étude. Dans cette même phase, nous incluons le choix du modèle au sens de la construction théorique censée représenter la réalité du phénomène que l'on cherche à prévoir. Ce choix, comme nous l'avons montré dans la section précédente, passe par la définition du champ spatio-temporel d'analyse et de la nature des interactions prises en compte entre les variables (d'entrée et de sortie du modèle). Ce choix peut être formalisé à travers la définition du cadre conceptuel d'analyse du problème. Cette étape est évidemment indispensable, bien que souvent mal identifiée dans nombre d'études.

La phase suivante permet le recueil des données nécessaires à une première validation du modèle théorique lors de l'opération de calage. Cette phase est évidemment fondamentale, car ces données sont considérées comme étant une mesure de la réalité que l'on cherche à reproduire. Il est évident que toute erreur rejaillira sur la capacité du modèle à prévoir le futur. Le recueil de données doit nécessairement suivre la définition des objectifs et non la précéder comme on peut

malheureusement l'observer fréquemment. C'est impératif lorsqu'il y a production de données. L'enquête doit alors fournir les données permettant de valider ou d'amender le modèle théorique proposé en testant certaines des hypothèses formulées (objet de la phase suivante). L'ordre inverse risque de conduire à l'impossibilité d'alimenter le modèle avec certaines données qui n'auraient pas été incluses dans l'enquête ou d'interdire le test de certaines hypothèses. Le modèle serait alors appauvri, sa capacité à répondre à certaines questions serait alors compromise.

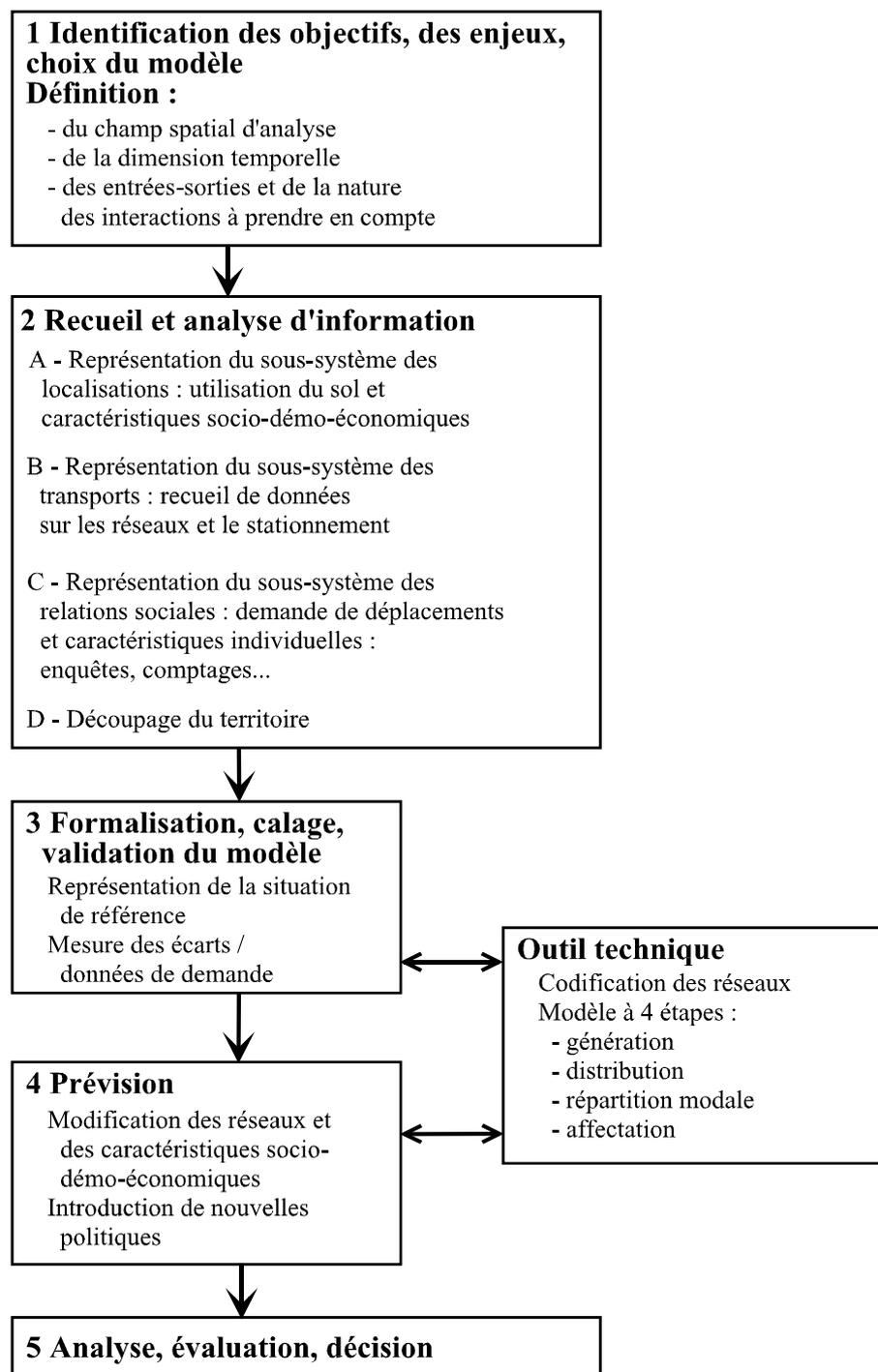


Schéma 2.1 : Procédure de prévision

La pratique oblige toutefois, dans certains cas, à amender cette séquence. En effet, lorsque l'analyste ne dispose pas des moyens pour produire ses propres données, il est obligé de composer avec les données disponibles. Dans ce cas, la production des données a de fait précédé la construction du modèle théorique. Toutefois, dans la séquence de travail de l'analyste, la réflexion sur les objectifs et enjeux doit rester première, car il n'y a pas de modèle sans objectifs (Legay, 1990). Par contre, il peut y avoir concomitance dans la construction du modèle théorique et l'examen de la disponibilité des données. Plus précisément, la construction du modèle théorique doit rester première, mais la confrontation avec les données disponibles peut conduire à modifier cette construction. Il y a fort à parier alors que cette modification est en fait une réduction de la portée du modèle. L'absence de prise en compte, faute de données, d'une dimension dont on pense qu'elle peut avoir une influence non marginale sur le phénomène étudié, risque effectivement de compromettre la capacité prévisionnelle du modèle. Elle peut dans ce cas conduire l'analyste à réviser ses objectifs pour en adopter de plus modestes.

Les opérations de formalisation, de calage puis de validation du modèle permettent de confronter la construction théorique avec les données disponibles. Elles visent à s'assurer de la capacité du modèle à reproduire la réalité, ou plutôt une certaine mesure de cette réalité, correspondant aux données d'enquêtes. Il s'agit donc de reproduire la situation observée par les données d'enquête. La mauvaise qualité de la reproduction peut amener à modifier le modèle ou à s'interroger sur la qualité des données et éventuellement à en produire de nouvelles. De nouveau, cette étape est fondamentale. Si le modèle n'est pas apte à reproduire la situation observée, il y a fort à parier que les écarts risquent de s'amplifier lors de la prévision. C'est à ce niveau qu'est mis au point l'outil technique : le modèle au sens plus classique du modèle mathématique. En fait à ce niveau deux opérations distinctes sont réalisées. La première consiste en la formulation mathématique du modèle, c'est-à-dire la mise en équations des relations entre variables. La seconde est le calage, c'est-à-dire la détermination des inconnues des équations mathématiques. Enfin, il reste à valider le modèle. Il s'agit de vérifier que le sens ou l'ordre de grandeur de certains indicateurs ou de certaines évolutions sont conformes à ce que l'on connaît du phénomène étudié. Il s'agit en quelque sorte de vérifier qu'au delà de la capacité à reproduire une situation donnée, le modèle est également capable de reproduire la dynamique du phénomène étudié.

Une fois le modèle calé et validé, c'est-à-dire apte à reproduire le phénomène étudié, le modèle mathématique peut être utilisé en prévision. La prévision ne prend donc place qu'au cours de la quatrième (et presque dernière) phase. Il s'agit alors de modifier les données d'entrée du modèle (des hypothèses sont également parfois faites relatives à l'évolution des coefficients ou des paramètres du modèle dans le temps (paramètres de distribution de la valeur du temps par exemple)) et grâce à l'outil technique d'obtenir en sortie les valeurs des indicateurs recherchés. La phase de prévision apparaît donc comme la plus simple de toute, si ce n'est que la modification (en fait prévision) des données d'entrée est un exercice périlleux. En effet, le modèle n'est apte qu'à prévoir ce qui se passera si le contexte évolue conformément à l'évolution des données d'entrée, toutes choses égales par ailleurs. Il est bien évident que ce « toutes choses égales par ailleurs », n'est qu'une fiction théorique, certes nécessaire, mais aussi simplificatrice. La qualité de la prévision dépendra bien évidemment de la qualité de la construction théorique, c'est-à-dire de sa capacité à représenter correctement le phénomène et ses évolutions, mais aussi de la qualité de la prévision des données d'entrée utilisées lors de la phase de prévision.

Enfin, vient la phase d'analyse des résultats, qui peut déboucher sur une phase d'évaluation puis de décision. A ce niveau, un bouclage est possible tant vers la quatrième phase pour tester de nouveaux scénarios que vers des phases amont si les résultats n'apparaissent pas satisfaisants.

Cette présentation est évidemment assez théorique. Il est fréquent que les objectifs soit en fait amendés lors du déroulement du processus. Par exemple, lors de la phase de calage, la confrontation avec les faits peut conduire à réduire la portée du modèle utilisé (nous en verrons une illustration lors

de l'exercice, section 4, p. 51). L'indisponibilité de certaines données peut conduire à modifier le modèle proposé... Dans la pratique, l'analyste est donc souvent confronté à de fréquents allers-retours entre ces trois phases. Ces différentes situations correspondent en fait à la problématique du modèle et aux réponses que l'on peut y apporter, que nous abordons section 3 (p. 45).

Même si cette présentation est schématique, elle nous semble importante, car elle permet d'insister sur l'importance de chacune des phases sur la qualité des résultats produits par le modèle. Trop souvent les trois premières phases sont traitées trop rapidement, notamment parce que les phases de production, de collecte des données et de calage sont très chronophages. De plus, elle nous permet d'insister sur la distinction entre le modèle au sens de la construction théorique, de sa formalisation mathématique, puis de son calage pour en faire un outil opérationnel pour la prévision.

Nous souhaitons à ce niveau faire une distinction entre modèles et logiciels de prévision de la demande, pour éviter de fréquentes confusions entre ces termes. Selon le Petit Robert (édition 1991), le logiciel est un « *ensemble de programmes, procédés et règles relatifs au fonctionnement d'un ensemble de traitement de l'information* ». Le logiciel apparaît donc plutôt comme une boîte à outils permettant à l'analyste de construire son propre modèle en utilisant certains des programmes et en choisissant certains des procédés ou règles disponibles dans le logiciel. C'est ainsi qu'il faut considérer la plupart des logiciels disponibles sur le marché international (comme VISEM-DAVISUM, EMME2, TP+, TRANPLAN, TRIPS...), qui ne sont pas à proprement parler des modèles, mais servent à la construction de modèles.

1.1. Phase 1 : objectifs, enjeux, définition du modèle

Cette première phase conduit tout d'abord à préciser les objectifs et enjeux relatifs au phénomène que l'on veut prévoir. C'est bien évidemment le préalable à toute étude, pas seulement de prévision. Si cela va de soi lorsqu'on l'énonce ainsi, force est de constater que la pratique connaît de nombreuses infractions à cette règle. Bien souvent, les objectifs et enjeux sont implicites ou encore flous. On conçoit pourtant aisément qu'une définition partielle des objectifs ne permettra pas de répondre précisément aux différentes questions conduisant au choix du modèle.

Nous avons indiqué dans le chapitre précédent la nature de ces choix, qui nécessite un positionnement des objectifs sur certaines dimensions : temporelle, spatiale et nature des interactions entre les trois sous-systèmes du système urbain. Il faut donc connaître le terme temporel de l'ensemble des questions auxquelles le modèle devra répondre.

De même, faut-il appréhender la zone d'influence des projets étudiés. L'étude d'un carrefour à feux ne mobilisera pas la même étendue géographique que l'étude d'une nouvelle voie en ville ou encore celle d'un contournement d'agglomération. De même, l'extension du stationnement payant dans une zone ne concernera pas forcément la même étendue géographique que l'augmentation des tarifs de stationnement pour l'agglomération... Le choix du champ spatial d'analyse consiste en quelque sorte à définir un isolat. Seuls les éléments présents dans cet isolat interagissent entre eux. En revanche, si les éléments extérieurs à cet isolat peuvent influencer sur ceux situés à l'intérieur, la rétroaction n'est pas prise en compte dans le modèle. Prenons un exemple pour illustrer le propos. Considérons une étude de PDU sur une agglomération. Le champ spatial retenu correspondra probablement à l'agglomération. De ce fait, le trafic de transit sera considéré comme un élément extérieur au champ d'étude. Cela ne signifie pas que l'on considère que le trafic de transit sera stable dans le temps. Il sera possible de retenir un taux de croissance du trafic de transit. Les effets de cette croissance sur les déplacements effectués dans la zone d'étude pourront être pris en compte. En revanche, la rétro-action ne sera pas prise en compte. C'est-à-dire que les conditions de circulation à l'intérieur de l'agglomération n'influenceront pas sur le volume du trafic de transit ou sur son taux de croissance. La croissance de ce dernier sera donc considérée comme exogène au modèle considéré. Il

est évident que ce choix est toujours réducteur. C'est donc l'analyse du contexte local et des objectifs de l'étude qui permettront de (in)valider ce choix.

Prenons un autre exemple pour insister sur cette dimension importante que constitue la définition de l'isolat d'analyse de l'objet d'étude. Pour une prévision du trafic sur le viaduc de Millau, on pourrait être tenté de retenir un périmètre d'étude à l'échelle de l'agglomération. Mais l'analyse risque fort d'invalider rapidement ce choix. Il n'est probablement pas très raisonnable de considérer que le trafic de transit tout particulièrement de longue distance Nord-Sud est exogène à l'investissement considéré. On peut penser que la suppression de ce goulot d'étranglement conduira une partie du trafic qui transitait par l'autoroute du Sud à se reporter sur un axe passant par le viaduc. L'hypothèse d'absence de rétroaction apparaît donc trop abusive obligeant à élargir le champ spatial d'analyse, au minimum pour la phase de choix d'itinéraire (nous revenons sur ces considérations dans le chapitre 10, p. 307).

Enfin, compte tenu de ces deux dimensions, il convient de sélectionner la nature des interactions entre les trois sous-systèmes qui peuvent influencer de manière significative sur l'objet (ou les objets) d'étude. Par exemple l'étude de l'impact à long terme d'un scénario de péage urbain de zone ne peut se passer d'une prise en compte des rétro-actions sur les localisations, alors que la mise à péage d'une infrastructure pourra dans certains cas en faire abstraction.... C'est donc à l'analyste en fonction de sa connaissance de l'objet d'étude de choisir son cadre d'analyse. Le choix de ne pas retenir certaines relations constitue bien évidemment une simplification dans la mesure où l'on pourra souvent exhiber le contre-exemple contredisant l'hypothèse ainsi prise. Mais il faut garder à l'esprit que le modèle, par définition, est une construction simplifiée de la réalité.

La construction du modèle théorique intervient après ces choix. Cela revient en quelque sorte à construire la maquette du modèle. Il s'agit tout d'abord de choisir les variables d'entrée du modèle, c'est-à-dire l'ensemble des variables pouvant influencer sur l'évolution de l'objet d'étude. Parmi celles-ci, on définira également les variables de commande, qui sont celles que l'utilisateur du modèle pourra faire évoluer pour simuler leurs effets sur les résultats du modèle. Les autres constituent des variables dont l'évolution est exogène (ou tout au moins considérée comme telle par simplification) à la sphère d'étude du problème. On trouve ensuite les variables de sortie. Elles constituent un groupe d'indicateurs caractérisant l'évolution du phénomène que l'on veut étudier plus un certain nombre d'indicateurs pouvant être utile à l'analyse ou au contrôle de la validité des résultats produits par le modèle (notamment un certain nombre de résultats intermédiaires). Enfin, il s'agit de déterminer la nature des interrelations entre l'ensemble des variables de ces deux groupes. Cette construction peut être schématisée par la présentation du cadre conceptuel d'analyse du problème (chapitre 4, p. 139).

L'absence d'objectifs clairs ne permet pas d'opérer ces choix. La chose est évidente pour les variables de sorties car elles doivent permettre de répondre aux... objectifs de l'étude. Mais il en est de même pour les variables d'entrée et les relations entre variables. Si une variable importante influant sur l'objet d'étude est omise, on peut émettre des doutes sur la capacité prévisionnelle du modèle si la grandeur associée à cette variable évolue dans le temps.

La construction du modèle théorique apparaît donc comme une étape fondamentale. Si l'analyste a une très grande maîtrise du phénomène étudié, il pourra proposer une maquette quasi définitive. En revanche, dans le cas contraire, la construction théorique permet de formuler un certain nombre d'hypothèses sur les variables pouvant avoir un rôle sur l'objet d'étude ou sur la nature des relations pouvant être mises en jeu. Il est important que ces hypothèses soient clairement formulées dès ce stade, car les phases suivantes permettront d'une part le recueil des données correspondantes et d'autre part le test, à l'aide des données recueillies, des hypothèses ainsi formulées sur l'influence des différentes variables (Daly, 1982). C'est à l'issue de ces deux phases que le modèle pourra être figé. En revanche, si certaines variables ne sont pas prises en compte, les données correspondantes ne seront

pas recueillies, oblitérant toute possibilité d'analyse ultérieure, sauf à reprendre les phases de recueil de données et de calage. Cette opération est toutefois généralement très coûteuse et chronophage.

1.2. Phase 2 : recueil et analyse d'informations

Cette phase permet le recueil des données nécessaires pour alimenter le modèle théorique. Ce recueil répond à plusieurs objectifs qui sont liés entre eux.

Le premier correspond aux tests des hypothèses qui ont été formulées lors de la construction du modèle théorique. Il s'agit de recueillir les données permettant de tester la pertinence de ces hypothèses. Le résultat de ce travail consiste à préciser la maquette du modèle. Cet objectif n'est pas toujours présent dans les études de prévision, notamment lorsque l'analyste a une bonne connaissance du phénomène étudié ou lorsqu'il utilise un modèle conçu préalablement ou encore lorsque son budget ne lui permet pas de consacrer du temps à cette étape.

Le second objectif concerne la formalisation mathématique des relations entre variables. Lors de la phase précédente, l'analyste a défini la nature des interactions existant entre les variables d'entrée et de sortie du modèle. Toutefois, la spécification mathématique de ces interactions n'a pas été abordée. Pour prendre un exemple, lors de la première phase, on établit que la motorisation est un facteur influant sur la répartition des déplacements entre la voiture et les transports collectifs. Toutefois, la forme mathématique de cette relation n'est pas précisée. Le recueil de données sur la motorisation des individus et la répartition des déplacements entre ces deux modes permettra de déterminer la forme mathématique de la relation entre ces deux grandeurs reproduisant au mieux le jeu de données. De nouveau, cette phase n'est pas présente dans toutes les études de prévision, notamment lorsque l'analyste utilise un modèle qui a déjà été spécifié préalablement pour d'autres études. Pour répondre à cet objectif, il faut recueillir des données relatives aux variables d'entrée et de sortie du modèle correspondant aux relations que l'on doit formaliser.

Le dernier objectif consiste à caler le modèle. Il s'agit cette fois d'estimer l'ensemble des inconnues du modèle de telle sorte que le modèle soit apte à reproduire la situation observée. Pour cette étape du travail, il faut disposer des données relatives à l'ensemble des variables d'entrée et de sortie du modèle correspondant à la situation observée. Cette étape est présente dans toutes les études de prévision sauf si l'analyste utilise un modèle déjà calé pour une étude précédente, sur la même zone géographique, avec le même zonage et la même période de référence. Si l'un de ces paramètres doit évoluer, il est souhaitable que la phase de calage soit actualisée.

Pour répondre à ces trois objectifs, qui bien évidemment ne sont pas exclusifs, il est généralement nécessaire de recueillir un volume important de données. Pour organiser ces données, nous reprenons l'analyse du système urbain dans sa décomposition en trois sous-systèmes.

La représentation du sous-système des localisations nécessite de recueillir des données sur l'utilisation du sol et les caractéristiques socio-démographiques et économiques des individus ou des ménages qui y résident. La nature des informations collectées dépend bien évidemment du modèle proposé, mais généralement il s'agit de réunir des données socio-démo-économiques sur chacune des zones du périmètre d'étude. En milieu urbain, pour des prévisions à moyen terme concernant l'agglomération, les données les plus couramment recueillies concernent : la population en nombre de ménages ou d'individus, le nombre d'actifs, le nombre de collégiens ou d'étudiants, la motorisation ou le revenu, le nombre d'emplois (avec une décomposition commercial - non commercial fréquente), le nombre de places dans les lieux d'études.... Ces données ne sont généralement pas produites pour les besoins de l'étude, mais collectées auprès des organismes producteurs de ce type de données (INSEE (Institut National de la Statistique et des Etudes Economiques), Agence d'Urbanisme et diverses administrations...).

La représentation du sous-système des transports conduit à collecter des données sur les réseaux et le stationnement (en théorie du moins, car peu de données sont disponibles sur le stationnement, ce qui limite sa prise en compte dans la modélisation). De nouveau la nature des données dépend du modèle choisi, dont découle notamment le mode de codification des réseaux. Le niveau de détail de cette codification dépend principalement du zonage utilisé. Il s'agit généralement de décrire les infrastructures avec leurs caractéristiques. Pour ce qui concerne le réseau viaire, les données généralement recueillies concernent la longueur, le sens de circulation, la capacité en nombre de véhicules par heure, la vitesse à vide, la courbe débit-vitesse permettant de prendre en compte la congestion, le coût d'usage.... Parfois, les caractéristiques des carrefours sont ajoutées. En ce qui concerne les modes collectifs, le mode habituel de représentation consiste à définir des lignes avec leurs caractéristiques : vitesse ou temps de parcours, fréquence, régularité.... Ces lignes peuvent utiliser les infrastructures viaires ou des infrastructures spécifiques. Ces données sont le plus souvent collectées auprès des organismes gestionnaires des différents réseaux. Des enquêtes spécifiques complémentaires sont parfois nécessaires (par exemple enquête temps de parcours, détermination de capacité réelle et non théorique (par exemple en cas de stationnement fréquent en double file la capacité réelle est réduite par rapport à une capacité théorique déduite du nombre de voies)...).

Le sous-système des relations sociales correspond à l'ensemble des relations et activités des acteurs du système urbain. Pour les individus, il transparait au niveau des pratiques urbaines, c'est-à-dire des comportements de mobilité. Il s'agit donc de recueillir à travers diverses enquêtes des informations sur les activités et les déplacements réalisés. Ces données doivent être recueillies simultanément à celles concernant les déterminants supposés de la demande de manière à pouvoir tester les hypothèses, formaliser les relations ou encore les caler. Certaines enquêtes permettant d'alimenter ce sous-système sont produites périodiquement : enquête ménages déplacements, cordon, d'utilisation des réseaux, comptages routiers.... Il est toutefois fréquent que cette information ne soit pas suffisante pour alimenter certaines études, qui nécessitent alors une production spécifique de données.

Nous avons ajouté dans ce bloc le découpage du territoire (nous précisons cette notion au chapitre 10) bien qu'il ne s'agisse pas en fait d'une donnée à recueillir. Le choix du découpage d'analyse est en fait le résultat d'un compromis entre les objectifs de l'étude de prévision dont découle un certain niveau de précision et les contraintes de coûts d'une part (le coût de production des données et de calage du modèle croissant plus que proportionnellement avec le nombre de zones) et des découpages préexistants pour les enquêtes réalisées périodiquement d'autre part (les découpages de chacune de ces enquêtes sont le plus souvent différents). Le choix du découpage est une tâche importante et souvent délicate conditionnant pour partie la qualité des prévisions du modèle.

1.3. Phase 3 : formalisation, calage, validation du modèle

Dans cette phase, on trouve en fait trois opérations distinctes, bien que fortement liées. La première concerne la formulation mathématique du modèle, c'est-à-dire la mise en équations des relations entre variables. La seconde est le calage proprement dit, c'est-à-dire la détermination des inconnues des équations mathématiques. Prenons un exemple pour illustrer le processus. Soit une grandeur A évoluant en fonction de plusieurs variables $X_1, X_2 \dots X_n$. On peut alors écrire :

$$A = f(X_1, X_2 \dots X_n). \quad (2.1)$$

La formalisation (on dit aussi la spécification) de l'équation consiste à poser une forme mathématique particulière de la fonction f.

La spécification doit bien être distinguée du calage. Par exemple si la spécification nous amène à proposer une forme linéaire, on pourra écrire :

$$A = \alpha_1 * X_1 + \alpha_2 * X_2 + \dots + \alpha_n * X_n + \beta \quad (2.2)$$

qui constitue la spécification du modèle.

Le calage consistera alors à estimer les inconnues du modèle, c'est-à-dire les coefficients $\alpha_1, \alpha_2 \dots \alpha_n, \beta$, par exemple par régression sur un jeu de données faisant intervenir les variables $A, X_1, X_2 \dots X_n$.

Le calage consiste à ajuster les diverses inconnues de la formalisation mathématique de telle sorte que le modèle soit capable de reproduire la réalité. Plus exactement, il s'agit de reproduire une certaine représentation de la réalité correspondant à la construction théorique élaborée dans la première phase. De plus, il ne s'agit que d'une certaine mesure de cette représentation correspondant aux données d'enquêtes collectées ou produites. Cette opération consiste donc à ajuster certains coefficients de telle sorte qu'en faisant « tourner » le modèle à partir des données d'entrée issues des enquêtes, le modèle produise des résultats similaires aux valeurs des variables de sortie (ces variables peuvent correspondre aux résultats finaux du modèle mais aussi à des résultats intermédiaires repérées dans les enquêtes).

Concrètement l'opération de calage peut le plus souvent être décomposée en deux parties (il s'agit ici d'une commodité de présentation, car ces deux parties sont fortement liées). La première correspond à la détermination des coefficients inconnus des diverses équations du modèle. La seconde concerne le calage (ou plus exactement la validation) de la codification des réseaux. En effet, il n'est généralement pas possible de recueillir des données détaillées pour chaque tronçon de chacune des infrastructures viaires ou de chacune des lignes des modes collectifs. Dès lors, la codification repose souvent sur la définition de classes de voies ou de lignes auxquelles sont associées certaines caractéristiques (par exemple un coefficient de régularité ou d'image pour une ligne de bus, une vitesse à vide ou une courbe débit-vitesse pour une voie routière...). Le calage doit permettre de valider ce classement. En général, si la première partie du calage peut faire appel à des techniques statistiques sophistiquées, elle nécessite un temps de travail plus court que la seconde (calage des réseaux). Le calage d'un réseau important peut nécessiter plusieurs mois de travail pour un chargé d'études.

Cette phase est fondamentale. On conçoit aisément que si le modèle n'est pas apte à reproduire la situation observée, il est peu probable qu'il soit apte à simuler une situation future. Toutefois, dans les faits, le modèle ne reconstitue que rarement (jamais probablement) complètement la situation observée. Plusieurs raisons permettent de l'expliquer. Le modèle n'est qu'une construction simplifiée de la réalité. C'est-à-dire qu'un certain nombre de facteurs pouvant influencer sur le phénomène étudié ne sont pas pris en compte (section 2.1, p. 42). De même, les enquêtes sont entachées d'un certain nombre d'erreurs que la statistique permet de définir lorsque les hypothèses nécessaires à ces calculs sont respectées (elles concernent principalement les méthodes d'échantillonnage (Ardilly, 1994 ; Richardson et al., 1995)). Il est donc « justifié », d'une certaine manière, que des différences subsistent entre les résultats du modèle et les données d'enquête. Toutefois, il est indispensable de mesurer ces écarts pour les analyser et tenter de les expliquer. Ils peuvent désigner une nécessité d'améliorer le calage, mais peuvent aussi désigner une insuffisance du modèle théorique. Il faudra dans ce cas reprendre le modèle théorique pour l'amender ou bien réduire la portée du modèle. Nous illustrons cette situation dans la section 4 (p. 51), pour aborder les réponses en cas d'insuffisance du modèle. L'exemple du calage du modèle prix-temps illustre mieux qu'un long discours l'importance de cette analyse des écarts entre les résultats du modèle et les données d'enquête.

A l'issue de cette étape, on dispose donc d'un modèle capable de reproduire la situation observée, qui peut donc en théorie être utilisé pour la prévision. En fait, compte tenu des réductions nécessaires (choix d'un nombre limité de variables, formalisation mathématique, calage...), des erreurs inhérentes aux recueils des données et des biais des procédures d'estimation, la capacité prévisionnelle du modèle ne peut être démontrée. Pour répondre à cette limite, il est nécessaire de procéder à des tests

de robustesse du modèle. C'est ici qu'intervient la troisième opération de cette phase 3 : la validation. Il s'agit en quelque sorte de valider le modèle. Plusieurs techniques sont disponibles. La plus courante consiste à faire varier certaines données d'entrée et à examiner les résultats en sortie pour apprécier leur plausibilité. Souvent ce travail passe par le calcul des élasticités directes et croisées sur les grandeurs étudiées. Il faut alors évaluer dans quelle mesure l'élasticité est conforme avec ce que l'on connaît du phénomène. Concrètement, il s'agit d'analyser le signe du coefficient et sa valeur (par exemple, il est raisonnable d'attendre que l'élasticité de la demande à l'offre soit positive, c'est-à-dire qu'une croissance de l'offre s'accompagne d'une croissance de la demande, et inversement qu'une décroissance de l'offre entraîne une décroissance de la demande). Cette analyse est évidemment délicate, car on ne peut disposer à ce stade de mesure effective de la validité de ces évolutions. De plus, certaines évolutions contre-intuitives peuvent s'avérer conformes à l'évolution effective des grandeurs correspondantes, nécessitant une certaine prudence dans cette phase d'analyse.

Dans certains cas, on n'exploite qu'une partie de l'information disponible pour le calage. La validation consiste alors à vérifier que le modèle est apte à reproduire les données observées disponibles et non utilisées dans la phase de calage. Cette validation reste toutefois statique et n'intègre pas la dimension dynamique du modèle nécessaire en prévision. Cette opération est assez fréquente pour la validation du codage du réseau. On procède au calage du réseau en prenant en compte par exemple 90 % des comptages et donc en laissant de côté 10 % des comptages. Une fois le réseau calé, on vérifie que le modèle reproduit correctement les comptages non utilisés.

Ainsi, la meilleure validation est a posteriori lorsque l'on est en mesure de confronter les prévisions du modèle avec l'observé à l'horizon de la prévision. On pourra objecter que c'est trop tard ! la prévision est déjà faite, les décisions auxquelles la prévision a contribué ont déjà été prises. Certes, mais l'exercice devrait être mené plus souvent car il permet de développer un regard critique sur la modélisation et donc de la faire progresser pour d'autres projets. Malheureusement cette opération est trop rarement menée (Mackinder, Evans, 1981 ; CEMT, 1982 ; Bonnafous, Tabourin, 1995 ; SEMALY, 1996 ; Gunn, Van der Hoorn, 1998 ; Quinet (1998) cite également d'autres travaux).

1.4. Phase 4 : prévision

Une fois le modèle calé, c'est-à-dire apte à reproduire la situation observée, le modèle mathématique peut être utilisé en prévision. Cette phase nécessite de modifier les données d'entrée du modèle pour obtenir en sortie les valeurs des indicateurs recherchés.

Cette phase apparaît donc comme relativement simple, si ce n'est que la modification des données d'entrée est en elle-même une prévision. Elle est donc sujette aux mêmes limites évoquées ci-dessus. De fait, les rares études visant à comparer les résultats des modèles de prévision de la demande de transport avec l'évolution effective des grandeurs modélisées, montre qu'une des principales source d'erreur vient de la prévision de l'évolution des données d'entrée du modèle (Mackinder, Evans, 1981 ; CEMT, 1982 ; SEMALY, 1996 ; Gunn, Van der Hoorn, 1998). Prévoir l'évolution de la population, des emplois pour chacune des zones d'une agglomération n'est pas un exercice très facile. On s'appuie en général sur les documents d'urbanisme, mais ceux-ci correspondent davantage à des orientations souhaitées par les collectivités qu'à une modélisation fine du fonctionnement du sous-système des localisations. Prévoir l'évolution des revenus ou de la motorisation n'est pas non plus un exercice facile en raison de la forte dépendance à de nombreux facteurs économiques, démographiques... qui ne sont pas eux-mêmes toujours très bien maîtrisés... On pourrait multiplier les exemples tant la prévision des données d'entrée est entachée d'incertitudes fortes. C'est cette incertitude qui commande de faire des études de sensibilité aux différents paramètres d'entrée du modèle afin d'encadrer l'univers probable d'évolution des grandeurs en sortie.

Une autre difficulté concerne la modification des réseaux. En apparence l'opération semble simple, il suffit de modifier les réseaux, tel qu'il est prévu de les faire évoluer. En fait, l'analyste se

heurte à deux difficultés principales. Nous avons mentionné que lors du calage, l'analyste devait adapter la codification des réseaux de telle sorte que cette codification corresponde au fonctionnement effectif de ces réseaux et non à ce qu'il devrait être en théorie sans phénomène perturbateur. Il faut donc être en mesure de codifier les nouveaux réseaux de manière cohérente à cette définition fonctionnelle et non pas à leur définition théorique. Prenons l'exemple du logiciel TERESE de la SEMALY (SEMALY, 2000) pour illustrer le problème. Dans la codification d'une ligne de transports collectifs, l'analyste doit fixer, entre autres, un paramètre : la pénalité d'accès (ce terme correspond notamment à l'image de la ligne pour l'utilisateur : type de matériel, confort, quartier traversé...). Elle est définie par l'appartenance à une classe de lignes. Le calage, combiné à la connaissance des lignes, permet de déterminer, pour chacune des lignes, à quelle classe elle appartient. En prévision, il n'est pas forcément facile de déterminer à quelle classe une nouvelle ligne peut être affectée. La seconde difficulté rejoint le problème évoqué pour la prévision des données d'entrée. Le nouveau réseau modélisé n'est pas toujours mis en place conformément à ce qui est prévu lors de l'étude de prévision. Par exemple, lors de la mise en service d'une nouvelle ligne de métro, une restructuration du réseau de surface est généralement prévue. Toutefois, il est fréquent que la consistance effective de cette restructuration ne soit pas totalement conforme aux prévisions initiales dans la phase projet, pouvant expliquer des différences de fréquentation sur le nouvel axe (SEMALY, 1996).

Enfin, une difficulté spécifique se pose lors de l'introduction d'une alternative nouvelle. Il peut s'agir de l'introduction d'un mode nouveau dans une agglomération : une ligne de métro dans une ville qui en était dépourvu, un métro automatique aux caractéristiques spécifiques, des véhicules électriques.... Mais ce peut être également l'introduction d'un nouveau mode de régulation des déplacements : introduction du péage urbain, tarification modulée dans le temps ou l'espace pour les transports collectifs, gestion du stationnement.... Dans ce cas, l'analyste se trouve face à une double difficulté. D'une part, la définition des paramètres de qualification de la nouvelle alternative est délicate en l'absence d'information ou d'expérience sur celle-ci. Il est alors fréquent de recourir aux méthodes de préférences déclarées permettant de tester et de qualifier des alternatives nouvelles ou des futurs inexpérimentés. D'autre part, l'introduction de cette alternative nouvelle est-elle de nature à modifier les caractéristiques des alternatives existantes et déjà calées ? (Hivert et al., 1988, nous revenons également sur cette question en examinant la propriété IIA (« *Independence from Irrelevant Alternatives* », Indépendance vis-à-vis des autres alternatives), chapitres 8, p. 233 et 9, p. 258).

Nous ne développons pas la dernière phase qui porte sur l'analyse et l'évaluation des résultats. Il y aurait beaucoup à dire sur cette phase et notamment sur les méthodes d'évaluation, mais nous sortirions de notre problématique.

2. LE MODELE : RESULTAT D'UN DOUBLE PROCESSUS DE SIMPLIFICATION

Comme nous l'avons déjà signalé, le modèle n'est qu'une « *représentation simplifiée de la réalité, destinée à mieux la comprendre ou à agir sur elle* » (Guitton, 1964). Dans un modèle mathématique, cette simplification peut être décomposée en deux opérations. Tout d'abord au niveau de la construction du modèle théorique qui conduit à ne retenir qu'une partie des variables et des relations causales pouvant influencer sur le phénomène étudié (section 2.1). Ensuite la formalisation mathématique conduit à une seconde simplification pour des questions de choix de fonctions mathématiques, de mesurabilité des variables, de cohérence interne ou encore de complexité des calculs (section 2.2).

2.1. Première simplification : la construction théorique

« *Il n'est pas d'objet économique qu'un nombre restreint de concepts permette d'élucider tout entier* » (Bonnafous, 1989). Cette remarque n'est pas probablement spécifique au champ de l'économie. Elle signifie que la compréhension d'un objet passe par la prise en compte d'un nombre très élevé de dimensions. La chose est connue depuis longtemps pour l'analyse de la mobilité. Une

lecture rapide de la littérature permettrait de mettre en évidence plusieurs dizaines de déterminants de cette mobilité (déterminants économiques, démographiques, individuels, du mode de vie, du ménage, des structures spatiales urbaines, de l'offre de transport...).

C'est pour relever « *le défi du multidimensionnel* » (Bonnafous, 1989), que l'analyste est amené à réduire son objet. Cette réduction revient à sélectionner quelques variables parmi l'ensemble de celles pouvant influencer sur l'objet étudié. Il s'agit donc bien d'un choix fait par l'analyste de privilégier certaines variables parmi un ensemble plus vaste. Ce choix est opéré à partir de la connaissance acquise sur l'objet qui découle pour une bonne part de l'analyse de données d'enquête. Ce choix revient à considérer que les autres variables n'ont aucune influence, ou tout au moins que l'on fait l'hypothèse que l'on peut négliger l'incidence des autres variables sur l'objet d'étude dans le contexte dans lequel le modèle doit être utilisé.

On peut formuler autrement ce choix. Il revient à formuler une clause *ceteris paribus*. C'est-à-dire que l'on considère que toutes choses sont égales par ailleurs ou plus exactement, si on se situe dans une perspective dynamique, que tout ce qui n'est pas pris en compte continue à évoluer comme dans le passé. Il est clair qu'il s'agit là d'une réduction forte de l'objet d'étude, car le « *prévisionniste ne peut le plus souvent pas être certain que le système qu'il a isolé contient effectivement la totalité des éléments à prendre en considération. Au mieux, il aura pu constater, par observation du passé, que ce système constitue une approximation satisfaisante, mais il ne peut être certain que le système réel continuera à fonctionner comme s'il était isolé* » (Matalon, 1978).

Comme nous l'avons précisé précédemment, la construction théorique nécessite le choix des variables prises en considération, mais elle nécessite également la détermination des relations de causalité entre les variables d'entrée et de sortie du modèle. L'analyste est ici amené à faire de nouveau des hypothèses sur la nature des relations entre ces variables et sur la séparation des différents facteurs.

Nous situant dans une perspective prévisionnelle, il faut faire des hypothèses sur l'évolution des relations prises en compte. Quelle que soit la forme de ces relations, l'analyste est conduit à postuler l'invariance de certaines relations qui deviennent de facto des lois. Par exemple, dire que la motorisation va croître n'a d'intérêt pour la prévision que si l'on définit son taux de croissance qui constituera alors un invariant pour la période considérée (hypothèse des premiers modèles utilisée dans le cadre des EPIT (SETRA, 1973)). Si on a des raisons de penser que le taux de croissance ne restera pas constant, on peut définir la loi de son évolution, en supposant par exemple une relation avec le revenu (hypothèse fréquente de nombreux modèles). C'est cette nouvelle relation qui deviendra alors invariante. On pourrait complexifier encore la relation en faisant intervenir le temps, ou en introduisant des phénomènes asymptotiques (saturation du marché)..., on en revient toujours à l'expression d'un invariant de relation. Pourtant, rien ne nous permet d'affirmer que ces lois resteront valides quelle que soit l'évolution du contexte environnant l'objet d'étude. En fait nous nous retrouvons placé dans la même situation que précédemment d'hypothèse d'indépendance vis-à-vis de l'extérieur du système considéré pour l'objet.

« *Appliquer un modèle à un objet, c'est faire implicitement une prévision sur la stabilité des conditions qui permettent son application.* » L'apparition d'événements « *qu'il s'agisse d'une conséquence de la transformation de l'« extérieur » ou d'un produit inattendu de la dynamique interne de l'objet lui-même, doivent être prévus si on veut que la prévision souhaitée soit correcte. Or, par définition, leur occurrence ne peut être déduite du modèle lui-même, à moins d'élargir celui-ci, ce qui ne fait que déplacer le problème* » (Matalon, 1978).

L'analyste doit donc faire une série de simplifications (le terme simplification est peut être parfois un peu excessif. Peut-être serait-il plus adéquat dans certains cas de parler d'approximations de ce que l'on soupçonne de la réalité) lors de la construction de son modèle théorique. Il doit tout

d'abord faire le choix des variables d'entrée, ce qui l'amène à considérer que l'influence des autres variables sur l'objet d'étude peut être négligée. Ensuite il définit la nature des relations entre les variables d'entrée et de sortie du modèle, ce qui l'amène à déterminer des invariants de relation. Ces hypothèses reviennent à considérer un système fermé, isolé du monde extérieur, ou plus exactement qui continue à subir les influences de cet extérieur sans pouvoir influencer sur lui.

2.2. Deuxième simplification : la formalisation mathématique

La formalisation mathématique ou encore la spécification mathématique consiste à choisir la forme mathématique traduisant les relations entre variables au sein du modèle. Elle entraîne également un processus de simplification. Plusieurs raisons peuvent être avancées :

- si l'on veut appliquer le modèle, les variables utilisées par les fonctions mathématiques doivent pouvoir être mesurables, quelle que soit la mesure utilisée. Pourtant, les variables utilisées ne sont pas toujours facilement mesurables. Prenons un exemple pour illustrer le problème. Dans un modèle d'affectation (choix de l'itinéraire), on fait le plus souvent l'hypothèse que l'individu connaît l'ensemble des chemins possibles pour aller d'une zone i à une zone j et qu'il est capable de les comparer pour en choisir un (en général le plus court). En fait, on sait bien que cette hypothèse ne correspond pas à ce que l'on connaît des comportements individuels. En fait, les usagers ne connaissent pas tous les chemins et ils ont une certaine perception de chacun des chemins qu'ils connaissent qui n'est pas forcément la même que celle d'un autre usager. On est donc en quelque sorte en présence d'un filtre déformant. L'introduction de ce concept de filtre déformant de l'offre de transport pose un évident problème de mesure. Comment mesurer la perception d'une ligne de bus, d'un trajet de marche à pied, l'attente à un arrêt ? Comment mesurer la différence de perception entre deux itinéraires routiers ? Comment mesurer le filtre qui traduit la connaissance partielle des itinéraires possibles... L'introduction de la fonction de coût généralisé est la manière la plus courante de répondre au problème. Elle reste toutefois imparfaite, notamment en ne traitant pas le problème du filtre. De plus, la formalisation de cette fonction, puis son calage ne sont pas aisés. D'autres techniques à partir d'enquêtes de préférences déclarées permettent de préciser la mesure du coût généralisé, mais celle-ci reste le plus souvent incomplète. Enfin pour tenir compte de la variabilité de la perception, une approche stochastique du coût généralisé est parfois introduite (chapitres 5, p. 154 et 9, p. 246). Mais quelle que soit la sophistication de la mesure de la fonction, celle-ci reste une approximation de la réalité de la situation de chacun des individus ;
- de nombreuses fonctions mathématiques utilisées nécessitent des variables continues. Pourtant, certaines variables ne sont pas continues ou bien les données disponibles concernant ces variables ne se présentent pas de manière continue. Le cas le plus fréquent est celui du revenu. Par exemple, le modèle QUINQUIN (Bouf, 1989 ; Tabourin, 1989) s'appuie sur une relation liant mobilité et revenu. Pourtant, lorsque le revenu est disponible, il ne l'est généralement pas en tant que variable continue, mais plutôt par classes de revenu (car le refus de réponse est généralement moins grand quand on demande à l'enquêté de se positionner dans une classe, que lorsque l'on demande un montant). Il est bien évidemment toujours possible de lisser la courbe, mais il est nécessaire de faire des hypothèses de distribution au sein de chacune des classes pour cela ;
- le choix de la fonction mathématique ne correspond pas exactement à la distribution effective dans la population. Il est fréquent que l'on choisisse une forme linéaire, éventuellement après transformation des variables à l'aide de fonctions logarithmiques, exponentielles... ou des distributions statistiques connues (distribution normale, de Gumbel, de poisson...). Ces distributions sont en effet plus faciles à formaliser et à caler. Toutefois, les données disponibles ne permettent pas toujours un très bon ajustement aux distributions choisies ;
- les données ne sont pas toujours disponibles pour alimenter le modèle pour une variable ou pour une relation. C'est souvent le cas du stationnement dans un modèle de choix modal. Il est connu que la disposition d'une place de stationnement sur le lieu de travail est un facteur

- explicatif important du choix du mode (Massot, Orfeuill, 1991-93). Mais l'information est souvent indisponible, obligeant l'analyste à modifier son modèle théorique, ce qui conduit le plus souvent à un appauvrissement de ce dernier ;
- les variables explicatives sont fréquemment corrélées entre elles. On sait que des variables comme le revenu ou la motorisation influent sur l'évolution du trafic d'échange et de transit. Si l'on retient les deux variables pour expliquer l'évolution de ce trafic, par exemple dans un modèle linéaire estimé par régression, l'estimation des coefficients risque d'être peu précise du fait de la forte corrélation entre les deux variables. Face à ce problème, on est souvent conduit, soit à ne retenir qu'une seule des deux variables, soit à transformer l'une des deux variables pour réduire la corrélation ou encore à construire une variable synthétique agrégeant les deux variables... mais dans tous les cas, on procède à une simplification du modèle explicatif.

Cette présentation illustre bien le processus de simplification mis en œuvre lors de la formalisation mathématique du modèle théorique. Elle souligne également l'imbrication entre les différentes phases de la construction du modèle mathématique. Les exemples donnés démontrent la nécessité de traiter la phase de construction du modèle théorique avant celle de collecte ou de production des données, mais aussi l'importance de la qualité de la production des données. Des données biaisées peuvent conduire à choisir une fonction mathématique différente de celle qui aurait été retenue sur des données non biaisées, car elle permet un meilleur ajustement. Lors de la construction de son modèle théorique, l'analyste doit tenir compte des problèmes de formalisation. Il doit chercher à privilégier dans la mesure du possible des variables posant moins de problèmes de mesure. De même, il doit privilégier des relations pour lesquelles on estime pouvoir trouver des formulations mathématiques permettant un bon ajustement. Ce choix risque toutefois de se faire au détriment de la pertinence du modèle. Cette contradiction est une des illustrations de la problématique du modèle.

3. OPERATIONNALITE ET PROBLEMATIQUE DU MODELE

Si le modèle est le résultat d'un double processus de simplification, il est légitime de se poser la question de la qualité du modèle. Nous en proposons tout d'abord une définition. « *Un modèle apte à réaliser les performances que l'on attend de lui sera dit opérationnel. L'opérationnalité du modèle constitue donc sa qualité fondamentale* » (Bonnaïous, 1989). L'opérationnalité d'un modèle ne peut donc pas être définie de manière générale. Elle ne peut être caractérisée qu'à partir d'une définition des objectifs qui sont assignés au modèle. Le modèle peut donc s'avérer opérationnel pour répondre à certaines questions formulées dans un certain contexte et ne plus l'être du tout pour répondre soit à d'autres questions, soit aux mêmes questions formulées dans un contexte plus large. Nous en verrons une illustration dans l'exercice présenté plus loin, mais un exemple permet de préciser le propos. Prenons le cas d'un modèle de partage modal entre la voiture et les transports collectifs. Si le modèle est opérationnel pour répondre à des préoccupations de modifications de tracé ou de restructuration du réseau de transports collectifs, il n'est pas certain qu'il le reste si l'objectif devient la simulation de nouvelles lignes avec des technologies nouvelles. En effet, les lois de comportement de choix entre les transports collectifs et la voiture peuvent se trouver modifiées par l'introduction d'un mode collectif radicalement nouveau.

Bonnaïous dans le même ouvrage précise qu'il y a trois conditions à l'opérationnalité du modèle : pertinence (section 3.1), cohérence (section 3.2) et mesurabilité (section 3.3). Ces trois conditions sont nécessaires et suffisantes. Toutefois, elles sont contradictoires et constituent la problématique du modèle (section 3.4). Il est alors légitime de s'interroger sur la nature des erreurs commises lors de la conception ou de l'utilisation d'un modèle (section 3.5).

3.1. Pertinence

La pertinence correspond à la capacité du modèle à être conforme à ce que l'on croit savoir de la réalité, de l'état des choses. On conçoit aisément que cette condition soit nécessaire à l'opérationnalité du modèle. Un modèle non pertinent ne saurait être conforme à l'état des choses. L'absence de pertinence trouve principalement sa source dans deux dimensions. La première concerne le choix des variables explicatives. La seconde concerne les relations entre variables. Dans les deux cas l'oubli de certaines variables ou relations ou les problèmes de définition ou de spécification conduisent à une réduction de la pertinence. Dès lors, on voit mal comment le modèle pourrait produire une prévision conforme à ce que sera l'état des choses à l'horizon de la prévision. Au contraire, il y a tout lieu de croire que l'écart s'amplifiera avec la prévision, sauf « heureux » concours de circonstance.

Cette définition mérite une précision. *La conformité à ce que l'on croit savoir de l'état des choses* n'apparaît pas comme une définition très précise. Elle traduit le fait que l'on ne connaît pas la réalité, que « *la réalité que le modèle s'efforce de représenter désigne seulement le réel variable des étapes successives de la science, un état des choses constitué de ce que l'on croit savoir de l'objet du modèle, c'est-à-dire d'une certaine accumulation de connaissances, parmi lesquelles un ensemble organisé d'informations. La pertinence est donc en somme, la qualité de l'approximation d'une réalité très provisoire par sa représentation simplifiée* » (Bonnafous, 1989). La pertinence du modèle ne pourra donc s'apprécier que dans la mesure où la construction simplifiée de l'objet d'étude aura été explicitée, ce qui souligne une fois de plus, s'il en était encore nécessaire, la nécessité de bien expliciter la première phase de ce que nous avons appelé la procédure de prévision (schéma 2.1).

3.2. Cohérence

La cohérence du modèle recouvre deux notions : tout d'abord une condition de cohérence interne du modèle, ensuite une condition de cohérence d'objectifs du modèle.

La cohérence interne du modèle correspond à une condition de cohérence logico-mathématique du modèle. C'est-à-dire que le modèle ne doit pas comporter de contradictions internes que ce soit au niveau de la définition des variables, des valeurs prises par celles-ci ou au niveau des relations de causalité retenues. De plus, le modèle doit respecter les hypothèses d'application des modèles mathématiques. Cette première condition semble bien être le moins que l'on puisse attendre d'un modèle. Toutefois, sa mise en œuvre peut parfois soulever quelques difficultés.

Prenons deux exemples pour illustrer cette condition de cohérence interne. Dans le cas du modèle linéaire, il est souhaitable que les variables explicatives soient linéairement indépendantes pour appliquer l'ajustement au moindre carré ordinaire. Dans le cas contraire, il y a indétermination dans la détermination des coefficients de la régression. De même, les aléas doivent être statistiquement indépendants entre eux. Ils doivent avoir une même variance σ^2 et une covariance nulle. C'est l'hypothèse d'homoscédasticité. Des tests existent pour vérifier ces hypothèses (Cohen, Pradel, 1993 ; Wonnacott, Wonnacott, 1991, chapitre 6, p. 174). Le second exemple concerne l'estimation du coût généralisé. Dans un modèle à quatre étapes, le coût généralisé est utilisé dans toutes les étapes à l'exception de la génération. L'application de la condition de cohérence interne signifie qu'une même catégorie de déplacements doit avoir tout au long des étapes le même coût généralisé. Pourtant, à l'issue des trois étapes, le coût généralisé d'un déplacement de la zone i à j n'a pas la même valeur, plus particulièrement pour la voiture. Cette situation résulte du fait que le temps pour aller de la zone i à j dépend du trafic sur l(es) itinéraire(s) emprunté(s) pour aller de i à j . Ce trafic n'est connu qu'à l'issue de la phase d'affectation, ne permettant pas sa prise en compte lors des deux étapes précédentes. Bien évidemment, il existe une parade qui consiste à faire un bouclage en reprenant de nouveau les trois phases autant de fois que nécessaire en attendant la convergence du processus (qui heureusement est en général assez rapide). Toutefois, la déformation de la matrice origines-

destinations par ce processus itératif peut parfois poser d'autres problèmes de cohérence entre les étapes de distribution et de répartition modale (Morellet, Marchal, 1999).

Cette condition de cohérence interne reste toutefois le plus souvent soluble, fut-ce au prix de calculs un peu plus longs ou complexes. Un examen attentif du modèle permet d'ailleurs le plus souvent de détecter les problèmes de cohérence interne.

La seconde condition est celle de la cohérence d'objectifs du modèle. Cela signifie que le modèle « doit reposer sur une organisation logico-mathématique, qui soit cohérente avec son objectif : pour rendre compte d'un équilibre général, le modèle du même nom devra comporter autant d'équations qu'il y a de quantités de bien à échanger sur les marchés et de prix de ces biens. Il faudra de surcroît que ces équations soient indépendantes, faute de quoi quelques-unes des inconnues du problème resteraient inconnues » (Bonnafous, 1989). L'ajout de nouvelles variables explicatives pour répondre à un objectif particulier pourra donc nécessiter l'introduction d'équations nouvelles ou d'hypothèses supplémentaires pour lever une indétermination. Elle nécessitera également le plus souvent des tests d'indépendance entre variables.

Cette condition est bien évidemment nécessaire à l'opérationnalité du modèle. Sa mise en défaut pourra conduire à des contradictions dans les résultats du modèle pour ce qui est de la cohérence interne, et à des indéterminations pour l'autre condition de cohérence. Dès lors, c'est l'opérationnalité qui sera prise en défaut du fait de la non conformité avec la réalité ou de l'impossibilité de produire des prévisions.

3.3. Mesurabilité

La mesurabilité constitue la troisième condition de l'opérationnalité du modèle. Cette exigence se décline « selon trois registres :

- la mesurabilité des relations causales ;
- la mesurabilité des quantités mises en jeu dans le modèle ;
- la mesurabilité des paramètres de ses équations » (Bonnafous, 1989).

La mesurabilité des relations causales correspond à la nécessité de pouvoir les formaliser. Une absence de formalisation ne permettrait pas d'utiliser le modèle en prévision. Quel serait l'intérêt en matière de prévision d'une relation causale du type « l'accroissement de la motorisation entraîne une croissance du nombre de déplacements en voiture particulière », pour laquelle on ne pourrait proposer une fonction reliant motorisation et mobilité automobile. Tout au plus, on pourrait prédire une croissance de la circulation automobile en cas de croissance de la motorisation, mais on serait bien incapable de la quantifier et ainsi de répondre à l'objectif.

La mesurabilité des quantités mises en jeu dans le modèle est bien évidemment nécessaire. La non-mesurabilité d'une des variables entrant en jeu dans le modèle est problématique à plus d'un titre. Tout d'abord elle ne permettrait pas de tester l'opérationnalité du modèle. Par exemple si celui-ci établit que telle variable doit croître sous l'effet d'une variation d'une autre variable, le moins que l'on puisse attendre est de pouvoir confronter cette relation à l'état des choses. La non-mesurabilité risque fort de rendre l'entreprise assez délicate. De plus, la non-mesurabilité de la variable d'entrée ne permettra pas d'établir la valeur de la variable de sortie.

Enfin, la mesurabilité des paramètres de ses équations est également indispensable. L'indétermination d'un coefficient ne permettra pas d'estimer le modèle qui restera... indéterminé. Par exemple dans un modèle linéaire, l'impossibilité de déterminer la valeur des différents coefficients, rend la formalisation mathématique de peu d'intérêt. Cette indétermination ne permettra pas de tester son opérationnalité, face à l'impossibilité dans laquelle on se trouve de le confronter à l'état des

choses, pas plus qu'elle ne permettra la prévision de certaines grandeurs pour répondre aux objectifs visés.

C'est pour répondre à ces trois dimensions de la mesurabilité que la production de données est nécessaire. En l'absence de données d'enquête, les quantités des variables risquent fort de ne pouvoir être définies. Mais surtout, la formalisation ne pourra pas être confrontée à la « réalité » représentée par le jeu de... données. Enfin, la détermination des coefficients du calage sera impossible. Certes, dans ce cas, il est toujours possible de reprendre la formalisation et les coefficients de calage retenus lors d'un autre exercice de modélisation, mais alors de nouveau l'opérationnalité du modèle ne pourra être éprouvée, faute de confrontation possible à la « réalité ».

3.4. Problématique du modèle

Les trois conditions de pertinence, cohérence et mesurabilité sont donc nécessaires à l'opérationnalité du modèle, et elles sont aussi suffisantes (Bonnafous, Gérardin, 1976). Mais elles sont contradictoires. Ainsi l'amélioration de l'une des conditions se fait le plus souvent au détriment des deux autres, constituant ainsi la problématique du modèle. Examinons tour à tour ces contradictions.

La recherche d'une plus grande pertinence passe souvent par l'ajout d'un plus grand nombre de variables explicatives. En effet, la lecture de la littérature sur l'analyse de la mobilité illustre l'étendue des facteurs influant sur le nombre de déplacements quotidiens, sur le choix du mode de transport, sur le choix des itinéraires.... L'introduction de l'ensemble de ces paramètres dans le modèle risque fort de se faire au détriment tant de la cohérence que de la mesurabilité.

L'introduction de l'ensemble des variables nécessitera l'introduction des relations correspondantes au sein du modèle mathématique. La cohérence d'objectifs imposera des tests beaucoup plus lourds de robustesse du modèle pour vérifier que le modèle correspond bien à l'état des choses lorsque chacune des variables varie indépendamment ou conjointement. La cohérence interne risque également fort d'être mise à mal, notamment sur le plan de l'indépendance entre les variables lorsque l'hypothèse est nécessaire.

La contradiction est encore plus évidente sur le plan de la mesurabilité. Elle peut être déclinée sur les trois registres évoqués plus haut :

- mesurabilité des relations causales. Les données doivent être disponibles pour formaliser l'ensemble des relations. L'absence de données pour choisir la forme fonctionnelle correspondant le mieux à l'état des choses, imposerait de faire des hypothèses sur la fonction mathématique à choisir, ce qui risque de compromettre tant la pertinence que la cohérence d'objectifs ;
- mesurabilité des quantités mises en jeu. De nouveau, la multiplication des variables peut poser des problèmes de mesurabilité de ces variables. Reprenons, le problème du concept de filtre déformant de l'offre de transport. L'introduction du coût généralisé permet de rendre mesurable la déformation de l'offre de transport par la perception que les individus en ont (encore que cette notion ne soit pas exempte de critiques sur le plan de la pertinence, et sur la mesure moyenne qui en est généralement faite). En revanche, la notion de filtre pour traduire la connaissance partielle de l'offre de transport par les individus continue à poser des problèmes redoutables de mesurabilité ;
- mesurabilité des paramètres de ses équations. Là encore, l'analyste doit pouvoir disposer d'un jeu de données suffisant mettant en œuvre l'ensemble des variables en relation afin de pouvoir caler les fonctions.

L'accroissement de la pertinence risque donc fort de se faire (dans la mesure où elle est possible) au prix fort, face à l'inflation des données nécessaires. En effet, la production des données

est un poste budgétaire important dans le cas où une telle production est nécessaire face à l'absence de données disponibles.

La recherche d'une plus grande mesurabilité se fait également souvent au détriment des autres conditions. Prenons l'exemple du revenu que nous avons déjà évoqué. Le revenu est rarement demandé directement dans une enquête, car les refus sont trop fréquents, ce qui donne une piètre mesure du revenu effectif. De ce fait, on utilise parfois des classes de revenu en demandant à l'enquêté non pas de donner son revenu mais la classe à laquelle il appartient. Ce subterfuge permet de réduire le nombre de refus. On définit alors une nouvelle variable, la classe de revenu plus facilement mesurable. En revanche, elle pose des problèmes de cohérence interne. Pour répondre aux objectifs de quantification du modèle, il est généralement nécessaire de disposer d'un revenu variant continûment et non par classe. La solution consiste alors à lisser la courbe. Il faut pour cela introduire des hypothèses sur la distribution des revenus au sein de chacune des classes. Mais la pertinence de ces hypothèses n'est pas éprouvée, faute de la connaissance de la distribution effective... des revenus !!

L'introduction de la classe de revenu réduit les refus, mais ne les annule pas. De plus, selon les niveaux de revenus et les professions, les informations délivrées par les enquêtés sont fréquemment volontairement erronées à la hausse ou à la baisse. L'information reste donc de qualité médiocre. C'est pourquoi, il est fréquent de remplacer le revenu par l'appartenance à une catégorie socio-professionnelle. On dispose cette fois d'une variable aisément mesurable, qui de surcroît est disponible dans les données de l'INSEE, permettant des redressements à différents niveaux géographiques. Par contre, il faut bien convenir qu'il s'agit d'un substitut approximatif du revenu dont la pertinence est plus que discutable.

Enfin, la recherche d'une plus grande cohérence ne fait pas non plus bon ménage avec les deux autres conditions de l'opérationnalité. Une plus grande cohérence interne dans un modèle linéaire peut conduire à l'élimination de variables. Deux exemples peuvent être avancés. Si deux variables ne satisfont pas suffisamment les critères d'indépendance, il est alors souhaitable d'en supprimer une des deux ou d'en créer une nouvelle les agrégeant. Mais la pertinence risque fort de s'en trouver diminuée si les deux variables influent sur l'objet d'étude. De même, la création d'une macrovariable (variable en agrégeant plusieurs) peut conduire d'une part à une moindre pertinence, mais surtout à des problèmes de mesurabilité. L'autre exemple concerne des variables dont les coefficients ne sont pas significativement différents de zéro. La statistique commande généralement de les enlever de la liste des variables prises en compte par le modèle linéaire. La pertinence peut s'en trouver affectée.

Mais c'est la cohérence d'objectifs qui entre le plus en conflit avec pertinence et mesurabilité. La recherche d'un objectif plus ambitieux nécessitera une approche plus étendue de l'objet d'étude. Prenons un modèle de choix modal opérationnel dans une agglomération pour les modes existants. La volonté d'élargir la cohérence d'objectifs face à l'introduction de modes nouveaux risque fort de se heurter aux autres conditions. Elle nécessite de s'assurer de la pertinence des lois de comportement retenues dans le modèle pour les modes nouveaux. La chose n'est pas évidente par exemple pour l'introduction de la voiture électrique ou d'un métro en l'absence de site propre. Sur le plan de la mesurabilité, le défi est tout aussi grand. Par exemple, l'absence d'expérience face à la voiture électrique risque fort d'imposer la mise en place de protocoles d'enquêtes assez complexes (Faivre d'Arcier et al., 1998).

Face à la problématique du modèle, Bonnafous propose deux solutions. « *La première est tout bonnement la réduction du modèle : réduction de son objectif, de son objet ou des deux* » (Bonnafous, 1989). La seconde consiste à traiter le problème par l'innovation méthodologique. Elle nécessite le plus souvent de déplacer le problème pour le resituer à un autre niveau. C'est évidemment la sortie que l'on préférera, mais elle est nettement plus exigeante.

3.5. Classification des erreurs de modélisation

Si les trois conditions de l'opérationnalité sont antagonistes, il est légitime de se poser la question des limites de la modélisation et de la manière d'appréhender ces limites. Nous en proposons deux lectures complémentaires (pour un approfondissement et une analyse plus systématique de cette question : Leurent, 1997 ; et pour une illustration sur un modèle : Raux et al., 1998). La première s'appuie sur les trois conditions de l'opérationnalité du modèle. Elle consiste à examiner ces trois conditions et à déterminer les infractions commises. Elle reste toutefois insuffisante pour une analyse de l'ensemble des erreurs liées à un exercice de modélisation. La seconde s'appuie sur notre schéma de procédure de prévision (schéma 2.1) en déterminant la nature des erreurs pouvant être commises à chacune des phases de cette procédure :

- erreurs relatives à la construction théorique. Elles découlent notamment des hypothèses faites sur les comportements individuels, que ce soit au niveau du choix des variables ou des relations entre variables. Ces choix nécessitent le plus souvent un certain nombre d'hypothèses qu'il convient d'explicitier. On mentionnera tout particulièrement les hypothèses de la théorie micro-économique du choix du consommateur, qui repose sur la notion de la maximisation de l'utilité individuelle du consommateur. Nous reviendrons sur ce point dans le chapitre 5 ;
- erreurs relatives à la production des données. La mesure réalisée lors d'une enquête est toujours entachée d'une erreur. Cette erreur peut être calculée lorsque les hypothèses nécessaires à son calcul sont respectées (chapitre 4, p. 112). En fait à ces erreurs évaluables, il faut en ajouter d'autres découlant des erreurs d'échantillonnage (base de sondage incomplète, refus de réponse, individus non contactés...), des erreurs de mesure (questions mal comprises, qualité insuffisante du questionnaire, biais liés à l'enquêteur ou à l'enquêté...) (Bonnel et al., 1997) ;
- erreurs relatives à la spécification des relations. Nous avons déjà évoqué ce point. Le choix des fonctions mathématiques ne reflète pas toujours parfaitement la réalité des choses ;
- erreurs de calage. L'estimation des coefficients inconnus s'effectue à l'aide de procédures statistiques nécessitant le respect d'un certain nombre d'hypothèses statistiques qui de fait ne le sont pas toujours. Leurent (1997) parle aussi d'erreurs algorithmiques (non convergence ou précision de la convergence). Elles constituent ce que l'on appelle généralement les biais d'estimation ;
- erreurs de prévision des variables d'entrée du modèle. Nous avons déjà eu l'occasion d'évoquer ce type d'erreur pour mentionner que les rares études de comparaison entre les prévisions des modèles de la demande de déplacements et les réalisations à l'horizon du modèle mettent en avant ce type d'erreur (Mackinder, Evans, 1981 ; CEMT, 1982 ; SEMALY, 1996 ; Gunn, Van der Hoorn, 1998).

L'intérêt d'une telle classification est de permettre des tests sur ces différentes erreurs pour analyser leur incidence sur les résultats de la prévision. Si l'on n'est pas en mesure de résoudre complètement la problématique du modèle au moins peut-on tenter d'évaluer les conséquences des manquements à l'opérationnalité.

L'existence de ces différentes sources d'erreurs potentielles justifie la nécessité des études de sensibilité du modèle à l'incertitude sur certains paramètres. Mais au-delà, il est souhaitable de bâtir des scénarios contrastés, notamment à partir des marges de fluctuations possibles sur les variables d'entrée. L'expérience accumulée en prévision doit nous rendre modeste. On sait bien que l'avenir est en partie incertain, tout au plus peut-on définir l'univers des futurs possibles. La construction de scénarios contrastés peut aider à identifier cet univers des futurs possibles et indiquer dans cet univers ceux qui semblent les plus probables.

4. EXEMPLE : CALAGE DU MODELE PRIX-TEMPS

Ce chapitre met en avant de nombreuses dimensions qui peuvent paraître assez abstraites pour celui qui se lance dans la modélisation. C'est pourquoi nous l'illustrons à l'aide d'un modèle très simple, mais qui pour autant n'est pas sans intérêt. Il s'agit du modèle prix-temps appliqué à l'analyse du partage du marché entre les opérateurs ferroviaires et aériens. Nous présentons tout d'abord rapidement le modèle et le principe du calage (section 4.1) pour ensuite illustrer les différents éléments de ce chapitre (section 4.2).

4.1. Présentation du modèle prix-temps et du calage du modèle

4.1.1. Principe du modèle

Ce modèle repose sur l'hypothèse que le choix d'un voyageur entre le train et l'avion s'effectue en fonction de la valeur que celui-ci attribue à son temps, et des caractéristiques de coût et de temps de transport de chacun des modes (Abraham et al., 1969). Ainsi, l'utilisateur k choisit le mode dont le coût généralisé, compte tenu de sa valeur du temps h_k , est le plus faible.

Si P_F et P_A sont les prix respectifs du fer et de l'avion, et si T_F et T_A sont les durées de trajet (y compris terminaux), les coûts généralisés pour l'utilisateur k sont définis par :

$$Cg_A^k = P_A + h_k T_A \quad (2.3)$$

$$Cg_F^k = P_F + h_k T_F \quad (2.4)$$

Ces coûts généralisés sont illustrés graphiquement (graphique 2.1).

Sur une relation donnée i , il existe une valeur du temps h_i telle que :

$$Cg_A^i = Cg_F^i$$

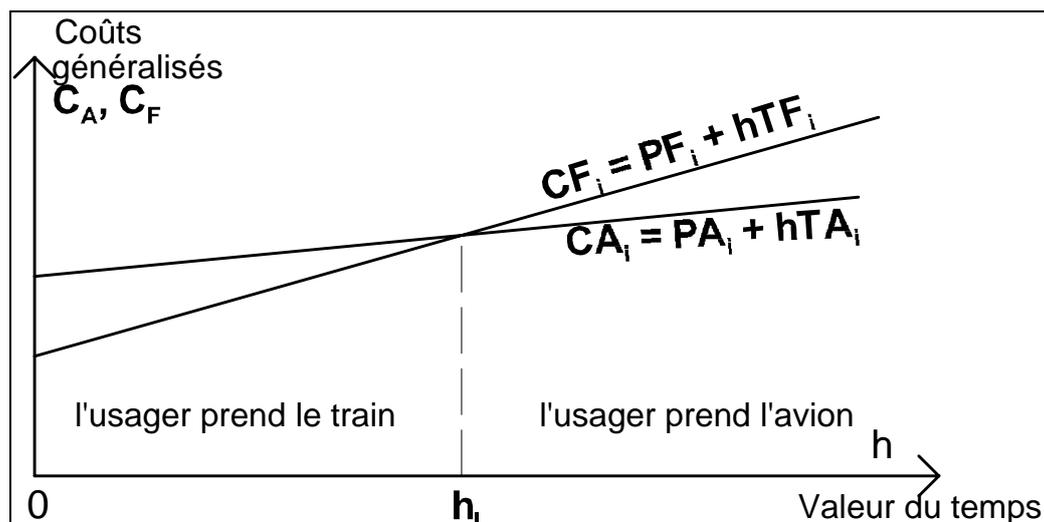
$$\text{soit : } h_i = \frac{P_A - P_F}{T_F - T_A} \quad (2.5)$$

qui est appelée valeur d'indifférence du temps sur la liaison i . Si $h_k < h_i$, le voyageur k choisit le fer, sinon l'avion.

On suppose que la population des voyages est caractérisée par une distribution de la valeur du temps des voyageurs $f(h)$. La fonction de répartition :

$$F(h) = \int_0^h f(x) dx \quad (2.6)$$

est égale à la proportion de voyages dont la valeur du temps est inférieure à h .



Graphique 2.1 : Courbe de coût généralisé fer-avion

Dans ces conditions, si h_i désigne la valeur d'indifférence sur la liaison i , la proportion X_i d'utilisateurs du train dans le trafic total sera donnée par :

$$X_i = F(h_i) = \int_0^{h_i} f(x)dx \quad (2.7)$$

et la proportion Y_i d'utilisateurs de l'avion dans le trafic total sera donnée par :

$$Y_i = 1 - F(h_i) = \int_{h_i}^{\infty} f(x)dx \quad (2.8)$$

Ces équations sont illustrées par les deux graphiques 2.2 et 2.3 (pour une distribution log-normale).

Soit N_A = trafic avion en valeur absolue ;

et N_F = trafic fer 1ère + 2ème classes en valeur absolue,

$$X_i \text{ est donc égal à : } \frac{N_F}{N_A + N_F}, \quad (2.9)$$

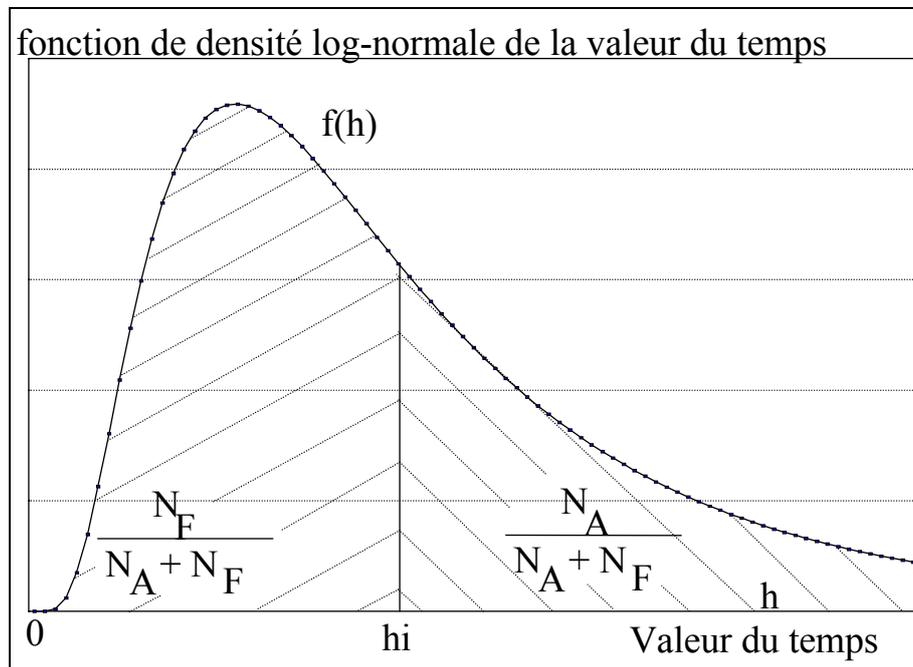
$$\text{et } Y_i \text{ est donc égal à : } \frac{N_A}{N_A + N_F} \quad (2.10)$$

La forme de la fonction de distribution des valeurs du temps n'étant pas connue, on pose par hypothèse une analogie avec la distribution des revenus dans la population. On a donc retenu une fonction de densité $f(h)$ log-normale (Abraham et al., 1969, ces auteurs ont également proposé une distribution de Pareto car « il a notamment été montré que la fonction de répartition des revenus s'ajustait mieux à une loi log-normale pour les bas revenus et à une loi de Pareto en $\frac{1}{x^\alpha}$ pour les

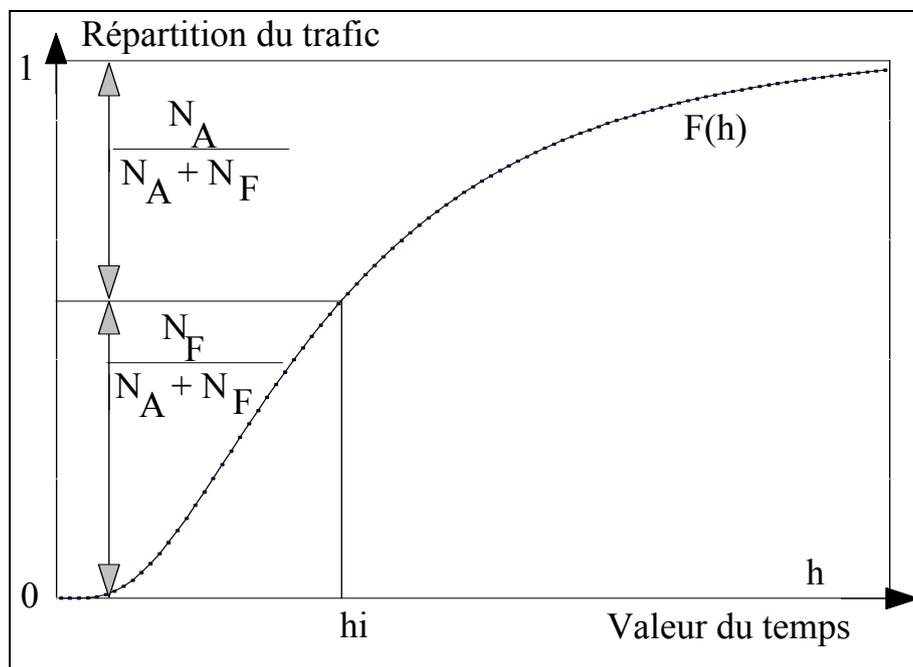
revenus les plus élevés, α étant voisin de 3,2 en 1968 » ; la loi log-normale s'est toutefois imposée, en raison de l'analogie avec la distribution des revenus et pour des considérations empiriques : Arduin et al., 1994) ce qui conduit à l'équation 2.11 (Dagnelie, 1973 ; Calot, 1975) :

$$f(h) = \frac{1}{h\sigma\sqrt{2\pi}} \exp\left(\frac{-(\ln(h)-m)^2}{2\sigma^2}\right) \quad (2.11)$$

où σ est l'écart-type du logarithme népérien des valeurs individuelles du temps, et m la moyenne du logarithme népérien des valeurs individuelles du temps, sont les deux inconnues de la distribution log-normale.



Graphique 2.2 : Forme de la fonction de densité log-normale



Graphique 2.3 : Forme de la fonction de répartition log-normale

La fonction de répartition est de la forme suivante :

$$F(h) = \Phi\left(\frac{\ln(h - h_0) - m}{\sigma}\right), \tag{2.12}$$

où Φ est la distribution cumulative de la loi normale standard ;
et h_0 , la plus petite valeur de h .

On peut donc prendre $h_0 = 0$ F/heure, ce qui simplifie la formulation :

$$F(h) = \Phi\left(\frac{\ln(h)-m}{\sigma}\right) \quad (2.13)$$

4.1.2. Ajustement du modèle

L'ajustement du modèle consiste à caler les paramètres σ (écart-type du logarithme népérien des valeurs individuelles du temps) et m (moyenne du logarithme népérien des valeurs individuelles du temps) de la distribution log-normale.

Si l'on reprend la fonction de répartition (équation 2.13), on peut alors poser :

$$u = \frac{(\ln(h)-m)}{\sigma} \quad (2.14)$$

où u est la transformée de Laplace-Gauss. Elle correspond donc à la fonction inverse de la distribution cumulative de la loi normale standard. La droite ainsi obtenue est appelée la droite de Henry.

La part de marché du fer est fournie par $F(h)$ pour la valeur $h = h_i$. La transformée de la part de marché du fer est donc :

$$u_i = \frac{(\ln(h_i)-m)}{\sigma} \quad (2.15)$$

La loi normale standard étant symétrique, la transformée de la part de marché de l'avion est donnée par :

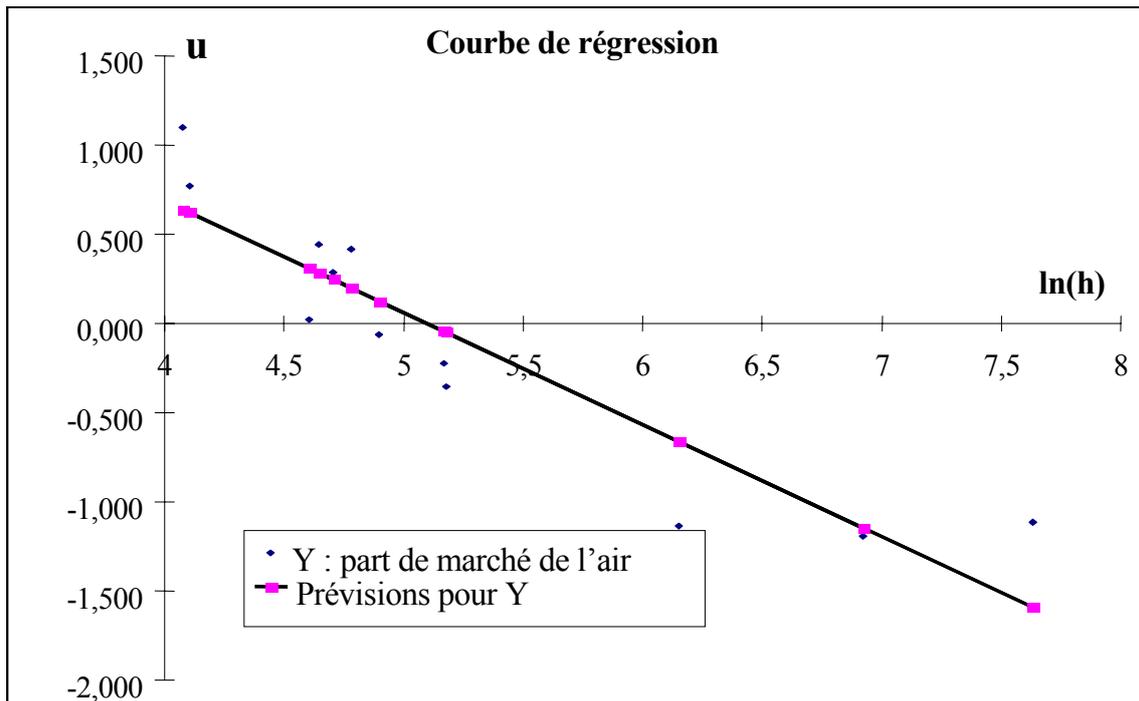
$$-u_i = \frac{-(\ln(h_i)-m)}{\sigma} \quad (2.16)$$

Pour calculer m et σ , il faut déterminer les paramètres de cette droite. Pour cela, nous disposons, sur un certain nombre de liaisons où l'avion et le train sont en concurrence, des données de trafic. Elles nous permettent de calculer les parts de marché respectives du fer et de l'air. Les valeurs de u s'obtiennent en appliquant l'inverse de la transformée de Gauss sur la part de marché du train (ou de l'avion). Deux méthodes sont applicables pour cela :

- utiliser la table de la fonction intégrale de la Loi de Laplace-Gauss donnant u_i à partir de $\Phi(u_i) = F(h_i)$ pour le train ($\Phi(-u_i) = 1 - F(h_i)$ pour l'avion) ;
- utiliser sous tableur la fonction statistique « loi.normale.standard.inverse » qui fournit u_i en fonction de $\Phi(u_i) = F(h_i)$ pour le train ($\Phi(-u_i) = 1 - F(h_i)$ pour l'avion).

Les valeurs de h_i (valeur d'indifférence du temps) sont obtenues à partir des données de temps et de tarif. Il suffit donc de déterminer la droite passant au plus près de ces points par régression linéaire (graphique 2.4).

Le calage du modèle sur les données des années 1967, 1972 et 1980 concernant les neuf relations françaises desservies par voie aérienne depuis le plus longtemps, a permis de fixer la valeur de σ et de vérifier sa stabilité au cours du temps. Il est également apparu qu'au cours des années 70, la médiane des valeurs du temps a augmenté, en valeur constante, au même rythme que la consommation des ménages par habitant en volume.



Graphique 2.4 : Calage de la droite de Henry

Les propriétés de la loi log-normale (Dagnelie, 1973), nous permettent d'écrire :

médiane (h) = $\exp(m)$

où m est la moyenne du logarithme népérien des valeurs individuelles du temps. m évolue donc comme le logarithme népérien de la médiane des valeurs du temps. Nous faisons donc l'hypothèse que m évolue comme le logarithme népérien de la consommation des ménages par habitant en volume.

4.2. Présentation des phases du schéma de procédure de prévision

Nous reprenons les 4 phases du schéma de procédure de prévision (schéma 2.1) afin d'illustrer les différentes notions évoquées.

4.2.1. Phase 1 : objectifs, enjeux, définition du modèle

L'objectif du modèle consiste à prévoir la part du marché aérien et ferroviaire attribuée à chacun des deux modes, en fonction des caractéristiques d'offre des deux modes concurrents. Il s'agit uniquement de répartir le marché et non pas d'évaluer son volume global, ce qui nécessiterait la mise en œuvre d'autres modèles. Nous définissons volontairement l'objectif de manière assez large afin de mettre en évidence les réductions nécessaires du fait de la problématique du modèle appliquée au modèle prix-temps.

Quelle est la définition du modèle théorique ? Pour y répondre, nous analysons tout d'abord le positionnement de l'objectif sur les dimensions temporelles et spatiales. Ensuite nous analysons la nature des interactions mises en œuvre. Enfin, nous décrivons les différentes hypothèses faites sur le comportement des voyageurs et sur le contexte.

Sur le plan temporel, l'ambition du modèle se situe à moyen, voire à long terme. En effet, le modèle doit pouvoir répondre tant à des modifications de service, qu'à des créations de lignes nouvelles et notamment de TGV pour le ferroviaire.

Sur le plan spatial, la portée du modèle se situe sur le trafic entre deux agglomérations. Toutefois, pour simplifier le modèle celles-ci sont considérées dans leur globalité et représentées par une seule zone réduite à son centroïde.

Le modèle mis en œuvre étant très simple, les interactions prises en compte entre les trois sous-systèmes sont fortement simplifiées. On considère le système des localisations comme étant exogène au champ d'analyse de l'objet. En première analyse, cette approximation se justifie dans la mesure où l'objectif n'est pas de prévoir le marché, mais seulement la part du marché attribuée à chaque mode. On peut penser que l'évolution du système des localisations jouera peu sur ce partage. Entre les deux sous-systèmes des transports et des relations sociales, seules les incidences du premier sur le second sont prises en compte. Toutefois, lors de l'analyse (phase 5), si l'offre de transport n'apparaît pas en rapport avec les parts de marché prévues, il est toujours possible de modifier cette offre et de refaire tourner le modèle. Mais la rétro-action reste exogène au modèle.

Deux hypothèses sont donc formulées, dont on pourra discuter de la pertinence en 4.3. Le sous-système des localisations est exogène. Il n'y a pas de rétroaction du système des relations sociales sur le système des transports. On se situe dans le cadre du groupe A d'interactions (schéma 1.1, p. 30). Cette appréhension du système est donc posée comme hypothèse. Elle est bien évidemment discutable, nous y reviendrons plus loin. Elle illustre la nécessité qu'il y a à réduire la complexité des interactions pour n'en retenir que certaines considérées comme les plus importantes (pertinentes) par l'analyste.

Examinons maintenant les autres hypothèses (nous nous limitons aux principales, car elles sont en fait très nombreuses, bien qu'implicites).

Les premières sont relatives au comportement de l'utilisateur du marché de l'aérien et du ferroviaire. L'utilisateur choisit le mode ayant le coût le plus faible. Le coût pris en compte est un coût généralisé. Il comprend le prix du billet et le temps de transport (temps de trajet + temps terminal : temps pour aller de l'origine du déplacement à la gare ou l'aéroport et de la gare ou de l'aéroport à la destination finale). Le temps de transport est monétarisé par le biais de la valeur du temps.

Les secondes sont relatives à la définition de l'utilisateur. On considère qu'il s'agit d'un usager moyen. Cet usager moyen paye un prix moyen : tarif moyen unique pour le ferroviaire à partir des tarifs 1ère et 2ème classe, et unique également pour les différentes classes tarifaires de l'aérien. En cas de concurrence entre opérateurs d'un même mode, on définit également un tarif moyen unique par mode à partir des différents tarifs de chacun des opérateurs. Les temps de transport sont également uniques, tant pour le temps de trajet que pour le temps terminal, quelles que soient les dessertes mises en place et la localisation effective de l'origine et de la destination du déplacement. Seule la valeur du temps est considérée comme un attribut associé au déplacement et non pas évaluée en moyenne.

Les dernières (momentanément) hypothèses sont relatives au contexte. On considère en fait un système des transports qui se limite aux offres ferroviaires et aériennes sur la liaison étudiée. De ce fait, l'offre ferroviaire ou aérienne sur d'autres liaisons n'est pas prise en compte, elle est donc supposée ne pas être en interaction avec le système sur lequel on travaille. De même, l'offre des autres modes de transport n'est pas prise en compte. Elle est supposée n'avoir aucune influence. Si l'on travaillait sur l'évaluation du marché global, il est clair qu'une telle hypothèse ne serait pas tenable, notamment du fait de la concurrence de la voiture. Travaillant sur le partage modal, cela revient à dire que la concurrence des autres modes n'est pas de nature à modifier la concurrence entre le fer et l'air (il est clair que cette hypothèse est discutable. Mais nous verrons aux chapitres 8 et 9 qu'elle est analogue à l'hypothèse nécessaire à l'utilisation des modèles logit (chapitres 8, p. 233 et 9, p. 258)). Le contexte économique est également considéré comme exogène, il agit uniquement au niveau des variables d'entrée que sont les niveaux de tarif et surtout la valeur du temps.

Il est clair que l'on pourrait souligner d'autres hypothèses, en raison de la simplicité du système pris en compte, mais nous préférons nous limiter aux hypothèses principales découlant de la présentation du modèle.

Les variables d'entrée du modèle sont donc les variables entrant dans la détermination du coût généralisé : prix moyen du titre de transport, temps moyen de transport (temps de trajet + temps terminal) et la valeur du temps. Les variables de sortie sont constituées uniquement de la répartition du marché entre les deux modes de transport.

Les relations prises en compte sont également assez limitées. La première lie les variables d'entrée pour la détermination d'un coût généralisé. La seconde concerne la relation entre la valeur du temps des usagers pour un déplacement entre deux agglomérations et la distribution des revenus. La dernière a trait à l'hypothèse de comportement de choix du consommateur qui se traduit par le choix du mode ayant le plus faible coût généralisé. Ces relations sont accompagnées des hypothèses formulées ci-dessus.

4.2.2. Phase 2 : recueil et analyse d'informations

La phase de recueil des données doit permettre de collecter des données pour répondre aux trois besoins identifiés précédemment :

- données permettant le test des hypothèses (nous n'abordons pas cette dimension ici) ;
- données permettant la formalisation mathématique (de nouveau nous n'abordons pas cette question en considérant la formalisation proposée dans la présentation du modèle) ;
- données permettant le calage. Le principe de calage exposé en 4.1.2. nécessite donc la disposition des parts de marché de chaque mode (ou des trafics de chaque mode permettant le calcul des parts de marché), et les données entrant dans le coût généralisé : tarif et temps de transport (temps de trajet et temps terminal) pour quelques liaisons où les modes sont en concurrence. Ces données sont fournies en annexe 1.

4.2.3. Phase 3 : formalisation, calage, validation du modèle

La formalisation consiste à définir la nature des fonctions mathématiques pour les relations prises en compte. Nous avons donc à définir deux fonctions :

- la fonction de coût généralisé : $C_g = P + h \cdot T$ (P : prix ; h : valeur du temps de l'individu pour le déplacement considéré ; T : temps de transport) ;
- la fonction de distribution de la valeur du temps associée à chaque déplacement. L'hypothèse qui est formulée est que la distribution est de même forme que celle des revenus (évidemment les coefficients sont différents dans la mesure où la propension à se déplacer est également corrélée au revenu). Nous retenons donc une fonction de densité log-normale.

La dernière relation entre le choix du mode et les coûts généralisés des deux modes ne nécessite pas de formalisation particulière puisqu'elle est définie dans la nature de la relation : choix du mode ayant le plus faible coût généralisé.

Pour le calage, il ne reste donc plus qu'une seule équation pour laquelle il subsiste des inconnues. Il s'agit de la fonction de densité log-normale. Ces inconnues sont σ l'écart-type du logarithme népérien des valeurs individuelles du temps, et m la moyenne du logarithme népérien des valeurs individuelles du temps. La procédure de calage est décrite en 4.1.2..

	Modèle théorique	Données d'enquête	Ecart
Relation entre Paris et ...	part de marché	part de marché	différence en % absolue relative

	de l'aérien en %	de l'aérien en %		/air	/fer
Bordeaux	45,7	41,1	4,581	11,14	7,78
Brest	53	47,6	5,357	11,25	10,23
Clermont-Ferrand	45,4	36,2	9,242	25,56	14,48
Limoges	21,7	12,8	8,895	69,25	10,21
Lyon	9,6	11,7	2,111	18,08	2,39
Marseille	59,8	67,1	7,34	10,93	22,33
Metz/Nancy	23,1	5,7	17,38	306,1	18,43
Montpellier	58,2	61,2	3,07	5,014	7,921
Nantes	3,7	13,2	9,548	72,16	11,00
Nice	73,7	86,4	12,71	14,7	93,51
Strasbourg	56,2	66,1	9,878	14,94	29,14
Toulon	60,9	50,8	10,04	19,75	20,41
Toulouse	73,1	78,0	4,897	6,281	22,22

Tableau 2.1 : Comparaison modèle théorique, données d'enquête

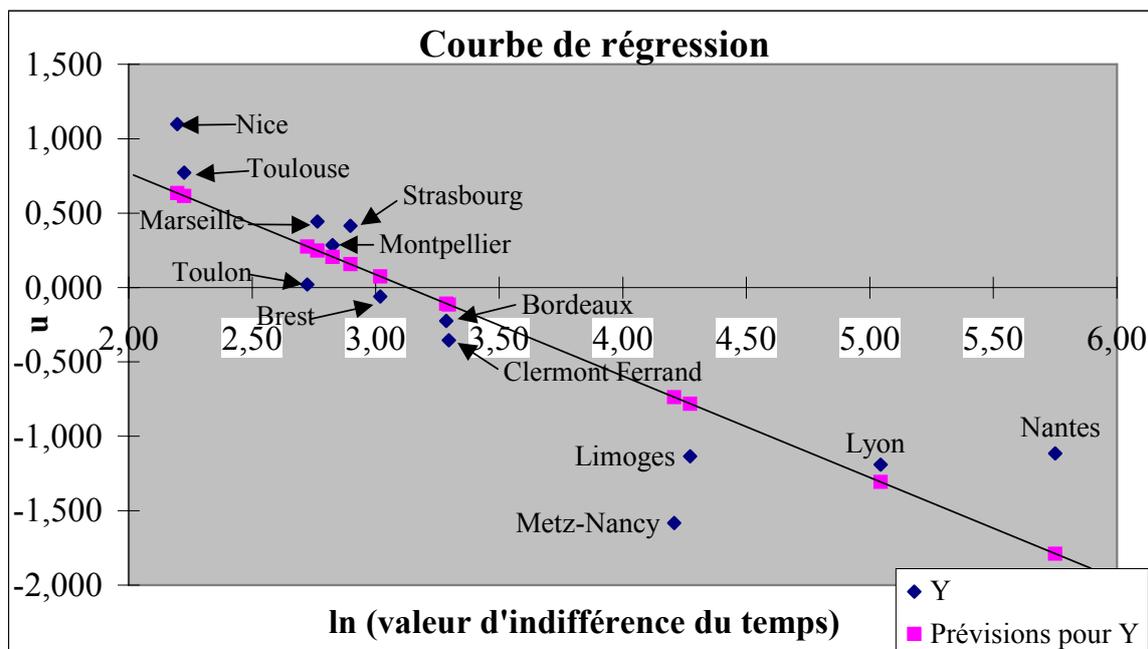
La phase de calage doit toujours s'accompagner d'une analyse des écarts entre la prévision et les données relatives à la situation observée. Les données et les résultats de la régression (sous Excel) sont fournis en annexe 1. Nous reprenons le tableau (tableau 2.1) de comparaison entre les résultats du modèle théorique et les données d'enquête.

Avant d'analyser le tableau, il convient d'étudier les statistiques de la régression. Nous ne développons pas ici ce point et renvoyons le lecteur vers le chapitre 6 (p. 174) ou vers les manuels correspondants de statistiques ou sur les modèles linéaires (notamment Cohen, Pradel, 1993 ; Wonnacott, Wonnacott, 1991). Il faut au minimum analyser le R^2 , les statistiques de Fisher et de Student (il existe également de nombreux autres tests (Cohen, Pradel, 1993 ; Wonnacott, Wonnacott, 1991). Nous nous limitons à ceux disponibles dans le module de régression d'Excel 5, que nous commentons succinctement en annexe 1). Il convient également d'examiner les points pouvant constituer des effets de levier (Ritschard, 1990) susceptibles de modifier la pente de la droite (c'est le cas de Nantes, graphique 2.5).

Nous nous limitons à l'analyse des points aberrants (Ritschard, 1990). C'est-à-dire des points s'écartant fortement de la droite de régression. Deux raisons justifient ce choix. Tout d'abord ces points sont susceptibles de modifier sensiblement l'équation de la droite. Ensuite, ils peuvent désigner des problèmes soit au niveau de la qualité de la donnée correspondante, soit au niveau de la pertinence ou de la cohérence d'objectif du modèle. Nous illustrons ce second point à travers les trois villes présentant les plus grands écarts relatifs par rapport à la part de marché de l'air ou du fer.

Prenons tout d'abord le cas de Metz-Nancy. La part de marché de l'aérien issue de l'enquête est de 5,7 % alors que le modèle conduit à une valeur de 23 %. Le modèle ne reconstitue donc pas du tout la situation de Metz-Nancy. Il convient alors de passer en revue l'ensemble des erreurs possibles pour tenter de comprendre pourquoi le modèle est défaillant. On peut en effet craindre que le modèle ne soit pas apte à faire de la prévision, s'il n'est déjà pas capable de reproduire correctement la donnée de départ. Nous nous limitons ici au premier type d'erreur à savoir les erreurs relatives à la construction théorique. Nous avons postulé que l'utilisateur est toujours en situation de choix. Comme dans bon nombre de modèles, il s'agit en fait d'une hypothèse implicite. Mais qu'en est-il réellement ? La demande aérienne annuelle n'est que de 75.000 passagers. Un calcul sommaire (en considérant que le trafic annuel équivaut à 300 jours moyens de semaine) nous fournit 125 passagers par sens et par jour. On conçoit aisément que l'offre aérienne ne peut dans ce cas être abondante. Tout au plus aura-t-on deux à trois vols par jour, les jours ouvrables. La fréquence ferroviaire est par contre beaucoup plus

importante. Dès lors, on ne peut réellement considérer que l'on est en situation de choix. Si l'on conserve le modèle en l'état, il faut alors expliciter cette hypothèse et restreindre l'opérationnalité du modèle aux relations pour lesquelles on considère que les usagers sont réellement en situation de choix au regard de la fréquence des avions et des trains (nous verrons plus loin que l'on peut apporter d'autres solutions à ce problème de pertinence). Le modèle ne peut donc pas être utilisé à Metz-Nancy, Limoges et dans une moindre mesure à Clermont-Ferrand dans la configuration actuelle de l'offre aérienne.



Graphique 2.5 : Application du modèle prix-temps sur le partage air-fer - courbe de régression

La seconde ville présentant des problèmes importants est Nantes. Cette fois, la demande est nettement supérieure à ce que prévoit le modèle, malgré une offre aérienne assez faible en raison de la forte concurrence du TGV. C'est une autre hypothèse qui se trouve ici prise en défaut. Nous considérons dans notre modèle un usager moyen en ce qui concerne les tarifs et les temps d'accès à la gare ou à l'aéroport. Donc, par hypothèse implicite, nous considérons que tous les usagers sont identiques à cet usager moyen pour ces deux dimensions (c'est une hypothèse commune à tous les modèles agrégés. Les modèles désagrégés tentent de s'en affranchir le plus possible. Nous explicitons ces distinctions au chapitre 5). Bien que cette hypothèse soit évidemment abusive, on espère toutefois que les erreurs commises se compensent du fait de la définition moyenne de cet usager. C'est bien souvent le cas, mais Nantes offre un contre-exemple. La part de marché de l'aérien étant très faible, les clients qui restent sur ce mode sont en fait pour une part importante des clients qui se différencient fortement de l'usager moyen. Ce n'est pas que leur comportement ne soit pas rationnel, mais plutôt que les hypothèses relatives à leur situation sont erronées. Parmi la clientèle aérienne, il y a une part d'usagers faisant un vol international transitant par Paris, dans ce cas, le tarif aérien peut fortement diverger par rapport à celui que nous avons retenu. De même, nous trouverons des usagers pour lesquels les trajets terminaux sont en fait beaucoup plus courts pour l'avion et à l'inverse plus long pour le train de sorte que l'application du modèle avec cette donnée leur ferait effectivement préférer l'avion au lieu du train (c'est le cas des usagers qui résident à proximité des aéroports et se destinent dans la zone proche de l'aéroport de destination). On peut également penser à l'ensemble des personnes acquittant un tarif très faible voire nul (personnels de compagnie aérienne, d'agences de voyage...). L'analyse peut évidemment être reprise pour le train. Ces usagers fortement différents de l'usager moyen sont donc présents sur toutes les liaisons à la fois sur le train et sur l'avion. Toutefois,

lorsque la part de marché de chacun des modes est proche de 50 %, les effets de cette clientèle sur chacun des deux modes s'annulent et sont implicitement pris en compte dans le calage de la fonction de distribution des valeurs du temps. En revanche, lorsque la part d'un des deux modes s'approche de 0 %, ces clients « atypiques » représentent une part importante de la clientèle de ce mode. Le modèle doit donc être utilisé uniquement sur des villes pour lesquelles la part de marché théorique ne s'approche pas trop de 0 %. Pour l'aérien, l'examen des résultats de Nantes et de Lyon pour laquelle la concurrence du TGV est également très forte, semble indiquer une limite de l'ordre de 10 % (une limite de l'ordre de 90 % doit probablement également s'appliquer par symétrie pour le train). Cette remarque est bien évidemment également applicable en prévision. Toute prévision conduisant à une part de marché inférieure à 10 % doit être considérée avec beaucoup de précaution (la remarque n'est pas inutile, si l'on veut utiliser le modèle sur des liaisons où existent des projets de TGV offrant des conditions de concurrence analogues à Nantes ou Lyon, comme Strasbourg par exemple). De nouveau d'autres solutions que la réduction de la portée du modèle peuvent être apportées, comme nous le verrons plus loin.

Nice est notre troisième cas problématique. La part de marché observée de l'avion est plus importante que la part de marché théorique. Le problème vient cette fois en grande partie des données utilisées. L'aire d'attraction de l'aéroport de Nice ne se réduit pas à la ville de Nice, mais englobe une partie de la Côte d'Azur. En revanche, compte tenu du maillage ferroviaire, l'aire d'attraction de la gare de Nice ne déborde probablement pas beaucoup l'agglomération niçoise. Pour utiliser correctement le modèle, il faudrait probablement additionner les trafics des gares de Nice, Menton, Cannes... afin de retenir des aires de chalandise comparables.

Nous pourrions prolonger cette analyse pour identifier d'autres limites à notre modèle, par exemple le cas de Marseille qui capte une partie du marché aérien de Toulon en raison de meilleures fréquences..., tant cette analyse peut être riche d'enseignement. Mais notre objectif n'est pas ici d'analyser le modèle prix-temps. Il est plutôt de l'utiliser comme explicitation des notions développées dans ce chapitre. Ainsi, ce cas d'école illustre mieux qu'un long discours l'importance de la phase de calage. Outre l'amélioration de la reproduction du jeu de données par le modèle qui est un des objectifs de la phase de calage, cette analyse des écarts prévisions-données est nécessaire au moins à trois niveaux :

- pour contribuer à expliciter certaines hypothèses restées implicites ;
- pour délimiter le domaine de pertinence du modèle ;
- pour détecter d'éventuels problèmes de production, voire de définition des données.

Enfin, vient l'étape de validation du modèle. Il s'agit de faire varier les variables d'entrée tour à tour et d'analyser les résultats en sortie. Il est alors possible de calculer des coefficients d'élasticité de la part de marché d'un mode à la variation des variables d'entrée. On parlera d'élasticité directe par exemple pour l'élasticité de la part de marché du train à la variation du tarif ferroviaire et d'élasticité croisée si l'on fait varier le tarif de l'aérien. On doit raisonnablement attendre de la première qu'elle soit de signe négatif, une croissance des tarifs ferroviaires, toutes choses égales par ailleurs, doit contribuer à la réduction de la demande pour ce mode, et que la seconde soit de signe positif.

4.2.4. Phase 4 : prévision

Pour utiliser le modèle en prévision, l'exercice est très simple, il suffit de modifier les variables d'entrée et d'obtenir en sortie la part de marché de chacun des modes. La moyenne du logarithme népérien des valeurs du temps doit également être modifiée pour tenir compte de sa croissance en euros constants. On retient en général une croissance égale à celle du PIB pour la médiane des valeurs du temps.

Concrètement, on peut tester la mise en service d'une ligne TGV sur une desserte. Il suffit pour cela de modifier le prix et le temps de transport ferroviaire sur la liaison, de déterminer la médiane des

valeurs du temps à l'année horizon d'étude pour obtenir la part de marché du TGV et de l'avion sur la liaison concernée.

4.3. Le modèle : résultat d'un double processus de simplification

Nous reprenons ici rapidement les deux niveaux de simplification.

4.3.1. Première simplification : la construction théorique

Nous avons déjà évoqué dans les sections précédentes certaines des simplifications opérées lors de la construction du modèle théorique. Nous avons tout d'abord celles relatives à l'appréhension des interactions entre les trois sous-systèmes :

- le sous-système des localisations est considéré comme exogène. C'est-à-dire que les modifications dans l'implantation des activités économiques ou de l'habitat sont sans effet sur la part de marché de chacun des modes. L'hypothèse est évidemment assez forte sur le long terme. Par exemple, on peut penser que le développement dans une ville de province d'activité à forte valeur ajoutée nécessitant des relations fréquentes avec Paris est de nature à accroître la part de marché de l'avion toutes choses égales par ailleurs ;
- l'interaction entre le sous-système des transports et le sous-système des relations sociales est univoque ;
- le sous-système des transports est réduit aux deux modes étudiés. En particulier, l'offre routière n'est pas prise en compte. On considère donc qu'elle est sans effet sur la part de marché entre l'avion et le train. De plus, seule l'offre sur la relation étudiée est considérée (les exemples de Nice d'une part, et de Toulon et Marseille d'autre part montrent la limite de cette hypothèse).

Nous avons également un jeu d'hypothèses déjà mentionnées :

- l'utilisateur est toujours en situation de choix entre les deux modes. Cela suppose donc une fréquence suffisante pour les deux modes, mais aussi l'existence de places disponibles pour tous les trains et avions ;
- l'utilisateur choisit le mode ayant le coût le plus faible ;
- le coût pris en compte est un coût généralisé faisant intervenir le temps de trajet et le temps terminal, le prix du billet et la valeur du temps. Les autres éléments sont donc sans influence sur le choix de l'individu ;
- on considère un usager moyen censé représenter la diversité des usagers ;
- ...

L'analyse du calage illustre l'importance de bien expliciter l'ensemble des hypothèses (simplifications) formulées. Elles facilitent la détermination du domaine de validité du modèle.

4.3.2. Deuxième simplification : la formalisation mathématique

Les relations prises en compte étant peu nombreuses, la formalisation est assez simple.

La première formalisation concerne la formule du coût généralisé : $C_g^k = P + h_k T$. Dans ce cas, la formalisation en elle-même n'apporte pas de simplification, si ce n'est l'hypothèse d'additivité du prix et du temps par l'intermédiaire de la valeur du temps. Les simplifications ont en fait été définies précédemment (prix et temps unique commun à tous les usagers, non prise en compte d'autres facteurs dans la formulation du coût généralisé).

La seconde formalisation concerne la distribution de la valeur du temps. Cette distribution étant inconnue, on choisit la distribution habituellement utilisée pour les revenus. Il est clair qu'il s'agit

d'une approximation de ce que l'on soupçonne de la réalité dans la mesure où l'on ne connaît pas la distribution effective de la valeur du temps. Toutefois, on est incapable d'évaluer l'importance de la simplification opérée, sauf à réaliser une enquête auprès des clientèles des deux modes pour quantifier la valeur du temps associée aux déplacements réalisés (l'incertitude ne serait en fait pas complètement levée, la mesure pouvant varier selon la méthodologie d'enquête retenue. De plus, comme nous l'avons déjà mentionné, les données sont elles-mêmes une approximation de la réalité (Bonnafous, 1989)).

4.4. Opérationnalité et problématique du modèle

Nous avons déjà eu l'occasion d'aborder l'opérationnalité du modèle prix-temps. Nous le traitons donc rapidement, pour insister surtout sur la problématique du modèle et les pistes permettant de la résoudre.

4.4.1. Pertinence

Les problèmes de pertinence ont déjà été évoqués. Ils résultent pour une large part des hypothèses formulées. Il ne nous semble donc pas utile de les mentionner de nouveau.

4.4.2. Cohérence

Du fait de la simplicité du modèle mathématique, la cohérence interne du modèle ne pose pas de problème particulier.

Il en est apparemment de même de la cohérence d'objectifs dans la mesure où l'on utilise le modèle de manière cohérente avec ce pour quoi il a été conçu. Par exemple, il serait illusoire de tester la sensibilité du partage modal entre l'avion et le train en fonction du niveau d'offre routière, ou en fonction de la fréquence des avions ou des trains.... Cette remarque semble être une lapalissade. Elle n'est toutefois pas aussi anodine qu'il paraît. Il est fréquent de voir des modèles utilisés en dehors de l'objectif pour lequel ils ont été conçus. Ceci est notamment dû à l'insuffisance (voire l'absence) fréquente de la première phase dans les exercices de modélisation, phase au cours de laquelle les objectifs et le domaine de validité du modèle doivent être précisés.

En fait, la phase de calage a mis en évidence une cohérence d'objectifs plus limitée que celle que nous avons implicitement définie. Le modèle ne peut pas être utilisé pour modéliser le partage air-fer dans tous les contextes. En particulier, lorsque la fréquence de l'un des deux modes est trop faible, la structure du modèle n'est plus cohérente avec l'objectif visé (cas de Metz-Nancy). De même, lorsque la part de marché de l'un des deux modes s'écarte de la plage 10-90 %, la cohérence n'est plus garantie (cas de Nantes). Ces deux exemples posent évidemment problème au regard de l'une des applications possibles du modèle prix-temps : la simulation de la perte de marché suite à la mise en service du TGV. Dans ce cas, la part de l'avion risque fort de chuter fortement avec, de ce fait, des fréquences limitées.

4.4.3. Mesurabilité

Malgré la simplicité du modèle, nous sommes confrontés à de multiples problèmes de mesurabilité. Nous abordons tout d'abord la mesure du coût généralisé. Nous évoquons ensuite les données nécessaires au calage.

Le calcul du coût généralisé nécessite la détermination d'un prix et d'un temps. Si l'opération peut sembler simple, elle est en fait très complexe. Comment calculer un prix moyen pour le ferroviaire ou l'avion, dans la mesure où il existe une multitude de tarifs en fonction : de la classe, des suppléments éventuels pour le train, des multiples réductions commerciales ou sociales, des gratuités

totales ou partielles, de la nature du voyage complet (domestique ou international)... On pourrait chercher à déterminer un tarif moyen en recherchant chacun des tarifs et leur poids. Toutefois, l'information n'est pas disponible. La réalisation d'une enquête permettrait de résoudre le problème. Mais peut-on transposer les résultats d'une relation sur les autres ? Et surtout, la répartition actuelle peut-elle être maintenue en prévision ? On peut en douter compte tenu de l'évolution assez rapide des tarifs dans un monde de plus en plus concurrentiel.

Le problème est de même nature pour les temps. Pour les temps de trajet, la diversité est moins grande, mais le problème reste de même nature tant pour déterminer une valeur à la date de référence qu'en prévision, car les programmes de desserte ne peuvent être connus avec prévision dans le futur. Les temps terminaux posent des problèmes analogues. Leur détermination suppose de connaître les temps d'accès de chacun des voyageurs. Cette donnée n'est pas disponible, nécessitant le recours à une enquête. Mais on retrouve toujours le même problème pour la prévision : comment faire évoluer ces temps terminaux ? L'étalement urbain, la congestion, les systèmes de transports urbains... sont autant de paramètres pouvant influencer sur ces temps.

Quelle que soit la solution adoptée pour déterminer tarif et temps, il importe que la mesure soit identique d'une part pour les deux modes et d'autre part lors du calage et en prévision. Dans l'exercice, un calcul très simple a été effectué. Il consiste à retenir les pleins tarifs de chacun des niveaux de suppléments et de chacune des classes en les pondérant par leur poids sur la relation considérée et de retenir le plein tarif de la compagnie dominante pour l'avion.

Pour le calage, il faut également disposer des trafics. L'accès à cette information n'est pas toujours très aisé compte tenu du développement de la concurrence entre opérateurs. Les problèmes de transit ne facilitent pas non plus la détermination de ces trafics.

Ces difficultés ne sont pas spécifiques au modèle utilisé. La production des données est ainsi bien souvent une opération délicate et coûteuse.

4.4.4. Problématique du modèle

La problématique du modèle résulte de la résolution contradictoire des trois conditions de l'opérationnalité des modèles. Pour l'illustrer nous reprenons les trois limites du modèle prix-temps évoquées lors du calage pour les villes de Metz-Nancy, Nantes et Nice, pour examiner les pistes de sortie.

Pour le cas de Metz-Nancy, nous nous trouvons face à un problème de pertinence, l'hypothèse selon laquelle tous les individus sont en situation de choix se trouvant invalidée par la faiblesse de la fréquence de la desserte aérienne. La première solution consiste à réduire la portée du modèle en ne l'utilisant que sur des liaisons où l'offre des deux modes est suffisante, comme nous l'avons mentionnée en 4.4.2. Cette solution n'est toutefois pas très satisfaisante. Une autre solution consiste à faire intervenir la fréquence dans la formulation du coût généralisé. Si l'on peut ainsi résoudre le problème de pertinence, on se trouve, en revanche, face à un problème de mesurabilité. Comment mesurer la fréquence ? faut-il faire intervenir l'amplitude horaire, distinguer heures creuses et heures de pointe, additionner l'offre d'opérateurs aériens concurrents ? De plus, quelle formulation mathématique retenir ?

Dans le cas de Nantes, nous nous trouvons de nouveau face à un problème de pertinence. Une partie de la clientèle peut en fait être considérée comme partiellement ou totalement captive de chacun des modes et s'éloigne donc trop fortement de la définition de notre individu moyen. La première solution consiste de nouveau à réduire la portée du modèle, puisque ce problème de pertinence ne se pose de manière importante que lorsque la part de marché de l'un des deux modes devient très faible. Une autre solution plus satisfaisante consisterait à définir une population captive de chacun des modes

et faire porter le modèle uniquement sur les individus en situation de choix. On résout ainsi le problème de pertinence, mais on se trouve face à des problèmes de mesurabilité et de cohérence. Comment mesurer le marché captif ? De plus, quelles hypothèses d'évolution de ce marché faut-il retenir dans le temps ? Le problème de cohérence interne tient à la détermination de la distribution des valeurs du temps. L'hypothèse a été formulée sur l'ensemble de la population des voyages. Est-il légitime de la conserver sur une population plus restreinte dans laquelle les captifs ont été ôtés ?

Enfin, le cas de Nice pose un problème à la fois de pertinence et de cohérence. Le problème de pertinence renvoie à la définition d'une aire d'attraction. Doit-on retenir une aire commune pour les deux modes ? Pour le mode aérien, peu d'aéroports sont en situation de concurrence pour le trafic domestique (à l'exception marginalement de Toulon et Marseille). En revanche, si l'on retient l'aire de chalandise d'un aéroport, les situations de concurrence entre gares sont plus fréquentes. Selon la définition choisie, on pourra se trouver ensuite face à des problèmes de mesurabilité (par exemple, si on décide de ne pas retenir la totalité du trafic d'une gare, la mesure du pourcentage à retenir pourra poser problème). Le problème de cohérence renvoie à la cohérence d'objectifs. Si l'objectif est de prévoir la part de marché de la gare de Nice, additionner le trafic de la gare de Nice avec celui des gares incluses dans la zone de chalandise de l'aéroport de Nice, ne permettra pas de répondre à la question posée (sauf à définir ensuite un modèle de répartition entre les gares concernées).

5. CONCLUSION

Ce cadrage méthodologique nous a permis d'introduire certaines notions fondamentales en matière de modélisation. L'utilisation d'un modèle ne doit pas se résumer à l'utilisation d'un outil. C'est un exercice dans lequel le modèle est certes un outil, mais qui doit s'intégrer dans une démarche reprenant les 5 phases exposées dans le schéma de procédure de prévision (schéma 2.1). Chacune de ces phases conditionne la qualité de la prévision ainsi que nous avons pu le montrer. La construction d'un modèle nécessite ainsi trois étapes qui doivent rester distinctes, même si elles sont fortement imbriquées : la construction du modèle au sens de la représentation théorique d'un objet, sa formalisation mathématique et enfin son calage. Tout cet ensemble nécessite une production des données qui peut venir alimenter ces trois étapes.

En tant que représentation simplifiée de la réalité, le modèle est le résultat d'un processus de simplification. Nous avons identifié principalement deux étapes dans ce processus : la construction théorique de l'objet d'étude et sa formalisation mathématique. Le modèle étant une représentation simplifiée, il convient de s'assurer de son opérationnalité, c'est-à-dire de s'assurer de sa capacité à reproduire la réalité, ou plutôt une certaine mesure de la réalité correspondant aux données recueillies. Nous avons établi les trois conditions de son opérationnalité : pertinence, cohérence et mesurabilité. Ces conditions étant contradictoires, elles forment la problématique du modèle.

La recherche de solutions à cette problématique nous conduit vers les chapitres suivants de ce manuel. La recherche d'une plus grande pertinence conduit vers l'analyse des comportements de mobilité (chapitre 3), tandis que la recherche d'une plus grande mesurabilité nous conduit vers la production des données (chapitre 4). Nous pourrions alors entrer dans le cœur de notre manuel, avec la présentation des modèles à quatre étapes (chapitres 5 à 10).

Chapitre 3 - Analyse de la mobilité

Notre propos n'est pas de fournir un panorama des analyses de la mobilité urbaine et de ses multiples déterminants (le lecteur intéressé par ces questions peut se reporter à de multiples publications, et notamment à Andan et al., 1988 ; INRETS, 1989 ; Salomon et al., 1993 ; Orfeuil, 1999 ; les actes des conférences de l'« *International Association of Travel Behaviour* » (International Association for Travel Behaviour, 1989 ; Stopher, Lee-Gosselin, 1997 ; Ortuzar et al., 1998 ; Hensher, 2001)...). Nous cherchons plutôt à situer en évolution les principaux indicateurs de mobilité urbaine (section 2) et à présenter quelques-unes des tendances lourdes qui influent sur la mobilité (section 3). Cette dernière section nous permet d'illustrer la complexité de l'analyse de la mobilité en raison de la conjonction de facteurs influant pour certains dans le même sens, pour d'autres dans des sens opposés. De plus, ces facteurs sont le plus souvent corrélés entre eux, rendant délicate leur séparation. Enfin, nous montrons que des angles d'analyses différents peuvent conduire à des conclusions radicalement opposées. Trois tendances structurelles sont retenues : l'étalement urbain, la motorisation et le vieillissement de la population. Les exemples d'analyse ont été choisis tout autant pour leur dimension fortement pédagogique, que pour les résultats produits. Mais auparavant, nous proposons une définition de la mobilité (section 1).

1. MOBILITE : UNE DEFINITION EN EVOLUTION

Qu'est-ce-que la mobilité ? En première analyse, c'est une quantité facile à cerner qui correspond au nombre de déplacements réalisés. En fait, la réalité n'est pas aussi simple. Qu'y a-t-il de commun entre un déplacement à pied de courte distance et un déplacement effectué en voiture de la banlieue vers le centre ? Pour le planificateur du système urbain, il ne s'agira pas du même objet, même si dans les deux cas il y a eu déplacement. Le premier n'utilise pas (ou presque pas) l'infrastructure de transport. Il ne donne pas lieu à modélisation et le plus souvent n'est pas pris en compte dans la planification. A l'inverse, le second sera l'objet de toutes les attentions, car il mobilise ces infrastructures que ce soit en roulant ou pour stationner. Ces objets différents doivent-ils renvoyer à des définitions différentes de la mobilité ?

Historiquement, la définition de la mobilité était réduite au nombre de véhicules circulant sur un axe. Cette première définition correspondait au besoin de dimensionnement des infrastructures. D'un axe isolé, la problématique a évolué vers le réseau de transport. La définition correspond toujours à un nombre de véhicules, mais cette fois l'origine-destination du véhicule devient nécessaire pour connaître son utilisation du réseau concerné. Cette définition correspondait aux données nécessaires pour faire tourner les premiers modèles de trafic. Progressivement, cette définition a été enrichie au fur et à mesure de l'évolution des besoins en matière de planification, mais aussi sous l'effet des recherches menées sur la connaissance de la mobilité urbaine (chapitre 1).

Aujourd'hui encore, on peut rencontrer plusieurs définitions selon les besoins d'études ou de recherches. La définition la plus couramment utilisée dans le milieu des études de prévision, correspond au nombre de déplacements réalisés par une personne au cours d'une journée. L'unité d'observation n'est donc plus le véhicule, mais la personne. De plus, on ne se limite plus aux déplacements motorisés, mais on prend en considération l'ensemble des modes de transport, y compris la marche à pied ou les deux roues. Enfin, les déplacements sont considérés avec certains de leurs attributs, comme l'origine, la destination, le motif, l'heure de réalisation ou le moyen de transport employé....

Les recherches menées sur l'analyse de la mobilité ont conduit à affiner cette définition. Le déplacement n'est plus seulement considéré comme moyen de se déplacer, mais plutôt comme moyen de réaliser certaines activités en différents lieux de l'espace. En effet, à l'exception de quelques cas particuliers (promenade...) le déplacement n'est pas un but en soi. Il sert à la réalisation d'activités spatialement définies. De plus, le déplacement et l'activité réalisée à destination ne sont plus considérés isolément, mais insérés dans l'ensemble du schéma d'activités réalisé au cours de la journée. De ce fait, les déplacements sont appréhendés comme un cheminement dans le temps et dans l'espace permettant de réaliser des activités inscrites dans ce temps et dans cet espace. La mobilité n'est plus alors seulement repérée par des déplacements. On est également passé d'un individu considéré isolément à un individu situé dans son entourage (familial, professionnel...). Cette dernière évolution vise à identifier les contraintes interpersonnelles des individus qui peuvent « expliquer » une part de la mobilité quotidienne (Bonnel, 1992). Cette définition permet de mieux analyser les déplacements et les contraintes qui pèsent sur leur déroulement dans la journée et sur leur localisation. Toutefois, si l'on a pu mettre au point des outils opérationnels de recueil de données (les schémas d'activités, schéma 3.1 (Bonnel, 1985 ; Axhausen, 1995)) correspondant à cette définition, ce n'est pas encore vraiment le cas des méthodes de prévision de la demande (même si les recherches sont nombreuses sur le sujet (notamment Ettema, Timmermans, 1997 ; Timmermans et al., 1997)).

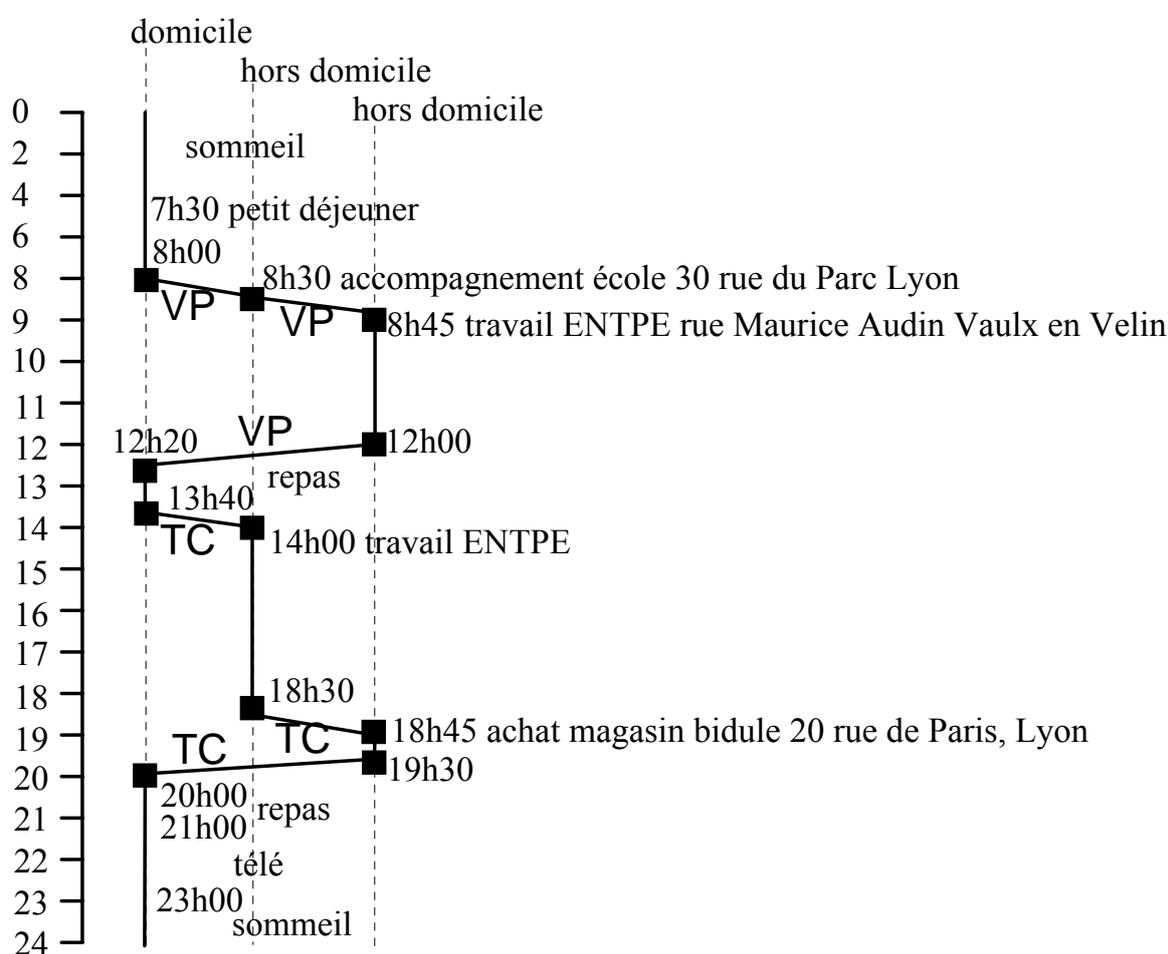


Schéma 3.1 : Diagramme espace-temps de recueil de la mobilité quotidienne

Ces multiples définitions illustrent encore une fois, les liens entre la définition de l'objectif de l'analyse, le modèle utilisé et la production des données. La première définition faisait appel à une production de données très simple, le comptage, et à un modèle lui-même très simple permettant de dimensionner les infrastructures. Les définitions suivantes ont fait évoluer progressivement la

production des données : des premières enquêtes origines-destinations auprès des ménages vers les enquêtes ménages déplacements et plus récemment dans certains pays vers les enquêtes basées sur les activités réalisées (Stopher, 1998). Les modèles ont eux-aussi connu des évolutions que nous avons décrites dans notre premier chapitre. Nous illustrons cette évolution à travers la présentation de l'enrichissement progressif du cadre conceptuel d'analyse de la mobilité (chapitre 4, p. 139).

2. DESCRIPTION DE LA MOBILITE ET DE SON EVOLUTION

Nous présentons quelques résultats issus de plusieurs enquêtes ménages déplacements (chapitre 4, p. 117) réalisées dans quelques villes françaises au cours des 20 dernières années. Nous pouvons ainsi illustrer l'évolution du nombre moyen de déplacements quotidiens par personne. Nous donnons tout d'abord une vision de l'ensemble des modes de transport pour avoir ensuite une approche modale. Dans cette section, nous utilisons le terme mobilité au sens du nombre de déplacements quotidiens par personne, pour simplifier le propos d'une part et par adéquation avec la définition retenue par les enquêtes ménages déplacements d'autre part.

L'analyse des évolutions de la mobilité doit être faite avec une certaine prudence dans la mesure où les effectifs ne sont pas toujours suffisants pour que les écarts entre deux enquêtes soient statistiquement significatifs (chapitre suivant). Nous présentons des données pour des périmètres d'enquête constants. Les données les plus récentes peuvent donc différer des statistiques présentées sur des périmètres plus étendus. C'est le cas à Lyon où les trois enquêtes ménages déplacements ont été menées sur des périmètres différents (nous utilisons, dans cette section 2, le périmètre de l'enquête de 1976-77).

Notre analyse est volontairement assez sommaire. Nous cherchons seulement à mettre en évidence les tendances lourdes d'évolution communes aux différentes agglomérations et à donner les principaux paramètres permettant d'interpréter ces évolutions. Une analyse plus fine conduirait probablement à compléter la liste des facteurs pouvant avoir une influence sur les indicateurs étudiés. Le lecteur souhaitant approfondir l'analyse sur les déterminants de la demande de mobilité ou sur les données de mobilité pourra consulter les documents suivants (Andan et al., 1988 ; INRETS, 1989 ; SYTRAL et al., 1990 ; DREIF, 1995 ; Raux et al., 1996 ; INSEE, INRETS, 1997 ; Orfeuil, 1999 ; DREIF, 2001...).

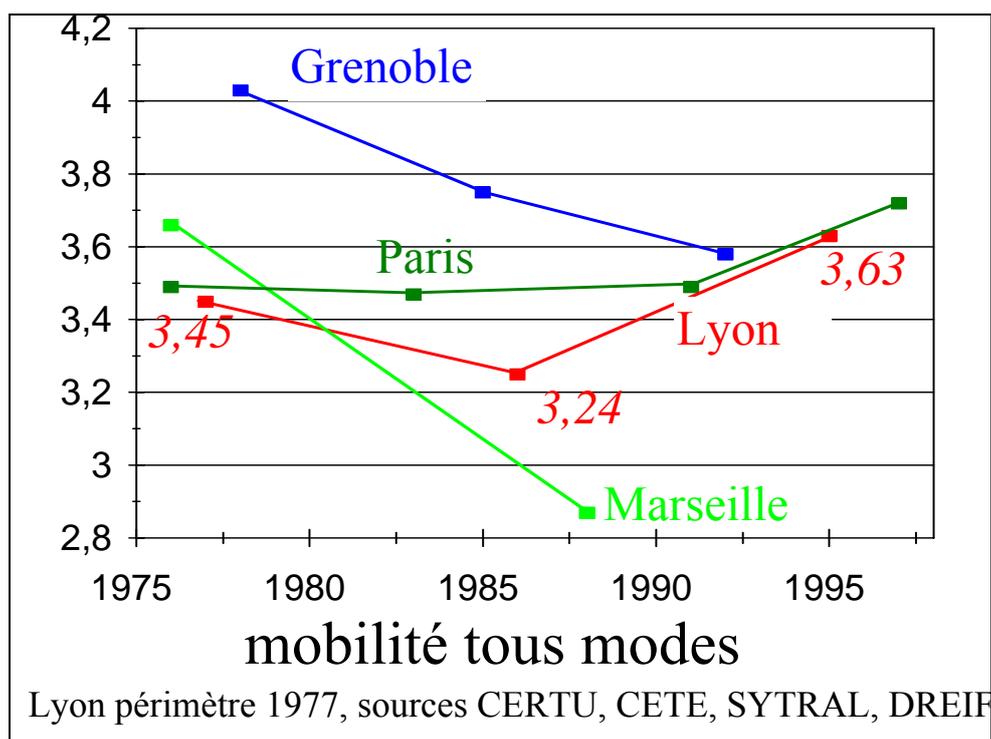
2.1. Une absence de tendance claire pour le nombre de déplacements quotidiens...

En moyenne chaque individu réalise 3 à 4 déplacements par jour dans une grande agglomération (graphique 3.1). Il semblerait donc que contrairement aux idées reçues, la mobilité est restée stable au cours des 20 dernières années (le conditionnel reste toutefois de rigueur, la difficulté du recueil des petits déplacements notamment à pied et la difficulté de faire comprendre la définition du déplacement au sens de l'enquête ménages déplacements fait dire à certains qu'une partie des écarts entre enquêtes est peut être également dû au plus ou moins grand professionnalisme avec lequel l'enquête est réalisée). Les évolutions sont pour partie contradictoires d'une agglomération à une autre, voire d'une enquête à une autre dans une même agglomération. Mais la tendance illustrée par ces données n'est pas à la croissance (sauf en ce qui concerne les deux dernières enquêtes de Lyon et Paris, qui traduiraient alors une inflexion de tendance si les données devaient être confirmées par d'autres enquêtes, comme cela semble être le cas au vu des résultats sur d'autres agglomérations de taille plus modeste). Cette idée de la croissance de la mobilité vient en fait souvent du développement du trafic automobile. Il faut toutefois distinguer ces deux notions qui ne recouvrent pas la même quantité.

Cette relative stabilité de la mobilité (d'après les données des enquêtes ménages déplacements) résulte de la conjonction de tendances contradictoires :

- développement de la journée continue qui supprime des déplacements le midi pour les personnes déjeunant sur leur lieu de travail ;

- vieillissement de la population, les personnes âgées se déplacent moins que les plus jeunes. Il faut toutefois considérer avec prudence ce type de données qui peut s'interpréter comme un effet d'âge ou bien comme un effet de génération (section 3.2, p. 88), conduisant alors à des conclusions différentes ;
- développement des grandes surfaces, qui réduit les déplacements nécessaires aux achats de proximité. De plus, ces achats sont souvent effectués en fin de semaine et ne sont donc pas pris en compte dans les statistiques présentées ;
- développement du travail féminin, qui accroît le nombre de déplacements réalisés par les femmes. En revanche, le développement du chômage a plutôt un effet inverse, tout comme un abaissement de l'âge de la retraite ;
- développement du temps libre qui peut conduire à réaliser davantage d'activités hors de chez soi, mais aussi à domicile. L'analyse des conséquences des 35 heures sera intéressante à cet égard ;
- spécialisation des espaces, qui peut conduire à accroître le nombre de déplacements pour réaliser autant d'activités ;
- développement du télé-travail, télé-achat et plus généralement des nouvelles technologies de communication qui pourraient conduire à diminuer la mobilité pour les motifs correspondants, mais pas forcément la mobilité totale. Il faut se méfier du mythe de la substitution des déplacements par les télécommunications (Claisse, 1997) ;
- motorisation qui offre une plus grande facilité de déplacement et peut accroître la mobilité.



Graphique 3.1 : Nombre de déplacements tous modes par jour et par personne

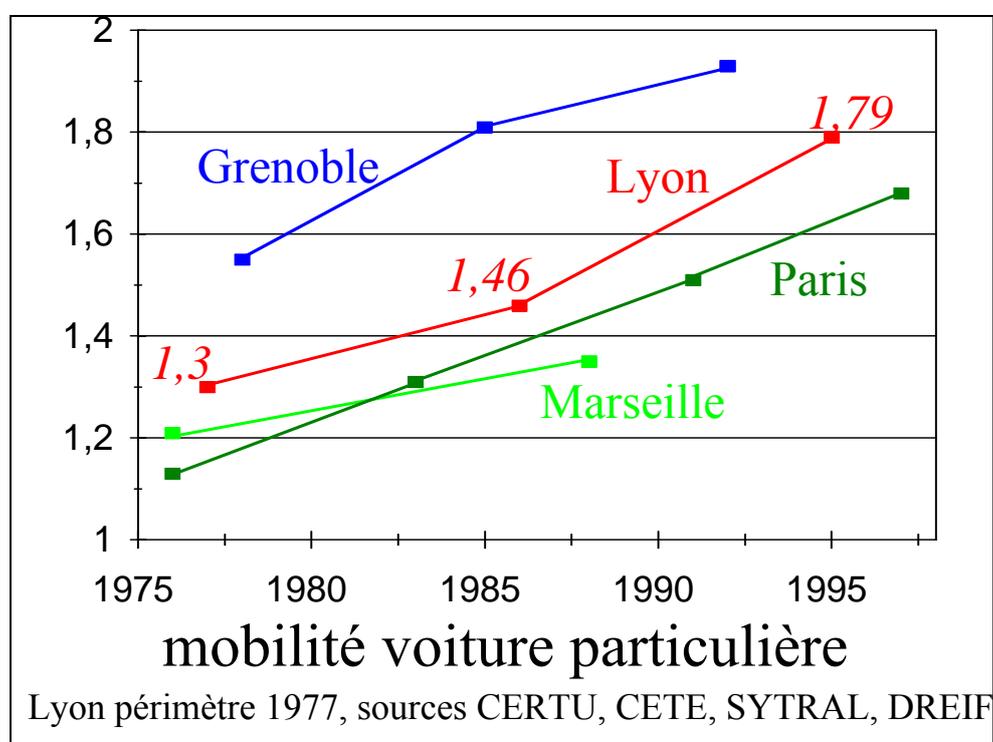
On pourrait allonger la liste, mais nous avons réuni selon nous les principales tendances lourdes d'évolution de la mobilité. On peut constater qu'elles sont toutes externes au sous-système des transports (à l'exception, dans une certaine mesure, de la motorisation). Cela ne signifie pas que le sous-système de transport n'a aucune influence sur le nombre de déplacements quotidiens par personne. Les nombreux exemples d'effets induits des infrastructures inciteraient à penser le contraire (il faut toutefois se méfier du mythe des effets induits, effets mécaniques d'une nouvelle infrastructure sur la mobilité (Bonnel, Thibaud, 1994)). Cela signifie plutôt que les principaux facteurs explicatifs de

la mobilité sont externes à ce sous-système. Ils sont beaucoup plus à rechercher du côté du sous-système des relations sociales à travers les activités réalisées et l'organisation de la société.

Après cet aperçu général, nous passons à une présentation modale : voiture particulière, transports collectifs, marche à pied et enfin deux-roues (graphiques 3.2 à 3.5).

2.2. ... Mais une forte croissance de l'usage de la voiture...

Si le nombre quotidien de déplacements ne s'accroît pas, ce n'est plus le cas de la mobilité voiture particulière qui connaît une forte croissance dans toutes les agglomérations, de près de 40 % sur 20 ans (graphique 3.2). Cette fois la tendance est identique dans les quatre villes étudiées et elle n'est pas propre à la France, même si les taux de croissance sont variables d'un pays à l'autre. Le nombre de déplacements semble décroître avec la taille de la ville (à l'exception de Marseille). Est-ce un effet de la congestion, de la plus grande proximité des activités, du système de transports collectifs ?... Les facteurs qui pourraient expliquer cette tendance sont assez nombreux.



Graphique 3.2 : Nombre de déplacements voiture particulière par jour et par personne

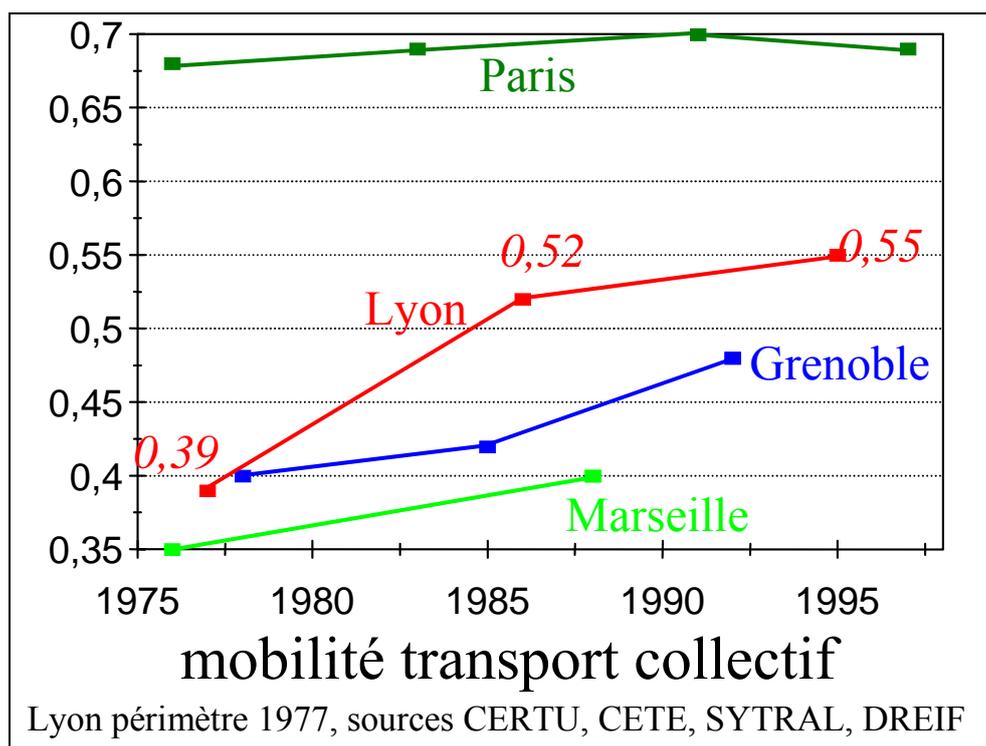
On peut avancer plusieurs facteurs pour tenter d'expliquer la croissance de la mobilité automobile :

- le développement de la motorisation. A moins que ne ce soit aussi le souhait ou la nécessité d'une certaine forme de mobilité permise par l'automobile qui explique la motorisation. La causalité n'est probablement pas totalement univoque. La relation, en revanche, est assez forte entre ces deux grandeurs ;
- le développement des infrastructures routières et de l'offre de stationnement tant privé que public dans la plupart des agglomérations (ce qui distingue la politique menée en France de celle menée dans d'autres pays du Nord de l'Europe et notamment en Suisse (Bonnel et al., 1994a)) ;

- l'étalement urbain de l'habitat et des activités (emploi, centres commerciaux), peu compatible avec une mobilité non automobile (la bi-motorisation est souvent généralisée dans les espaces péri-urbain) ;
- la souplesse qu'offre l'automobile dans la réalisation de schémas d'activités assez tendus ;
- l'évolution de l'organisation du travail avec des horaires plus variables peu favorables aux transports collectifs. Les 35 heures risquent d'amplifier cette tendance ;
- le développement du travail féminin qui s'accompagne souvent d'une motorisation des femmes ;
- en revanche, le développement de l'offre de transports collectifs pourrait contribuer à une évolution inverse. Toutefois, les autres facteurs sont beaucoup plus puissants de telle sorte que si le développement de l'offre de transports collectifs est nécessaire pour réduire l'usage de la voiture, il n'est en aucun cas suffisant pour atteindre un tel objectif (Bonnel et al., 1994a ; Bonnel, Chausse, 2000).

Nous pourrions poursuivre cette liste tant les facteurs expliquant cette croissance de l'usage de la voiture sont nombreux. A la différence de la mobilité globale, les facteurs explicatifs appartiennent cette fois aux trois sous-systèmes du système urbain. Ainsi une maîtrise de l'usage de la voiture doit passer par une action sur ces trois sous-systèmes pour atteindre son objectif, ce qui justifie une approche globale et cohérente d'un ensemble d'actions variées au sein de chacun des sous-systèmes.

2.3. ... Un accroissement plus modéré pour les transports collectifs...



Graphique 3.3 : Nombre de déplacements transports collectifs par jour et par personne

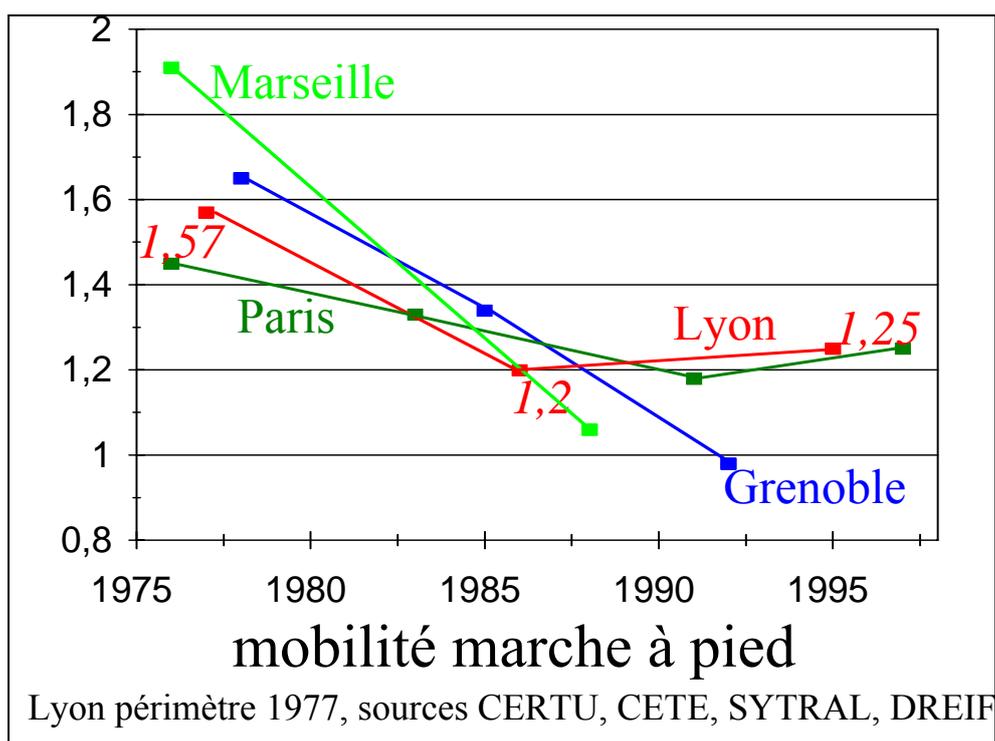
La mobilité transports collectifs est également en croissance (graphique 3.3). La croissance est toutefois plus modérée que celle de la voiture sur l'ensemble de la période. On observe de nouveau un effet taille, mais inverse de celui de la voiture. La mobilité transports collectifs s'accroît avec la taille de la ville (à l'exception de Marseille). Il y a sûrement un effet d'offre, mais probablement aussi un effet de la densité, de la congestion automobile...

La croissance des transports collectifs sur la période s'explique principalement par la croissance de l'offre, car la majorité des autres tendances structurelles sont plutôt défavorables à ce mode au profit de la voiture. Mais l'usage des transports collectifs n'est pas uniquement conditionné par l'offre proposée. Une comparaison internationale (Bonnel et al., 1994a) montre des rapports offre-demande significativement différents selon les pays. Il apparaît en fait que la politique mise en œuvre pour la circulation automobile et le stationnement est probablement l'un des principaux déterminants de l'usage des transports collectifs. Ainsi, si la qualité et la quantité de l'offre sont nécessaires pour un usage intensif des transports collectifs, ils ne sont en aucun cas suffisants.

2.4. ... Un effondrement de la marche...

La mobilité restant constante, la croissance de l'usage de la voiture et des transports collectifs, s'effectue au détriment d'autres modes, et en premier lieu de la marche (graphique 3.4). Celle-ci connaît une très forte décroissance similaire dans les agglomérations présentées (à l'exception de la dernière enquête pour Lyon et Paris, ce qui semble confirmer une inflexion de tendance récente de la mobilité qui reste toutefois à confirmer sur d'autres enquêtes).

Les déterminants de cette évolution sont en fait les mêmes que ceux qui ont fait le succès des deux autres modes et principalement de la voiture. De plus, la vie urbaine est de moins en moins centrée sur le quartier, notamment pour les achats ce qui limite la mobilité de proximité pouvant être aisément réalisée à pied.



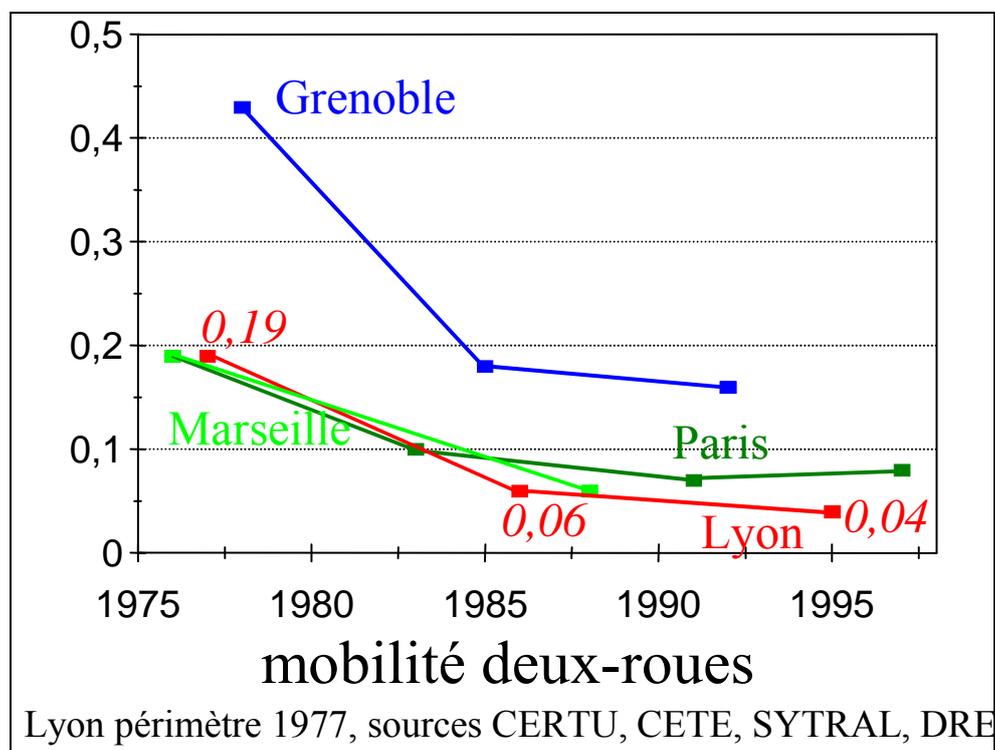
Graphique 3.4 : Nombre de déplacements marche à pied par jour et par personne

2.5. ... Et une quasi-disparition des deux-roues ?

Enfin, les deux-roues subissent également une chute spectaculaire à tel point qu'ils n'existent pratiquement plus au niveau de la mobilité quotidienne (graphique 3.5). L'inexistence d'une offre adaptée et sûre l'explique en partie. Grenoble, qui a fait plus d'effort en faveur de ce mode que la moyenne des villes, connaît une évolution similaire, même si la courbe reste au-dessus de celle des

autres agglomérations. De plus, l'ensemble des tendances structurelles qui expliquent la croissance de la voiture sont en fait plutôt défavorables à l'usage des deux-roues.

Est-il possible d'inverser la tendance ? Il est légitime de se poser la question au-delà d'un usage ludique ou sportif des deux-roues, même si les réponses apportées par les uns et les autres sont souvent contradictoires. Il reste que la croissance de leur usage est souvent souhaitée pour répondre aux problèmes de congestion ou d'environnement.



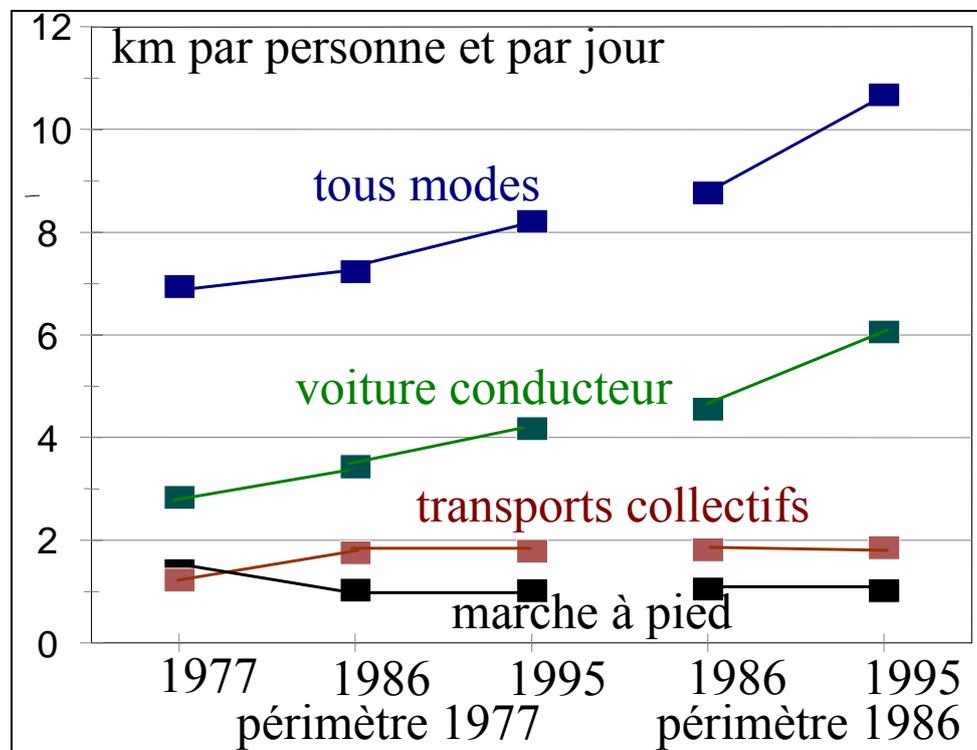
Graphique 3.5 : Nombre de déplacements deux roues par jour et par personne

2.7. Choix de l'indicateur de description de la mobilité

Notre ambition n'est pas de réaliser une analyse approfondie de la mobilité. On peut toutefois s'interroger sur la pertinence de l'indicateur nombre de déplacements pour décrire l'évolution de la mobilité. Les préoccupations environnementales sont de plus en plus prégnantes dans les procédures de planification. Un indicateur de distance parcourue serait probablement plus pertinent pour répondre à cet objectif. C'est une autre image de la mobilité qui apparaît alors. La stabilité relative du nombre de déplacements par personne laisse place à une croissance des kilomètres parcourus (graphique 3.6). La croissance reste assez modérée lorsque l'on reste à l'intérieur du périmètre de l'enquête de 1977 (périmètre plus restreint que la Communauté Urbaine de Lyon). En revanche, l'élargissement au périmètre de 1986 (équivalent au périmètre du SDAU), conduit à une croissance plus soutenue. Il est probable qu'un périmètre englobant l'ensemble de l'aire d'attraction de l'agglomération conduirait à une croissance encore un peu plus forte.

L'examen par mode de déplacement accroît encore le changement par rapport à l'indicateur nombre de déplacements. La chute pour la marche à pied est globalement identique à celle du nombre de déplacements, la distance moyenne par déplacement ne diminuant que très faiblement. La croissance des distances parcourues en transports collectifs est légèrement plus faible que celle du nombre de déplacements, car, on observe une diminution de près de 7 % de la distance moyenne par déplacement. En revanche, pour la voiture, on assiste à une explosion des distances. A l'accroissement

du nombre de déplacements s'ajoutent une croissance des distances parcourues par déplacement de près de 10 %. Cet effet s'amplifie à mesure que le périmètre s'élargit. Ainsi, 71 % des kilomètres effectués par les résidents du périmètre de l'enquête de 1986, à l'intérieur de ce périmètre, sont effectués en voiture (comme conducteur ou comme passager), alors que ce mode ne regroupe que 52 % des déplacements.



Graphique 3.6 : Distance quotidienne par personne selon le mode

(Sources : d'après enquêtes ménages déplacements de Lyon 1977, 1986, 1995, périmètre enquête 1977 (Bonnel, 2000a))

3. TENDANCES STRUCTURELLES D'EVOLUTION DE LA MOBILITE : EXEMPLE D'ANALYSE, EXEMPLE DE QUANTIFICATION

L'objectif de cette présentation n'est pas de décrire l'ensemble des facteurs influant de manière structurelle sur la mobilité. Nous privilégions trois dimensions : l'étalement urbain, la motorisation et le vieillissement de la population. Plusieurs raisons expliquent ce choix. Tout d'abord, il s'agit des facteurs ayant le plus le plus d'incidence sur l'évolution de la mobilité (entendu non plus en tant que nombre de déplacements, mais en tant qu'inscription dans l'espace et le temps d'un programme d'activités) (Kostyniuk, Kitamura, 19887 ; Banister, Bayliss, 1992 ; Orfeuil, 1992 ; Banister, 1993 ; Bonnafous, 1993 ; Tabourin, 1995 ; Bonnafous, Tabourin, 1998). De plus, les deux premiers facteurs sont présents dans la plupart des modèles de prévision de la demande au moins dans les trois premières étapes. Enfin, ils illustrent une nouvelle fois la nécessité d'appréhender la mobilité à travers une confrontation des trois sous-systèmes constituant le système urbain.

Nous ne souhaitons pas faire la synthèse des nombreux travaux menés sur ces trois facteurs. Notre choix se porte plutôt sur la présentation de deux exemples d'analyses débouchant sur une quantification et des possibilités de développement d'un modèle de prévision. Nous privilégions la dimension pédagogique dans la présentation de ces travaux. En effet, ils soulèvent des problèmes que l'on rencontre généralement dans la plupart des analyses de mobilité : corrélation entre les facteurs explicatifs, conjonction de phénomènes agissant de manière contradictoire masquant ainsi les

évolutions, prise en compte de la dimension temporelle nécessitant de séparer les effets d'âge et de génération, complexité des relations de causalité. La dernière étude est instructive à cet égard. Nous montrons comment à travers trois angles d'approche différents d'un même phénomène, on peut déboucher sur trois conclusions radicalement différentes.

3.1. Quantification des effets de l'évolution de la localisation des déplacements, de la motorisation et de l'offre de transport sur la part de marché des transports collectifs

L'objectif de cette première étude (Bonnel, Cabanne, 2000 ; Bonnel et al., 2001) est de quantifier l'incidence sur l'évolution de la part de marché des transports collectifs de 5 facteurs :

- l'étalement urbain, mesuré à travers l'évolution des localisations des origines-destinations des déplacements ;
- la motorisation ;
- l'offre de transports collectifs ;
- l'offre de transport privé ;
- un ensemble d'autres facteurs non identifiés explicitement.

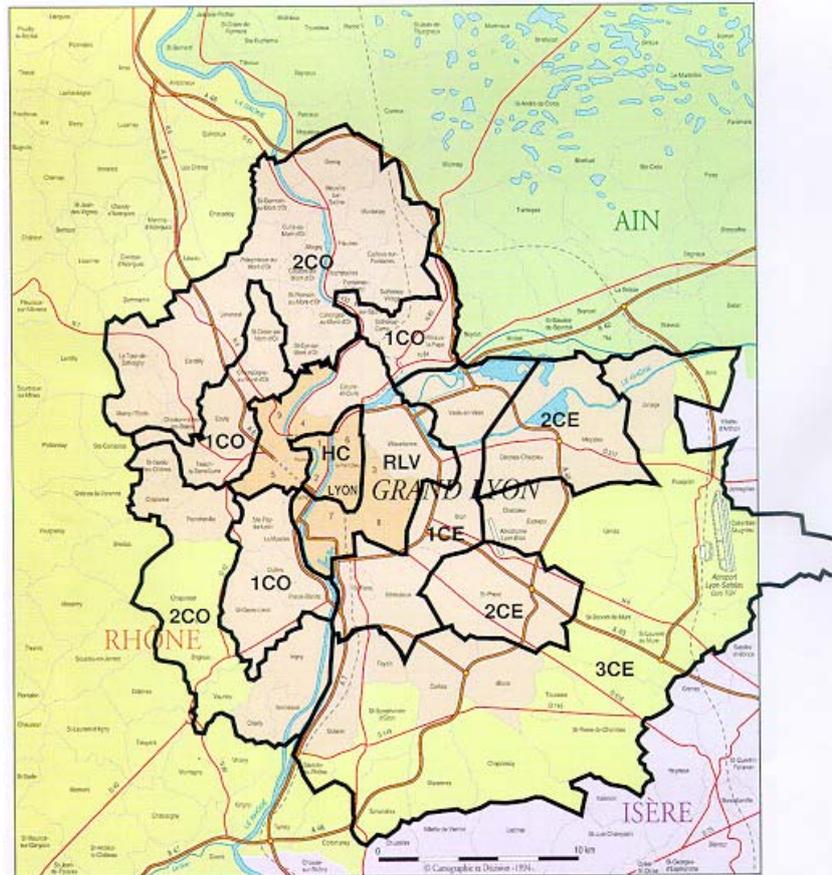
Ces facteurs étant fortement corrélés entre eux, il est nécessaire de proposer une méthode permettant de séparer la contribution respective de chacun des facteurs (section 3.1.1). Nous présentons ensuite l'évolution des facteurs étudiés et leur formulation (section 3.1.2), puis le choix d'une fonction de partage modal (section 3.1.3). Enfin, nous explicitons les principaux résultats (section 3.1.4) et les enseignements que nous pouvons retirer de cette recherche (section 3.1.5).

3.1.1. Principe de mesure et terrain d'étude

Nous utilisons les données des trois dernières enquêtes ménages déplacements réalisées sur l'agglomération lyonnaise en 1977, 1985 et 1995. Si les méthodologies d'enquête sont comparables, les périmètres retenus ne le sont pas. Le périmètre 1995 est un petit peu plus large que le périmètre 1985, qui est lui-même plus large que le périmètre 1977.

Pour évaluer les impacts des différents facteurs de choix modal à l'échelle de l'agglomération, nous découpons le périmètre d'étude en zones. Nous travaillons sur deux découpages différents : un découpage en 7 zones (carte 3.1, seules les zones 1 à 5 sont internes au périmètre 77) et un autre, plus fin, en 25 zones (Lichère, Raux, 1997a). Ce découpage en 25 zones permet une meilleure précision au niveau de la définition des temps de parcours entre zones et permet une multiplication des points d'observation. En revanche, les volumes de flux sont faibles sur de nombreuses origines-destinations. Ces deux découpages sont définis en s'appuyant sur une logique de couronnes, permettant de tenir compte de l'étalement urbain, à laquelle s'ajoute une segmentation Est-Ouest pour le découpage en 7 zones et une logique de bassin versant pour le découpage en 25 zones obtenu à partir de celui en 7 zones :

- zone 1 : hypercentre de Lyon (presqu'île + Part Dieu), notée (HC) ;
- zone 2 : reste de Lyon-Villeurbanne (zone interne au premier périphérique), notée (RLV) ;
- zone 3 : 1^{ère} couronne Est, notée (1CE) ;
- zone 4 : 1^{ère} couronne Ouest, notée (1CO) ;
- zone 5 : 2^{ème} couronne Est, notée (2CE) ;
- zone 6 : 2^{ème} couronne Ouest, notée (2CO) ;
- zone 7 : 3^{ème} couronne Est, notée (3CE) ;
- Ext : ensemble des localisations externes à ce périmètre.



Carte 3.1 : Périmètre de l'enquête ménages déplacements et découpage en 7 zones
 (Source : carte élaborée d'après l'Atlas du Grand Lyon, INSEE, Cartographie et Décision, 1994,
 (Lichère, Raux, 1997a))

A partir de ce découpage, nous constituons des matrices de volumes de flux en transport en commun et en voiture particulière entre les zones, des matrices de motorisation moyenne des individus effectuant le déplacement sur les différentes origines-destinations, des matrices d'offre en transport en commun et en voiture particulière entre zones. Nous calons alors un modèle de partage modal (un modèle logit (chapitre 8, p. 227)) sur les données ainsi constituées.

Les différentes variables explicatives sont généralement corrélées, de même que leurs évolutions. Cela soulève des difficultés en termes d'évaluation et d'interprétation des impacts respectifs de ces différents facteurs. Ainsi, motorisation et localisation des déplacements sont corrélés. La motorisation étant plus forte en périphérie, l'étalement urbain constaté depuis quelques décennies concourt à l'accroissement de la motorisation. Lorsqu'on simule l'incidence combinée de la motorisation et de l'étalement urbain, si l'on somme l'effet de chacune des variables, on risque fort de faire des doubles comptes sur l'évolution de la motorisation.

Pour faire face à cette difficulté, nous proposons une méthode de décomposition des effets qui permet de séparer la contribution de chacun des facteurs. Cette méthode repose sur deux principes. Le premier requiert des définitions originales de nos différentes variables afin de limiter les corrélations entre évolutions des variables. Nous proposons ainsi une définition de la motorisation à structure spatiale donnée. Cela nous permet de rendre indépendants évolution des localisations et évolution des motorisations. Le deuxième principe est d'étudier de façon systématique les effets simples et les effets combinés de nos différentes variables (encadré 3.1).

Encadré 3.1 : Effets simples, effets combinés et séparation des effets

Supposons qu'une variable à expliquer Y s'explique par une formulation de deux facteurs a et b ($Y = \sum_i f(a_i, b_i)$), somme que nous notons $Y = F(ab)$ par souci de simplification).

Entre deux dates d'enquête notées 1 et 2, l'effet simple E(a) du premier facteur correspond à la variation de Y due à la variation du premier facteur, le deuxième facteur restant inchangé. L'effet simple E(b) du deuxième facteur correspond à la variation de Y due à la variation du deuxième facteur, le premier facteur restant inchangé. Toutefois, le taux de variation de Y ne peut pas s'écrire simplement comme le produit de ces deux effets simples. Il existe un résidu E(a,b), effet combiné des deux facteurs :

$$\frac{Y_2}{Y_1} = \frac{F(a_2 b_2)}{F(a_1 b_1)} = \frac{F(a_2 b_1)}{F(a_1 b_1)} \times \frac{F(a_1 b_2)}{F(a_1 b_1)} \times \frac{F(a_2 b_2)}{F(a_2 b_1)} = E(a) \times E(b) \times E(a, b) \quad (3.1)$$

Si les effets des deux facteurs sont indépendants, cet effet double a toute chance d'être proche de 1. La variable à expliquer s'exprime alors sous forme de deux effets indépendants. Nous pouvons négliger l'effet combiné des deux variables et parler de séparation des effets. L'objectif est donc de trouver des définitions des facteurs et de la relation qui les lie, qui les rendent les plus indépendants possibles et de trouver une formulation de la variable à expliquer qui minimise cet effet combiné.

Dans la présente étude nous n'avons pas 2 mais 5 variables explicatives. Toutefois, nous pouvons généraliser sans problème (Bonnel, Cabanne, 2000). Avec l la localisation, m la motorisation, vp l'offre en voiture particulière, tc l'offre en transport en commun et a les autres facteurs, l'effet localisation est défini de la façon suivante :

$$E(l) = \frac{F(l_2, m_1, vp_1, tc_1, a_1)}{F(l_1, m_1, vp_1, tc_1, a_1)} \quad (3.2)$$

Les effets simples E(m), E(vp), E(tc), E(a) sont définis de façon analogue.

La variation de part de marché $\frac{P_{TC2}}{P_{TC1}}$ s'exprime alors de la façon suivante :

$$\frac{P_{TC2}}{P_{TC1}} = E(l) \times E(m) \times E(vp) \times E(tc) \times E(a) \times \underbrace{\text{résidu}}_{26 \text{ facteurs}} \quad (3.3)$$

Le résidu contient 26 facteurs : 10 effets doubles, 10 triples, 5 quadruples et 1 quintuple. Pour expliquer la part des transports en commun comme étant le produit des différents effets simples, ces 26 facteurs doivent être aussi proches de 1 que possible.

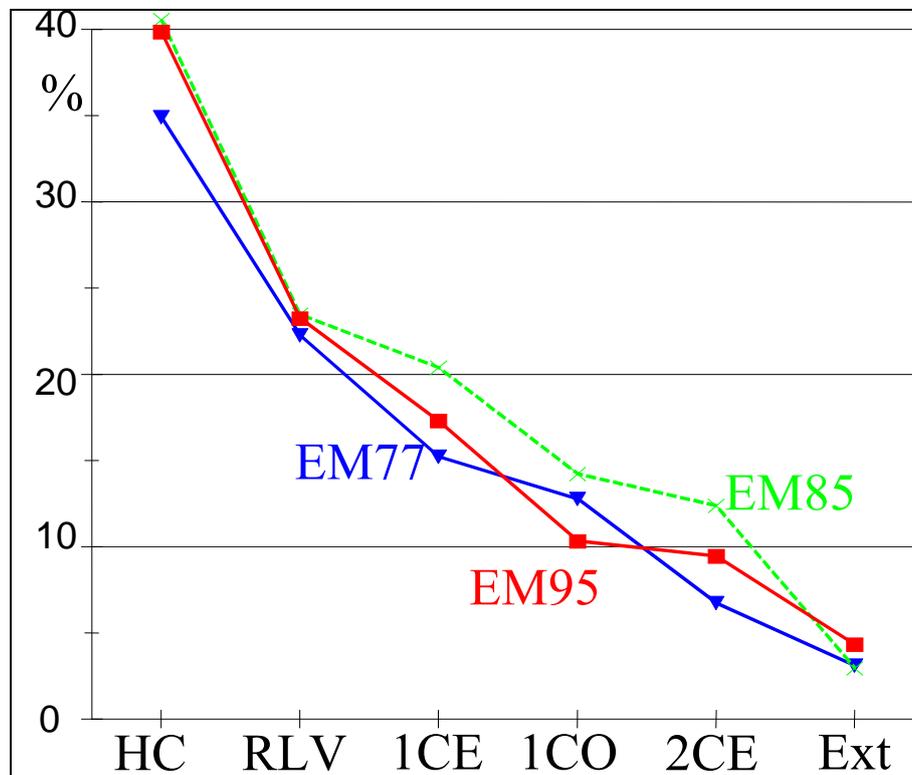
Après avoir explicité notre méthode de décomposition, nous proposons des définitions de nos différents facteurs qui limitent les corrélations.

3.1.2. Evolution et définition des facteurs étudiés

De 1977 à 1985 sur le périmètre 1977, la part des transports collectifs a augmenté de 22 % à 24 % pour retomber à 21 % en 1995. Cette période a été marquée par plusieurs phénomènes jouant dans des sens différents : étalement urbain, accroissement de la motorisation, amélioration forte de l'offre de transport en commun avec la mise en service du métro.

3.1.2.1. Eclatement des flux et définition du facteur localisation

A l'échelle de l'agglomération, la part des transports en commun a régressé entre 1985 et 1995. Pourtant, la part des transports en commun s'est maintenue sur de nombreuses origines-destinations, ou tout du moins sur les radiales (graphique 3.7).



Graphique 3.7 : Part de marché des TC sur (TC+VP) pour les déplacements ayant une origine ou une destination dans la zone..., périmètre 77

(Sources : enquêtes ménages déplacements Lyon 1977, 1985, 1995, voir carte 3.1 pour le zonage)

Mais dans le même temps, la part des déplacements ayant une extrémité en banlieue a augmenté (graphiques 3.8 et 3.9). Comme la part de marché des transports publics urbains est plus faible sur ce type des déplacements, la part de marché des transports en commun a globalement baissé sur l'ensemble de l'agglomération. Ceci montre bien l'importance de l'étalement urbain dans l'évolution de la part de marché des transports collectifs.

Pour démêler l'évolution de l'attractivité des transports en commun à origine-destination donnée et l'éclatement des flux, nous considérons le zonage précédemment défini. Nous pouvons alors construire une matrice origines-destinations l_{ij} où l_{ij} est la part des déplacements sur l'origine destination de i vers j parmi l'ensemble des déplacements (en transports collectifs et en voiture particulière) à l'échelle de l'agglomération.

Cette matrice l_{ij} est appelée matrice des localisations. L'évolution du terme localisation ne correspond donc pas exactement à l'étalement urbain au sens habituel de l'évolution de la localisation des activités dans l'espace urbain. Il correspond plutôt aux conséquences de cet étalement sur la géographie des flux parallèlement à l'évolution des choix de destination des individus. Cette définition a le mérite d'être opératoire et de permettre une formalisation très simple sur le plan mathématique,

tout en exprimant non pas l'évolution des localisations des activités, mais l'évolution des localisations des origines-destinations des déplacements.

Nous pouvons alors écrire la part de marché des transports collectifs en fonction des localisations grâce à la formulation suivante :

$$P_{TC} = \sum_{i,j} l_{ij} * P_{TCij} \quad (3.4)$$

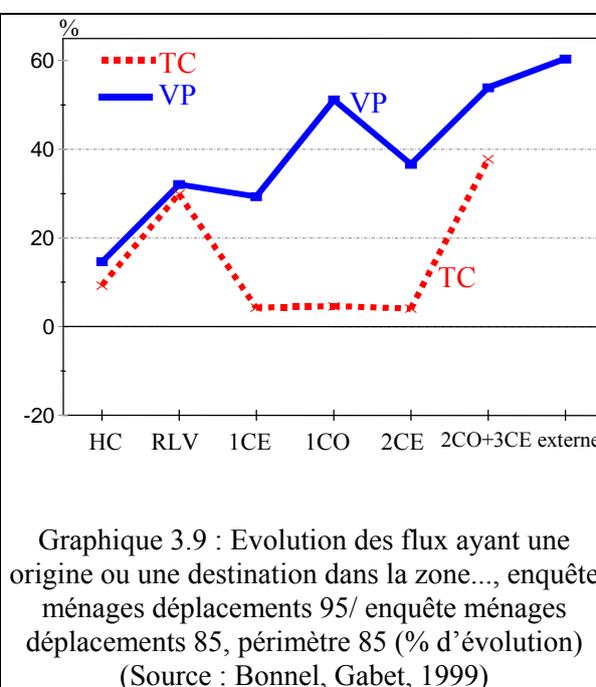
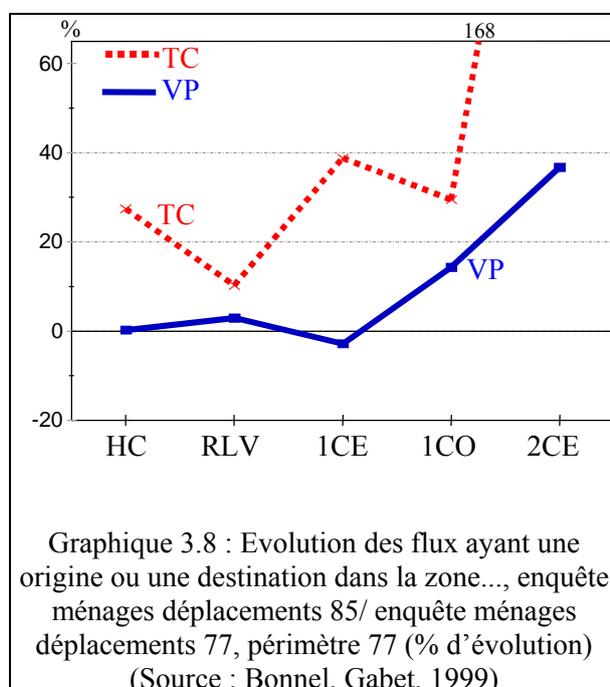
où P_{TC} est la part de marché des transports collectifs sur l'ensemble de l'agglomération ;

l_{ij} est la part des déplacements de l'origine-destination i/j parmi l'ensemble des déplacements (transports collectifs et voiture particulière) à l'échelle de l'agglomération ;

P_{TCij} est la part de marché des transports collectifs pour l'origine-destination entre i et j .

Les P_{TCij} peuvent s'écrire comme une fonction g (à déterminer) des autres facteurs explicatifs :

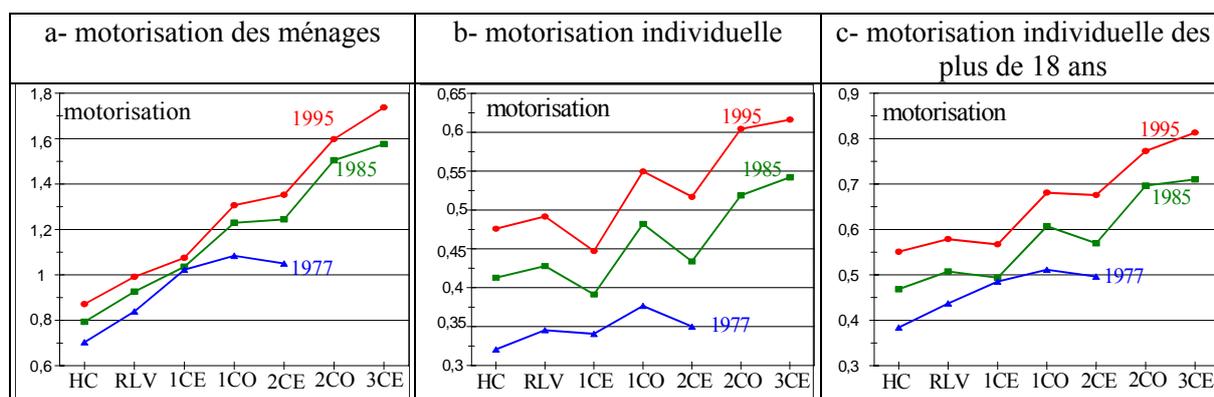
$$P_{TCij} = g(\text{mot}_{ij}, \text{vp}_{ij}, \text{tc}_{ij}, \text{a}_{ij}) \quad (3.5)$$



3.1.2.2. Facteur motorisation

La définition de la motorisation habituellement retenue en France est le nombre moyen de voitures par ménage (graphique 3.10a). Cette définition de la motorisation conduit à une croissance assez limitée de la motorisation dans le temps entre les trois enquêtes. En revanche, la corrélation entre la motorisation et la localisation est extrêmement forte. La motorisation de la plus lointaine banlieue est double de celle de la zone la plus centrale. Cette caractérisation de la motorisation des ménages est toutefois insuffisante pour préciser l'accès réel des individus à l'automobile. La taille des ménages n'est généralement pas uniforme dans l'espace. Le nombre de personnes susceptibles d'utiliser un véhicule dans le ménage est généralement plus important en périphérie que dans le centre. De plus, la taille des ménages a fortement chuté au cours des 20 dernières années. A motorisation des ménages donnée, l'accès individuel à la voiture est donc plus important aujourd'hui qu'il y a 20 ans.

Nous retenons donc une seconde définition de la motorisation, qui est la motorisation individuelle (graphique 3.10b). Elle est obtenue en divisant le nombre de voitures du ménage par le nombre de personnes enquêtées du ménage. Avec cette nouvelle définition, la croissance de la motorisation est très forte entre les trois enquêtes. En revanche, la corrélation entre localisation et motorisation est fortement réduite. Le rapport de motorisation entre la zone la plus éloignée et le centre n'est plus que de 1,4. Toutefois, on peut se demander si cette définition est très pertinente pour caractériser l'accès réel des individus à la voiture. Seules les personnes de 18 ans et plus sont susceptibles de conduire.



Graphique 3.10 : Evolution de la motorisation selon la zone de résidence et la définition de la motorisation (périmètre 1977)

(Sources : Bonnel, Gabet, 1999 ; découpage zonal carte 3.1 ; d'après les enquêtes ménages déplacements de Lyon 1977, 1985, 1995)

Nous retenons donc une troisième définition : la motorisation des individus de plus de 18 ans (graphique 3.10c). Elle est définie comme le rapport entre le nombre de voitures du ménage et le nombre d'individus de 18 ans et plus du ménage. Avec cette nouvelle définition, la croissance entre les trois enquêtes devient plus modérée. La corrélation redevient un peu plus forte. Toutefois, la motorisation de la zone la plus éloignée n'est supérieure que de 50 % à celle de la zone la plus centrale. La nouvelle définition de la motorisation individuelle des plus de 18 ans réduit la corrélation entre les deux variables par rapport à la motorisation du ménage. En effet, le nombre d'individus par ménage étant plus élevé en périphérie que dans le centre, les différences d'accès réel à la voiture sont moins marquées spatialement qu'avec une définition de la motorisation au niveau du ménage. Néanmoins une corrélation entre localisation et motorisation subsiste. Or nous cherchons à séparer les impacts des différents facteurs influant sur le choix modal. Avec la définition précédente, nous risquons de « lire » un effet évolution des localisations derrière les évolutions de la motorisation ou inversement.

Pour résoudre le problème posé par la corrélation entre l'évolution de la localisation et de la motorisation, nous proposons une formulation de la motorisation qui fait abstraction de la localisation. Nous le faisons en transférant l'attribut motorisation qui est attaché au ménage ou à l'individu, aux déplacements réalisés par ces ménages ou individus. Il est alors possible de calculer une motorisation pour chaque origine-destination en faisant la moyenne de la motorisation de chacun des déplacements concerné par l'origine-destination (encadré 3.2).

Cette formulation permet de faire varier indépendamment localisation et motorisation. Nous baptisons cette définition non habituelle de la motorisation, matrice des motorisations individuelles spatialisées.

Ce que nous nommons « effet motorisation » ne correspond donc pas forcément à la totalité des « effets de la motorisation » sur l'usage des modes. L'évolution de la motorisation peut avoir des incidences sur les choix de localisations qui sont alors inscrits dans notre étude comme effet localisation et non effet motorisation, tout comme les choix de localisations peuvent avoir une incidence sur la motorisation. Nous décomposons ici l'évolution de la motorisation à l'échelle de l'agglomération en deux composantes : évolution de la motorisation liée à l'évolution de la structure spatiale des flux ; évolution de la motorisation à structure spatiale donnée. Ce que nous nommons effet motorisation ne correspond qu'à l'effet de la seconde composante. Le premier facteur est contenu dans l'effet localisation, pour éviter les double comptes.

Encadré 3.2 : Formulation mathématique du taux de motorisation individuelle spatialisée

Soit t_k , le rapport nombre de voitures du ménage/nombre de personnes de 18 ans et plus du ménage auquel appartient l'individu réalisant le déplacement k sur l'origine-destination i/j ;

N_{ij} , le nombre de déplacements de l'origine-destination i/j .

La motorisation associée aux déplacements de i vers j est alors :

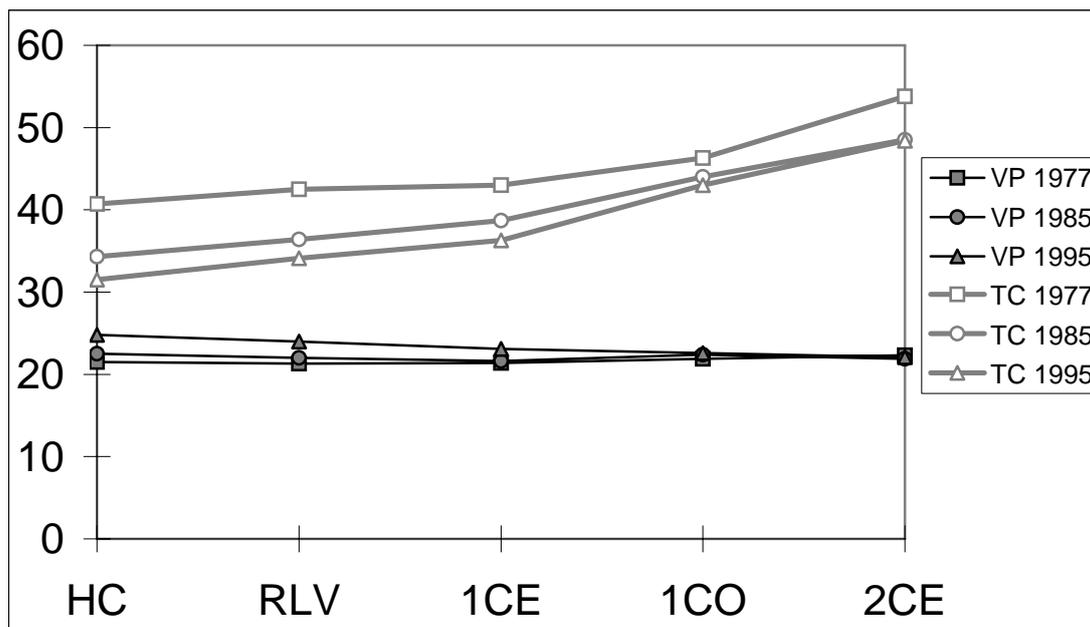
$$m_{ij} = \frac{\sum_{k=1}^{N_{ij}} t_k}{N_{ij}} \quad (3.6)$$

3.1.2.3. Offre de transport

Les deux décennies ont été marquées, à Lyon, par la mise en service du métro en 1978 puis par son extension au cours des années 1980 et 1990 (Raux et al., 1996). Les temps de parcours en transport en commun ont ainsi baissé de 15 % au cours des 20 dernières années particulièrement sur les trajets ayant une origine ou une destination en centre-ville. Les temps de parcours en voiture particulière ont légèrement augmenté principalement pendant la deuxième décennie.

Pour introduire l'offre de transport dans le modèle, nous calculons pour chacune des origines-destinations le temps généralisé en transports collectifs (noté tc_{ij}) et le temps généralisé en voiture particulière (noté vp_{ij}). Pour les transports en commun, ce temps généralisé agrège temps de parcours et recette moyenne d'un trajet en transport en commun (divisée par la valeur du temps). Pour la voiture particulière, il agrège temps de parcours et prix du carburant. La contrainte de stationnement n'est pas incluse dans l'offre de transport en voiture particulière, faute de posséder des données relatives à ce facteur. La contrainte de stationnement est donc incluse dans les « autres facteurs » non explicités. Le graphique 3.11 présente l'évolution des temps généralisés voiture particulière et transport en commun pour les déplacements ayant une (ou les deux) extrémité(s) dans une zone.

Lors de la première décennie, les baisses de temps de parcours en transport en commun sont assez importantes surtout dans les zones les plus centrales. En revanche, pendant la seconde décennie, les baisses sont beaucoup plus modérées. Pour la voiture, on constate une très légère progression surtout dans les zones les plus centrales au cours de la seconde décennie.



Graphique 3.11 : Evolution des temps généralisés voiture et transports collectifs de 1977 à 1995 pour les déplacements ayant une extrémité dans la zone ... (périmètre 1977)
(Source : d'après SEMALY, LET (Lichère, Raux, 1997a, b))

3.1.2.4. « Autres facteurs »

Nous ne connaissons pas le contenu des autres facteurs. Ils seront donc calculés par déduction.

La part des transports en commun à l'échelle de l'agglomération s'écrit :

$$P_{TC} = \sum_{ij} l_{ij} * P_{TCij} = \sum_{ij} l_{ij} * g(\text{mot}_{ij}, \text{vp}_{ij}, \text{tc}_{ij}, a_{ij}) \quad (3.7)$$

où l_{ij} est le poids du couple origine-destination parmi l'ensemble des déplacements réalisés en voiture ou en transports collectifs ;

P_{TCij} est la part de marché des transports collectifs pour l'origine-destination entre i et j ;

mot_{ij} est la motorisation des déplacements pour l'origine-destination entre i et j ;

vp_{ij} est le coût généralisé en voiture particulière des déplacements pour l'origine-destination entre i et j ;

tc_{ij} est le coût généralisé en transport en commun des déplacements pour l'origine-destination entre i et j ;

a_{ij} est le terme « autres facteurs » pour l'origine-destination entre i et j .

Les données des enquêtes ménages déplacements ou de l'étude SEMALY-LET (Lichère, Raux, 1997a, b) permettent de calculer chacune des matrices. Seuls les a_{ij} sont inconnus. Après avoir choisi puis calé une fonction g pertinente, nous en déduisons les a_{ij} par résolution de : $P_{TCij} = g(\text{mot}_{ij}, \text{vp}_{ij}, \text{tc}_{ij}, a_{ij})$ où les a_{ij} sont les inconnues.

Nous calculons alors les différents effets de nos 5 variables.

3.1.3. Choix d'une fonction de partage modal

Pour finaliser la formulation du partage modal, il est nécessaire de choisir une fonction reliant le partage modal sur chaque origine-destination aux variables explicatives de motorisation, offre de transport et autres facteurs.

Cette fonction est classiquement une fonction logit. Elle présente le mérite de fonder sur le plan théorique la formulation grâce à la théorie de l'utilité (chapitres 8, p. 227 et 9, p. 252, Ben-Akiva, Lerman, 1985 ; Domencich, McFadden, 1975) et est la formulation la plus couramment utilisée en modélisation pour exprimer le partage modal :

$$P_{TCij} = \frac{1}{1 + \exp(u)} \quad (3.8)$$

où u représente la différence d'utilité entre les deux modes de transport estimée à partir des variables mot , vp , tc , a .

Tout l'enjeu réside dans la détermination de u . Le choix de la fonction d'utilité est en effet fondamental. Il conditionne largement la pertinence des évaluations retenues et la validité de l'estimation des effets. Ces fonctions d'utilité doivent être pertinentes sur le plan de la théorie économique, elles doivent permettre de bien reconstituer les parts modales réelles et doivent minimiser les différents effets combinés.

Plusieurs formulations de la fonction d'utilité peuvent être retenues (encadré 3.3).

La formulation A1 est une formulation additive classique. Il est toutefois possible de s'interroger sur le mécanisme de choix entre la voiture et les transports collectifs. Il est probable que la comparaison entre les modes relève davantage d'une comparaison des performances en termes relatifs plutôt qu'en termes absolus. Nous testons donc une formulation dans laquelle les coûts généralisés des transports collectifs et de la voiture interviennent sous forme de rapport (formulation P1).

Dans ces formulations, la motorisation individuelle spatialisée intervient de manière additive. Certains travaux (Lichère, Raux, 1997a, b) ont montré que la motorisation pouvait aussi être interprétée comme un facteur influant sur la perception des deux modes motorisés. Plus la motorisation est élevée, plus la perception du coût généralisé en voiture est réduite. Et inversement plus la motorisation est élevée et plus le coût généralisé en transports collectifs est perçu négativement. Cette interprétation de la motorisation nous conduit de la formulation A1 à la formulation A2.

Les variables B_1 et B_2 permettent de schématiser la difficulté à stationner dans les zones les plus centrales (dans les autres zones, on peut considérer que la contrainte de stationnement sur l'espace public est très faible). En introduisant de façon schématique une variable de stationnement dans le calage de u , nous améliorons la fiabilité des paramètres de u sur les autres variables.

D'autres formulations peuvent être retenues (Bonnel, Cabanne, 2000) mais elles donnent des résultats souvent moins satisfaisants : mauvaises reconstitutions des parts modales, effets combinés trop importants.

Encadré 3.3 : Fonctions d'utilités utilisées

$$A1 : u = a \text{ mot} + b \text{ vp} + c \text{ tc} + b_1 B_1 + b_2 B_2 + d + \varphi(a) \quad (3.9)$$

$$P1 : u = a \text{ mot} + b \frac{tc}{vp} + b_1 B_1 + b_2 B_2 + c + \varphi(a) \quad (3.10)$$

$$A2 : u = a \frac{vp}{mot^m} + b tc mot^m + c + b_1 B_1 + b_2 B_2 + \varphi(a) \quad (3.11)$$

avec :

mot : taux de motorisation individuelle spatialisée ;

vp : coût généralisé vp ;

tc : coût généralisé tc ;

B_1 et B_2 sont deux variables destinées à schématiser la contrainte de stationnement :

B_1 telle que :

$B_1 = 1$ si 1 (l'hypercentre) est la zone d'origine ou de destination ;

$B_1 = 0$ sinon ;

B_2 telle que :

$B_2 = 1$ si 2 (le reste de Lyon et Villeurbanne) est la zone d'origine ou de destination ;

$B_2 = 0$ sinon.

La forme de la fonction d'utilité étant choisie, nous calons les paramètres de u par passage aux logarithmes, en supposant que les $\varphi(a)$ sont en moyenne nuls. Le calage est toutefois une opération délicate. Si sur une origine-destination donnée, les volumes enquêtés sont trop faibles, les parts modales estimées seront peu fiables. Il paraît alors nécessaire de rejeter l'origine-destination lors du calage sous peine de biaiser l'estimation des coefficients. Inversement, lorsque le nombre d'origines-destinations rejetées est trop important, les coefficients estimés seront peu fiables faute d'un nombre de degrés de liberté suffisant dans la régression. Dans notre étude (Bonnel, Cabanne, 2000), nous avons effectué le calage sur différents découpages plus ou moins fins de l'agglomération et pour des seuils d'acceptation des volumes enquêtés par origines-destinations différents. Nous avons testé ainsi la robustesse de nos coefficients. Les valeurs obtenues sont les suivantes :

$$A1 : u = a mot + b vp + c tc + b_1 B_1 + b_2 B_2 + d + \varphi(a)$$

	Constante	a (mot)	b (vp)	c (tc)	b_1	b_2
Coefficient	-0,13	3,26	-0,02	0,02	-1,11	-0,50
Statistique t	-0,25	4,20	-3,48	3,45	-10,60	-6,12

Calage par régression linéaire : $R^2 = 0,93$.

$$P1 : u = a \text{ mot} + b \frac{tc}{vp} + b_1 B_1 + b_2 B_2 + c + \varphi(a)$$

	Constante	a (mot)	$b \frac{tc}{vp}$	b_1	b_2
Coefficient	-1,17	3,47	0,52	-1,05	-0,45
Statistique t	-1,99	4,83	4,49	-10,26	-5,59

Calage par régression linéaire : $R^2 = 0,93$.

$$A2 : u = a \frac{vp}{\text{mot}^m} + b tc \text{ mot}^m + c + b_1 B_1 + b_2 B_2 + \varphi(a)$$

Coefficient	Constante	m (mot)	a (vp)	b (tc)	b_1	b_2
Valeur du coefficient	1,97	1,32	- 0,01	0,03	- 1,13	- 0,52
Statistique t	11,85	2,84	-2,23	4,86	-10,27	-6,36

Calage par utilisation de l'algorithme de Newton-Raphson (car l'utilité n'est plus de formulation linéaire) : $R^2 = 0,87$.

Une fois le calage effectué, nous n'avons plus qu'à calculer les différents effets simples. C'est l'objet de la section suivante.

3.1.4. Des effets simples importants

Pour les trois formulations considérées, les effets combinés sont négligeables. Nous ne présentons que les effets simples.

L'effet localisation est identique quelle que soit la formulation puisqu'il ne dépend pas du calage de la fonction d'utilité. En définissant l'impact des localisations $w(l)$ comme suit : $w(l) = E(l) - 1$, avec $E(l)$ effet des localisations défini à la section 3.1.2 (équations 3.2, p. 78 et 3.4, p. 80), et $w(l)$ exprimé en pourcentage, nous obtenons les impacts figurant dans le tableau 3.1. Il apparaît ainsi que durant la première décennie, les changements de localisation des origines-destinations des déplacements expliquent une baisse de part de marché des transports collectifs parmi les modes motorisés de 3,9 %. Cet effet s'est encore amplifié durant la seconde décennie avec une chute de 6,5 %. L'étalement urbain et la périphérisation des flux, en augmentant la demande de déplacements dans les zones où les transports collectifs sont peu performants au détriment des zones centrales, ont ainsi contribué à une chute de 8,6 % de la part de marché des transports collectifs sur 20 ans.

	Impact entre les enquêtes ménages déplacements		
	1977 et 1985	1985 et 1995	1977 et 1995
impact du facteur localisation	-3,9 %	-6,5 %	-8,6 %

Tableau 3.1 : Impact des localisations sur l'évolution de la part de marché des transports collectifs (sur l'ensemble VP + TC)
(Source : Bonnel, Cabanne, 2000)

Les effets motorisation individuelle spatialisée, offre de transport et « autres facteurs » sont plus délicats à quantifier, car ils ne relèvent pas d'un simple calcul mais dépendent de la validité de la formulation de l'utilité retenue (réalisme des variables retenues, pertinence de la formulation, fiabilité des paramètres). Ces différents effets varient suivant la formulation. Toutefois, ils restent comparables en ordre de grandeur (tableau 3.2).

	Impact des différents facteurs entre les enquêtes ménages déplacements											
	1977 et 1985				1985 et 1995				1977 et 1995			
	mot	vp	tc	a	mot	vp	tc	a	mot	vp	tc	a
A1	-11,7%	1,1%	7,1%	21,6%	-1,1%	3,9%	2,5%	-8,7%	-12,7%	5,2%	10,1%	11,7%
P1	-12,4%	1,6%	11,3%	17,8%	-1,1%	4,0%	3,4%	-9,4%	-13,5%	6,5%	15,6%	7,2%
A2	-7,0%	1,1%	6,9%	15,7%	-1,0%	3,3%	2,5%	-8,3%	-8,1%	5,0%	9,8%	6,3%

Tableau 3.2 : Impact des effets simples autres que localisation sur l'évolution de la part de marché des transports collectifs (sur l'ensemble VP + TC)
(Source : Bonnel, Cabanne, 2000)

L'effet motorisation peut être estimé à environ -7% à -12% pour la première période (1^o colonne du tableau). Pour la deuxième période (1985-1995), la motorisation individuelle spatialisée a peu évolué, ce qui a entraîné un effet très faible (colonne « mot » entre 1985 et 1995).

Pour l'offre de transport, l'impact peut être évalué entre 15% et 20% sur l'ensemble de la période (colonne « vp » et « tc » entre 1977 et 1995). L'effet de l'offre de transport peut être décomposé entre les deux modes. La diminution importante des coûts généralisés en transports collectifs due à l'amélioration importante de l'offre de transports collectifs sur la période (construction de quatre lignes de métro, restructuration du réseau de bus, forte croissance des véhicules*kilomètres...) explique une croissance de 10% à 15% de la part de marché des transports collectifs. L'accroissement de l'usage de la voiture combiné à un allongement des distances a eu un impact assez modéré sur la croissance des coûts généralisés de la voiture du fait de l'amélioration du réseau et des mesures de régulation. L'incidence sur l'évolution de la part modale des transports collectifs est donc limitée autour de $+5\%$ à 6% entre 1977 et 1995.

Les « autres facteurs » ont eu un effet important sur l'ensemble de la période (colonne « a »). Les valeurs des différents effets sont à retenir en ordre de grandeur, mais il apparaît que ces « autres facteurs » jouent un rôle du même ordre de grandeur que les différents facteurs explicatifs : de 6% à 12% de hausse sur les 20 ans avec un renversement de tendance. L'évolution positive de la part modale des transports en commun se révèle être plus forte que prévue sur la première période, alors qu'en seconde période d'autres facteurs semblent avoir joué un rôle de frein. A cela, plusieurs explications peuvent être avancées. L'effet « marketing » de l'ouverture du métro en 1978 aurait renforcé l'attractivité des transports en commun au-delà de l'amélioration de l'offre. L'extension du réseau après 1985 n'a pas eu le même effet. Entre 1985 et 1995, les politiques de développement de l'offre de stationnement dans le cœur de l'agglomération ont pu contribuer à infléchir la part modale des transports en commun, de même que la complexification des chaînes de déplacement par rapport à de simples allers-retours depuis le domicile. Il existe peut-être aussi une évolution des comportements liée à la perception des modes de transport au bénéfice de la voiture, que la stabilité des coefficients des fonctions de coûts généralisés sur l'ensemble de la période n'est pas à même de prendre en compte. Enfin, il reste probablement encore d'autres facteurs parmi l'ensemble des déterminants du choix du mode dont l'influence a pu évoluer au cours des deux décennies.

3.1.5 Conclusion : précautions méthodologiques

Sur le plan méthodologique, il nous semble utile de revenir sur l'interprétation que l'on peut donner de ces résultats. Tout d'abord, nous nous gardons bien de les interpréter comme étant le signe d'une relation de causalité univoque. Le terme effet est en fait impropre, car il réfère à une relation de causalité que nous ne souhaitons pas aborder ici, car notre travail ne nous permet pas de traiter pleinement cette question. En revanche, nous mettons en évidence qu'une certaine évolution des localisation peut être reliée à une certaine évolution de la part de marché des transports collectifs et que l'on peut quantifier cette relation. La même constatation peut être faite pour la motorisation et les autres facteurs.

La seconde remarque a trait à l'indépendance entre les facteurs. Le résultat que nous produisons ne signifie pas que nos cinq facteurs sont indépendants entre eux. L'évidence démontrerait facilement le contraire. Par contre, nous avons réussi à trouver une formulation qui permet de réduire (si ce n'est supprimer) la dépendance entre l'évolution des éléments que nous avons pris en compte.

Ceci nous amène à la dernière remarque. Ce que nous nommons effet motorisation, par exemple, est en fait un abus de langage. L'évolution de la motorisation peut avoir des incidences sur les choix de localisation qui sont alors inscrits dans notre étude comme effet localisation et non effet motorisation, tout comme les choix de localisations peuvent avoir une incidence sur la motorisation. Ce que nous nommons « effet motorisation » ne correspond donc pas forcément à la totalité des « effets de la motorisation » sur l'usage des modes. Il traduit plutôt « l'effet » de l'évolution de la motorisation à structure spatiale des déplacements constante. Nous établissons ainsi une quantification de la relation entre l'évolution de la matrice origines-destinations des motorisations et l'évolution de la part de marché des transports collectifs. Tout comme nous quantifions la relation entre l'évolution de la matrice origines-destinations des déplacements et l'évolution de la part de marché des transports collectifs ; de même pour chacun des facteurs. Et surtout l'évolution combinée de ces cinq matrices signifie une évolution de la part de marché des transports collectifs égale au produit des cinq relations, avec des relations croisées négligeables. On peut exprimer ce résultat par la formulation mathématique suivante :

$$\Delta P_{TC} = E(\Delta \text{matrice localisation}) * E(\Delta \text{matrice motorisation}) * E(\Delta \text{matrice offre transports collectifs}) * E(\Delta \text{matrice offre voiture}) * E(\Delta \text{matrice autres facteurs})$$

avec ΔP_{TC} : variation de la part de marché des transports collectifs entre deux dates ;

E : effet de l'évolution de chacun des facteurs.

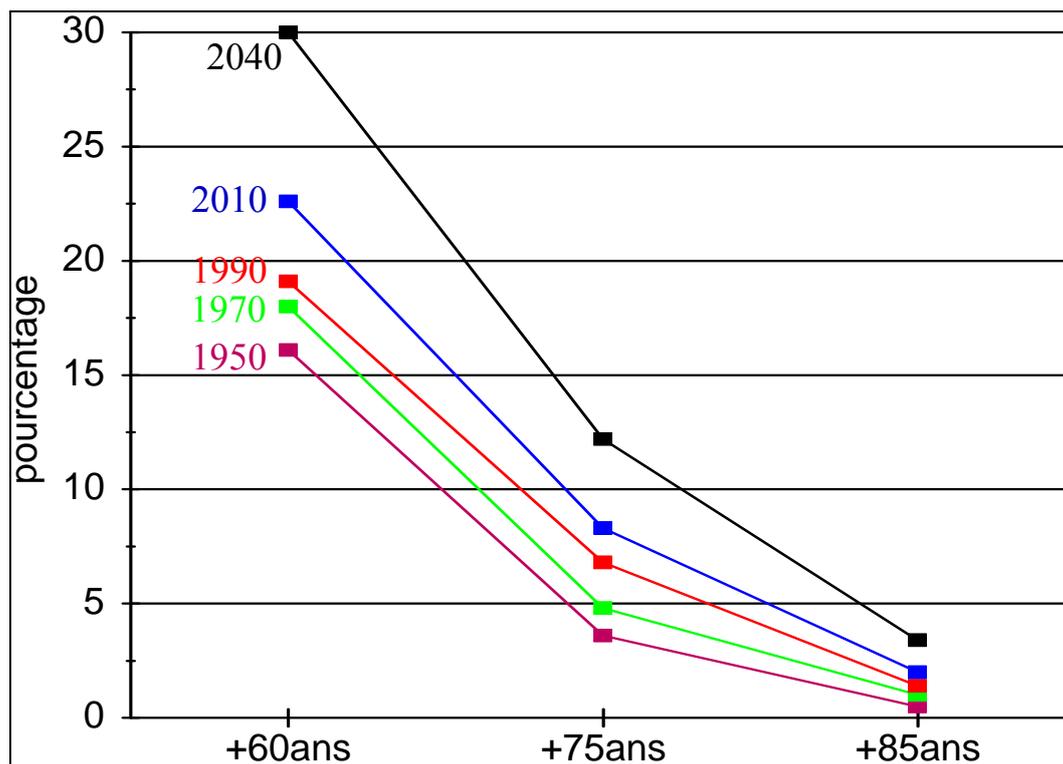
C'est le résultat de l'évolution de ces matrices que nous dénommons par facilité effet localisation, motorisation, offre de transport et autres facteurs, même s'il ne désigne en fait qu'une relation entre une certaine formulation de chacun des facteurs et la part de marché des transports collectifs.

3.2. Vieillesse de la population : perspective d'évolution de la mobilité automobile des personnes âgées

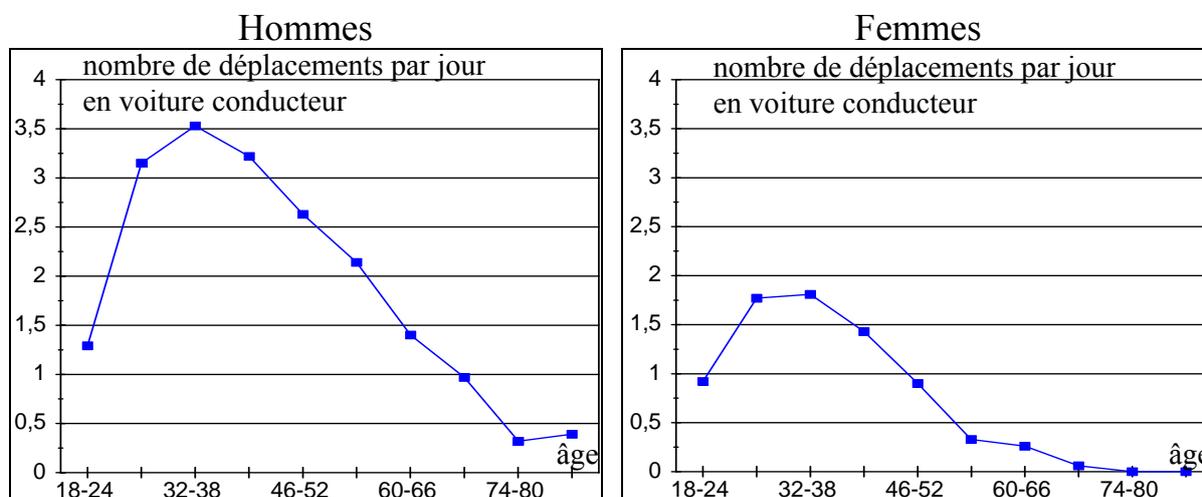
Cette présentation est issue entièrement des travaux de la thèse de Pochet (1995). La population française, comme celle de la plupart des pays développés, connaît un vieillissement progressif qui doit se poursuivre au cours du siècle selon toutes les prévisions démographiques. Le graphique 3.12 illustre cette situation.

3.2.1. Vieillesse de la population... solution des problèmes de congestion ?...

Le nombre de déplacements réalisés quotidiennement, ainsi que l'usage des modes varient fortement selon l'âge de la personne, le vieillissement de la population devrait avoir des incidences fortes sur ces indicateurs. L'évolution du nombre de déplacements quotidiens en voiture selon l'âge de la personne pour l'agglomération de Grenoble (graphique 3.13) semble confirmer ce constat. Le nombre de déplacements réalisés en voiture décroît très fortement avec l'âge. Ainsi le vieillissement de la population française devrait fortement réduire le nombre de déplacements automobiles dans le futur si l'on en croit ces données. De plus, les femmes vivent plus vieilles que les hommes et leur mobilité est encore plus faible que celle des hommes ce qui ne fait que renforcer notre conclusion.



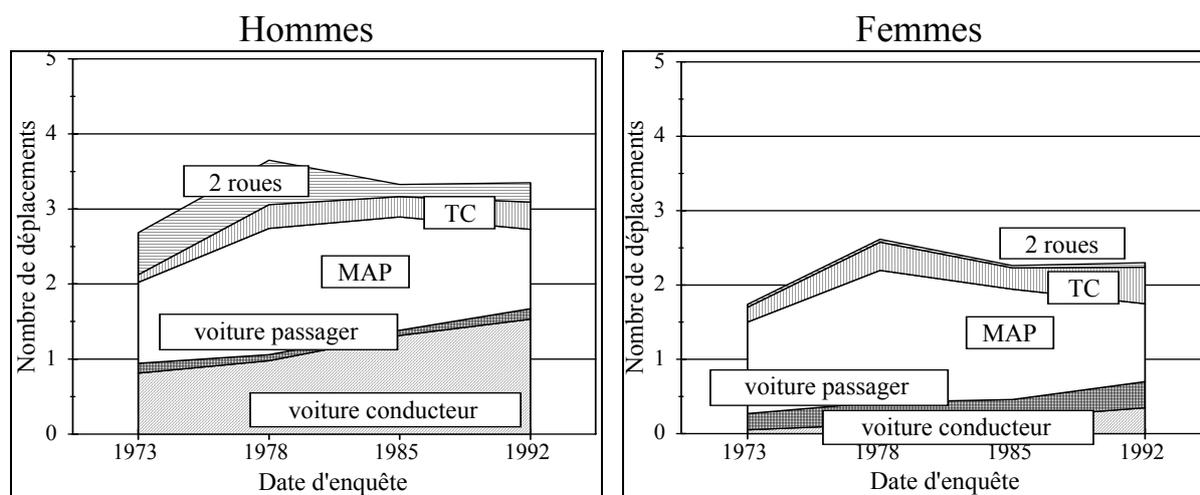
Graphique 3.12 : Poids des personnes âgées dans la population française entre 1950 et 2040
(Sources : Pochet, 1995 ; données INSEE)



Graphique 3.13 : Nombre de déplacements quotidiens en voiture comme conducteur selon l'âge
(Sources : Pochet, 1995 ; données enquête ménages déplacements de Grenoble 1992 (SMTC, 1992))

3.2.2. ... ou amplificateur de la congestion ?... dû à l'usage de la voiture par les hommes ?...

L'analyse n'est peut être pas aussi simple. Le graphique 3.13 représente en fait le comportement des grenoblois en 1992. Rien ne permet d'affirmer que les personnes âgées se comporteront de manière identique dans 10 ou 20 ans. Une analyse en coupe instantanée ne peut renseigner sur des tendances d'évolution. Pour valider notre hypothèse, il nous faut donc analyser l'évolution des comportements au cours des dernières années. Le graphique 3.14 présente les résultats des quatre dernières enquêtes ménages déplacements de Grenoble.



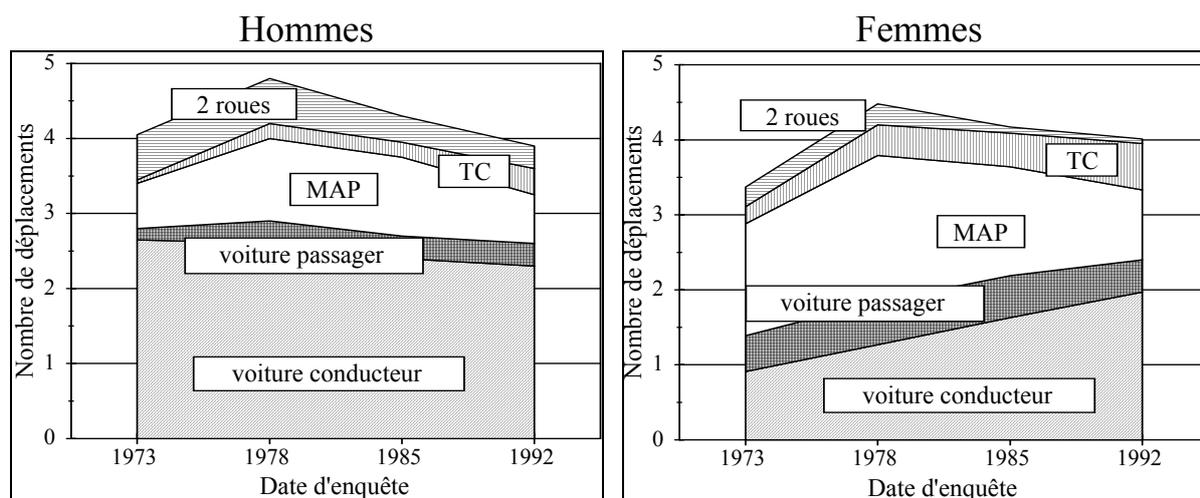
Graphique 3.14 : Evolution du nombre de déplacements quotidiens des personnes de 60 ans et plus (Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1973, 1978, 1985, 1992)

La lecture du graphique 3.14 ne peut qu'invalider notre hypothèse. Au contraire, les personnes âgées vont se déplacer davantage en voiture si l'on prolonge les tendances passées. La faible croissance de l'usage de la voiture par les femmes en tant que conductrice pourrait toutefois tempérer un peu cette conclusion dans la mesure où elles sont plus nombreuses que les hommes dans la population âgée de plus de 60 ans.

Cet exemple illustre tous les dangers d'une analyse en coupe instantanée lorsqu'il s'agit de faire de la prospective. Il n'est pas possible de faire l'économie d'une analyse temporelle. Celle-ci reste toutefois encore incomplète. Les personnes âgées de demain ne sont pas toutes âgées aujourd'hui. Ainsi une prospective de la mobilité des personnes qui auront 60 ans et plus dans l'avenir ne peut occulter l'évolution des comportements des personnes ayant aujourd'hui moins de 60 ans.

3.2.3. ... A moins que ce soit à cause des femmes ?...

L'évolution des comportements des personnes plus jeunes (graphique 3.15) nous conduit de nouveau vers une conclusion différente. Certes l'usage de la voiture devrait continuer à croître, mais il devrait être le fait des femmes principalement et non des hommes. En effet les hommes de 18 à 59 ans utilisent un peu moins la voiture en 1992 que lors des enquêtes précédentes à l'inverse des femmes. Il semblerait même que la croissance chez les femmes pourrait encore se poursuivre si les différences entre les sexes dans l'usage des modes devaient continuer à s'estomper comme le graphique 3.15 nous invite à le penser. Il reste toutefois une inconnue dans cette analyse. Est-ce que le passage à la retraite (ou le passage à la soixantaine) entraîne un abandon de l'usage de la voiture ?



Graphique 3.15 : Evolution du nombre de déplacements quotidiens des personnes de 18 à 59 ans (Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1973, 1978, 1985, 1992)

3.2.4. ... Ou les pièges de l'analyse de la mobilité

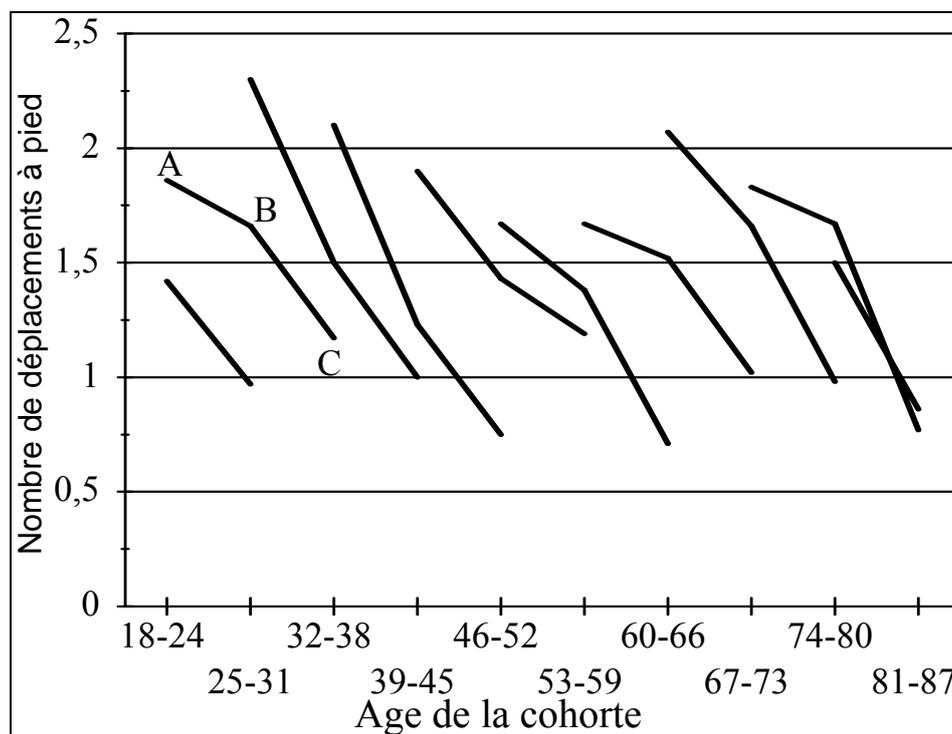
Sans répondre pour le moment sur le plan prospectif, les trois graphiques (graphiques 3.13 à 3.15) que nous avons montrés induisent des réponses radicalement différentes illustrant les risques de conclusions hâtives s'appuyant sur des corrélations et des relations de causalité non établies. Dans un premier temps (graphique 3.13) nous avons supposé une relation de causalité forte entre l'âge et l'usage de la voiture. Le vieillissement entraîne un abandon progressif de la voiture. En fait, le graphique peut tout aussi bien être interprété comme un effet de génération. Les générations anciennes n'ont pas eu accès à l'automobile de manière massive durant leur jeunesse, tout particulièrement pour les femmes, il est donc naturel qu'elles ne l'utilisent pas. Inversement les jeunes générations y ont un accès très large, elles pourraient donc continuer à l'utiliser de manière massive sans forte distinction de sexe. C'est le sens de la lecture du troisième graphique (graphique 3.15) qui s'appuie sur la constatation de la croissance dans le temps de cet usage chez les personnes de 60 ans et plus (graphique 3.14). Il est probable que l'on soit face tout à la fois à des effets d'âge (le passage à la retraite, puis le vieillissement physique réduisant le nombre déplacements quotidiens) et de génération avec l'évolution de la diffusion de l'automobile depuis la seconde guerre mondiale.

3.2.5. Effet d'âge ou de génération ?

Il existe même un troisième type d'effet. C'est l'effet de période. Il touche tous les individus de manière comparable indépendamment de l'âge. Le graphique 3.16 illustre un tel effet. Le nombre de déplacements en marche à pied diminue de manière analogue chez toutes les femmes indépendamment de leur âge et de la cohorte à laquelle elles appartiennent (pour une formalisation de ces approches âge-cohorte : Madre, Lambert, 1989 ; Madre, Armoogum, 1996).

Pour séparer les trois effets d'âge, de génération et de période, il faut développer une analyse âge-cohorte. C'est-à-dire qu'il s'agit de suivre une cohorte dans son vieillissement. Une cohorte est définie par les individus appartenant à la même classe d'âge à un instant donné. Pour faciliter la lecture des graphiques issues de ce type d'analyse, nous avons marqué la position d'une cohorte par les 3 lettres A, B, C sur le graphique 3.16. Le point A correspond à la classe d'âge 18-24 ans lors de l'enquête de 1978. Le point B correspond à cette population lors de l'enquête de 1985. La cohorte a donc vieilli et a maintenant entre 25 et 31 ans. Enfin, le point C correspond toujours à la même

population lors de l'enquête de 1992 qui a entre 32 et 38 ans. Il est ainsi possible de suivre une même cohorte dans son vieillissement (de A, vers B puis C) pour chacune des cohortes.



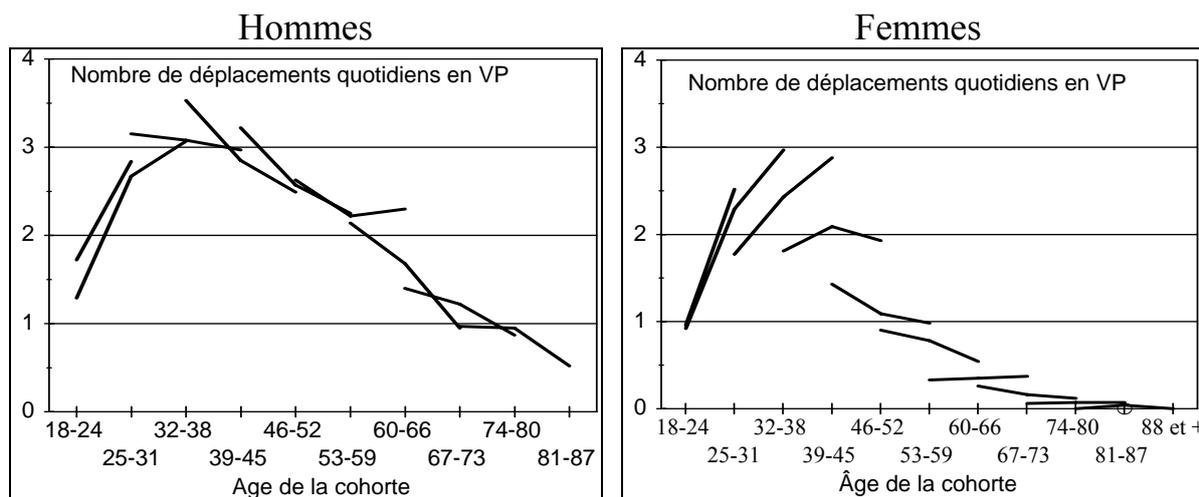
Graphique 3.16 : Evolution du nombre de déplacements à pied des femmes par cohorte
(Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1978, 1985, 1992)

La lecture de tels graphiques permet de séparer les trois types d'effet (des exemples sont fournis dans (Pochet, 1995) qui décrivent chacun des effets simples et combinés). Nous l'appliquons au nombre de déplacements automobiles quotidiens des hommes et des femmes (graphique 3.17). Pour les hommes, nous observons principalement un effet d'âge. En effet, globalement les cohortes se superposent. Il apparaît toutefois un faible effet de génération chez les plus jeunes (jusqu'à 45 ans) qui se traduit par une diminution du nombre de déplacements automobiles. En revanche, pour les personnes plus âgées et notamment pour la classe d'âge 60-66 ans, on observe un effet de génération conduisant à une faible augmentation du nombre de déplacements automobiles.

En revanche, pour les femmes, on assiste à un très fort effet de génération chez les plus jeunes (écarts verticaux entre les cohortes). Toutefois, cet effet se ralentit progressivement pour s'estomper chez les plus âgées. Plusieurs interprétations peuvent être données de ce résultat. La diffusion de l'automobile chez les femmes est assez récente. Elle correspond en fait au fort développement de la bi-motorisation depuis une vingtaine d'années, la première voiture étant celle de l'homme et la seconde celle de la femme. Elle correspond également à l'entrée en force des femmes sur le marché du travail qui se déplacent ainsi davantage. Cette interprétation conduit à supposer que l'effet de génération va se prolonger dans le temps chez les plus âgées et se traduire par une forte croissance de l'usage de la voiture pour ces dernières. On pourrait par contre estimer qu'au delà de l'effet de génération chez les plus jeunes, passées la soixantaine, les femmes ont moins besoin d'utiliser leur voiture et alors, on continuerait dans l'avenir à avoir uniquement un effet d'âge qui se traduirait par une relative stabilité de l'usage de la voiture des femmes de plus 60 ans.

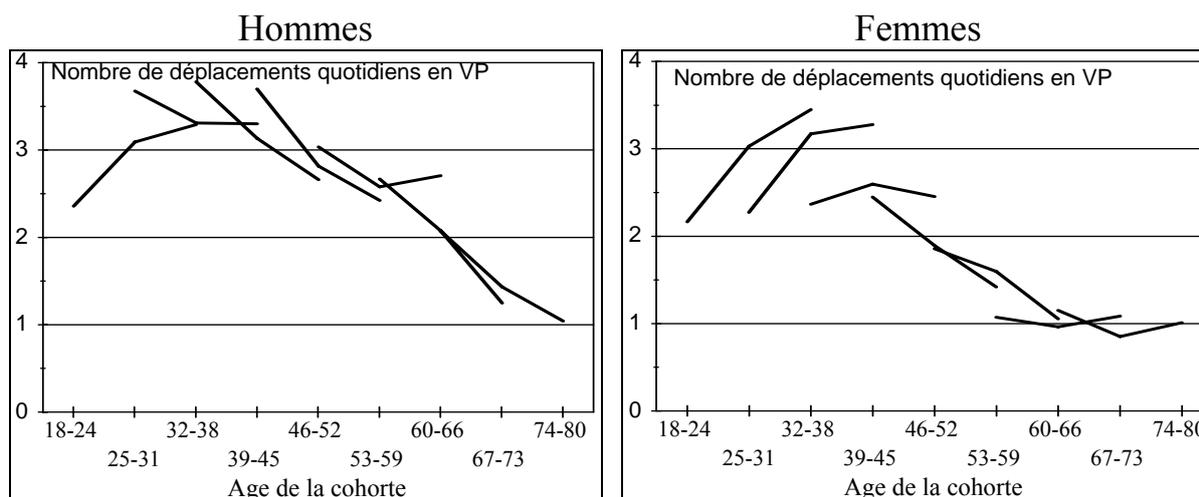
Au-delà du constat, il reste maintenant à essayer d'expliquer l'effet générationnel chez les femmes dans une perspective prévisionniste. Est-il dû à la croissance de la motorisation (ou plus

particulièrement de la bi-motorisation) ou est-il dû à un effet de comportement des femmes motorisées qui se déplacent davantage en voiture aujourd'hui qu'autrefois ? Il faut pour cela analyser l'évolution du nombre de déplacements des femmes ayant accès à la voiture. L'enquête ménages déplacements ne précise pas qui dans le ménage utilise régulièrement chacune des voitures. Une règle d'accès à la voiture a donc été définie : s'il n'y a qu'une voiture, elle est attribuée à l'homme s'il a le permis sinon à la femme si elle a le permis ; s'il y a deux voitures, les deux membres du couple ont accès à la voiture s'ils disposent du permis. On définit alors le conducteur potentiel comme étant celui qui a accès à une voiture selon cette définition. Le graphique 3.17 illustre l'évolution de l'usage de la voiture des conducteurs potentiels.



Graphique 3.17 : Evolution du nombre de déplacements automobiles par cohorte
(Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1978, 1985, 1992)

Pour les hommes, on retrouve principalement l'effet d'âge, mais avec une diminution plus lente du nombre de déplacements avec l'âge pour les personnes âgées. Cette similitude des courbes tient à l'accès massif à la voiture des hommes.



Graphique 3.18 : Evolution du nombre de déplacements automobiles par cohorte
des conducteurs potentiels
(Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1978, 1985, 1992)

En revanche, pour les femmes la courbe change légèrement de physionomie. L'effet de génération subsiste fortement pour les plus jeunes. Les conductrices potentielles ont donc accru leur utilisation de la voiture. Parallèlement le nombre de jeunes femmes ayant accès à la voiture a fortement augmenté. On se trouve donc face à un double effet : de génération qui se porte sur la diffusion de l'automobile d'une part et d'autre part sur l'utilisation plus intensive de ce mode pour celles qui y ont accès. En revanche, passé la cinquantaine, l'effet générationnel n'est presque plus présent. Les femmes plus âgées ayant accès à la voiture n'ont donc pas modifié leur comportement face à l'automobile.

3.2.6. Séparation et quantification des effets

Pour tenter une prospective, il devient nécessaire de séparer les différents facteurs influant sur l'évolution de l'usage de la voiture. Nous avons identifié un effet d'âge sur l'utilisation de la voiture. Il induira des modifications d'usage de la voiture dans l'ensemble de la population en raison de l'évolution du poids de chacun des groupes d'âge. Nous qualifions cette évolution d'effet démographie.

L'effet générationnel se traduit par un renouvellement des générations par d'autres ayant un autre usage de la voiture. Cette évolution de l'utilisation de la voiture est en fait due tout à la fois à une évolution de la motorisation et à un autre usage de la voiture chez ceux qui ont accès à ce mode de transport. Nous définissons ces deux termes comme étant respectivement un effet motorisation et un effet comportement.

La méthodologie de décomposition des effets est identique à celle développée précédemment (encadré 3.1, p. **Erreur ! Signet non défini.**). Nous pouvons donc reprendre les réserves d'interprétation que nous avons exposées pour la décomposition des effets de localisation et de motorisation.

Le produit des effets doubles et triple est généralement assez proche de 1, attestant de la capacité de la méthode à séparer les trois effets simples. Sur l'ensemble des personnes de 18 ans et plus (tableau 3.3), l'effet démographie est assez limité. Il conduit à une faible diminution de l'usage de la voiture. Pour les hommes, les autres effets sont également faibles. On observe une croissance limitée de la motorisation entre 1978 et 1985, probablement liée à la croissance de la motorisation plus soutenue chez les plus âgés. En revanche, l'effet comportement conduit à une diminution de l'usage de la voiture durant la première période. Chez les femmes, le résultat est bien différent. La croissance de la motorisation est forte durant les deux périodes. Elle se double également d'un effet comportement : les femmes motorisées se déplacent davantage en voiture qu'auparavant.

	hommes		femmes	
	1985/1978	1992/1985	1985/1978	1992/1985
Croissance globale	0,98	0,97	1,29	1,21
Produit des 4 effets doubles et triple	1,01	1,00	1,01	0,99
Décomposition				
effet « démographie »	1,00	0,97	0,96	0,97
effet « motorisation »	1,05	1,00	1,20	1,15
effet « comportement »	0,92	1,00	1,11	1,09

Tableau 3.3 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 18 ans et plus

(Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1978, 1985, 1992)

La décomposition par tranche d'âge (tableaux 3.4 à 3.6) permet d'affiner ces résultats. L'effet démographie devient sensible chez les plus âgés dont le nombre s'accroît au cours des deux périodes. Par contre la relative stabilité du nombre de personnes dans les deux autres groupes d'âge se traduit par un effet démographie nul sur l'ensemble des deux périodes.

Pour les hommes, l'effet motorisation devient sensible de nouveau pour les plus âgés, surtout au cours de la première période, le renouvellement des générations âgées imposant progressivement la motorisation massive observée chez les plus jeunes. Les effets comportement se traduisent par une croissance faible pour les plus âgés et par une diminution chez les plus jeunes (est-ce un effet de l'arrivée plus tardive sur le marché du travail des jeunes et du développement du chômage ?)

	hommes		femmes	
	1985/1978	1992/1985	1985/1978	1992/1985
Croissance globale	1,36	1,19	1,82	1,82
Produit des 4 effets doubles et triple	0,97	0,94	1,00	0,96
Décomposition				
effet « démographie »	1,06	1,02	1,06	1,03
effet « motorisation »	1,28	1,11	1,55	1,44
effet « comportement »	1,03	1,13	1,11	1,28

Tableau 3.4 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 60 ans et plus
(Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1978, 1985, 1992)

	hommes		femmes	
	1985/1978	1992/1985	1985/1978	1992/1985
Croissance globale	0,93	1,03	1,38	1,56
Produit des 4 effets doubles et triple	1,01	1,00	1,02	1,00
Décomposition				
effet « démographie »	0,98	1,01	0,93	1,06
effet « motorisation »	1,05	1,03	1,34	1,24
effet « comportement »	0,90	0,98	1,09	1,19

Tableau 3.5 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 39 à 59 ans
(Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1978, 1985, 1992)

	hommes		femmes	
	1985/1978	1992/1985	1985/1978	1992/1985
Croissance globale	0,97	0,93	1,19	1,04
Produit des 4 effets doubles et triple	1,00	1,00	0,94	0,98
Décomposition				
effet « démographie »	1,03	0,97	1,00	0,94
effet « motorisation »	1,03	0,97	1,12	1,09
effet « comportement »	0,92	0,99	1,11	1,03

Tableau 3.6 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 18 à 38 ans
(Sources : Pochet, 1995 ; données enquêtes ménages déplacements de Grenoble 1978, 1985, 1992)

Chez les femmes, on observe une conjonction des effets motorisation et comportement conduisant à une forte croissance de l'usage de la voiture. L'effet motorisation est prépondérant, tout particulièrement pour les plus âgées. Toutefois, pour ces dernières, si les taux de croissance sont très élevés, les niveaux d'usage restent très faibles (graphique 3.18). Inversement, pour les plus jeunes, le niveau déjà élevé de l'usage en 1978, se traduit par une augmentation sensible en niveau, même si elle est plus modérée en taux de croissance. En fait, pour les plus jeunes, les deux effets semblent s'estomper, probablement parce que le « rattrapage » (par rapport aux hommes de même âge) est pratiquement terminé, ce qui n'est pas le cas des deux autres groupes pour lesquels les taux de croissance sont similaires durant les deux périodes.

4. CONCLUSION : LES PIEGES DE L'ANALYSE DE LA MOBILITE

Ce chapitre nous a permis de présenter quelques-unes des caractéristiques de la mobilité urbaine et quelques-uns des principaux facteurs influant sur son évolution. Si les résultats décrits ne sont pas sans intérêt, cette présentation a surtout un autre mérite. C'est celui d'attirer l'attention sur les pièges de l'analyse de la mobilité. Nous en citerons quatre :

- le problème de la causalité ;
- la dépendance entre les facteurs dits explicatifs d'un phénomène ;
- la prise en compte du temps ;
- la définition des indicateurs utilisés pour mesurer des phénomènes.

4.1. *Corrélation ou causalité*

Nous avons eu l'occasion d'entrevoir le problème à de multiples reprises. Mais l'exemple le plus clair est probablement celui présenté à travers les graphiques 3.13 à 3.15. Chacun de ces graphiques nous a permis de mettre en évidence une corrélation forte entre des grandeurs. Nous avons volontairement transformé ces corrélations en relation de causalité. L'application de ces relations nous a conduit vers des conclusions radicalement différentes pour le même phénomène. Pourtant, ce ne sont pas les données qui sont en cause. Elles sont issues d'une même production de données présentant certaines garanties de représentativité. Les corrélations ne sont pas non plus en cause. C'est donc le passage à la causalité qui doit être remis en cause.

Un autre exemple (graphique 3.7, p. 79) pourrait conduire à l'issue d'une analyse sommaire à diagnostiquer à une absence de causalité. La part de marché des transports collectifs sur l'agglomération est restée quasiment stable entre 1977 et 1995. Pourtant, l'offre s'est fortement accrue. Il semble donc qu'il n'y ait pas de relation entre l'offre et l'usage de transports collectifs. Dès lors que l'on décompose spatialement les données, on comprend tout de suite que l'interprétation est erronée. La méthodologie développée dans ce chapitre permet même de calculer l'intensité de la relation entre ces deux grandeurs.

L'observation d'une corrélation entre des grandeurs n'a pas de signification en elle-même. A partir de données statistiques, il est aisé de bâtir de nombreuses corrélations. Mais c'est l'interprétation que l'on fera de ces corrélations qui leur donnera éventuellement sens ou valeur de causalité. Cette interprétation doit s'appuyer sur des données complémentaires (par exemple les données présentées dans les graphiques 3.8 et 3.9, p. 80) ou sur des analyses faites par ailleurs ou encore sur un jeu d'hypothèses qui permet de la valider.

Le passage de la corrélation à la relation de causalité est toujours assez délicat, car nous sommes généralement en présence de phénomènes complexes mettant en jeu de multiples dimensions. L'interprétation causale est pourtant nécessaire pour faire de la prévision. Le modèle est en effet un système d'interprétation causale dans lequel la modification des données d'entrée produit une évolution des données de sortie.

4.2. Dépendance entre les facteurs « explicatifs »

La mobilité est un objet complexe. Nous avons déjà eu l'occasion de l'illustrer ne serait-ce que parce qu'elle est située à l'interaction entre trois sous-systèmes. De ce fait, de multiples facteurs agissent sur cet objet. Cette multiplicité nous conduit à une approche multidimensionnelle du problème. Tous ces facteurs ne peuvent être analysés simultanément en raison de la forte dépendance existant entre nombre d'entre eux.

L'analyse de la motorisation nous a permis d'illustrer le problème de la corrélation avec les localisations. La plupart (si ce n'est toutes) des analyses de la mobilité sont confrontées à ce problème. Par exemple, la prise en compte de variables socio-économiques amène à utiliser des variables fortement corrélées : la profession ou l'activité est fortement corrélée au sexe, à l'âge, aux revenus, à l'accès à la voiture....

Pourtant, la construction d'un modèle explicatif nécessite de séparer la contribution de chacun des facteurs. Il faut pour cela que les variables soient indépendantes. Il devient alors nécessaire de trouver une formulation des facteurs conduisant à cette indépendance.

4.3. Prise en compte du temps

Toute prévision fait par nature appel au temps. Ce temps doit donc être pris en compte dans l'analyse. L'analyse de données en coupe instantanée nécessite ainsi une certaine prudence. Ces données correspondent à une « réalité » à un instant donné. La relation que l'on mettra en évidence à partir de ces données n'est donc qu'une relation traduisant une « réalité » de cet instant. Il n'est pas possible de lui donner une valeur d'invariant dans le temps. L'exemple de l'évolution de la mobilité automobile avec l'âge illustre bien cette constatation. C'est une « réalité » de la mobilité urbaine de Grenoble à la date de l'enquête, mais ce n'est pas un invariant atemporel ainsi que l'attestent les données des enquêtes réalisées les années précédentes.

L'analyse de données chronologiques permet d'introduire l'effet du temps dans l'analyse. C'est ce que nous avons fait en essayant de séparer un effet d'âge, un effet de génération et un effet de période. Toutefois, de nouveau, cette analyse permet d'établir des relations qui traduisent une « réalité » sur une période donnée. Les invariants établis sur cette période ne sont peut être plus valables sur une plus longue période. C'est ce que nous avons pu observer en analysant l'évolution de la mobilité des hommes et des femmes à Grenoble. La croissance observée chez les hommes âgés ne se prolongera probablement pas, tandis que la stabilité chez les femmes âgées est trompeuse.

C'est de nouveau l'interprétation que l'on donnera de ces évolutions, « réalité » d'une période donnée, qui permettra de faire de la prévision.

4.4. Quelle mesure ?

La mesure de la motorisation nous a permis d'aborder le problème. Nous avons présenté trois mesures de la motorisation (graphique 3.10, p. 81) qui conduisent à des éclairages différents du même objet. Tout particulièrement l'évolution temporelle de la motorisation est assez contrastée selon la définition retenue. Pourtant, ces mesures sont l'illustration d'un seul et même objet, la motorisation.

Cet exemple illustre une nouvelle fois la multidimensionnalité des objets d'étude utilisés dans les analyses de la mobilité. Les mesures que l'on utilise sont une réduction de cette complexité. Toutefois, lorsqu'il s'agit d'interpréter les évolutions de ces mesures, ou des corrélations, la multidimensionnalité doit être réintroduite. Dans le cas présent, les différences observées dans les différentes mesures de la motorisation ne découlent pas seulement d'une évolution du nombre total de voitures à Lyon, mais elles tiennent aussi à l'évolution des ménages sur la période d'étude. Ceux-ci

ont vu leur taille diminuer fortement, et dans le même temps la proportion de personnes de 18 ans et plus s'est accrue. L'interprétation de l'évolution de la motorisation doit donc intégrer également ces dimensions.

Ce chapitre illustre la complexité de l'analyse de la mobilité urbaine et la multidimensionnalité de cet objet d'étude. Nous pouvons en tirer deux enseignements qui renforcent s'il en était encore besoin les propos développés dans le chapitre 2. Le premier a trait à l'importance de l'analyse de l'objet d'étude en vue de la construction du modèle théorique. Une analyse sommaire peut facilement conduire à des constructions non pertinentes (l'analyse sur les conséquences du vieillissement de la population est explicite à cet égard). Le second concerne la production des données. Celle-ci nécessite une réflexion préalable sur l'objet d'étude. Sa multidimensionnalité impose une réduction de l'objet qui doit alors être approché sous plusieurs angles différents. L'enjeu de la production des données est de tenter grâce à ses visions partielles de proposer une vision globale de l'objet. Pour y parvenir, cette production doit répondre à une séquence méthodologique.

Chapitre 4 - Production de données

Comme pour le chapitre précédent, il ne s'agit pas de dresser le panorama des méthodes d'enquêtes. De nombreux ouvrages existent sur la question (parmi lesquels : Ampt et al., 1985 ; Ghiglione, Matalon, 1985 ; Groves et al., 1988 ; Richardson et al., 1995 ce dernier étant le plus complet en ce qui concerne les enquêtes déplacements ; Ardilly, 1994 pour l'échantillonnage). Nous présentons plutôt les données existantes en milieu urbain en France (la portée du chapitre sur la production des données déborde largement des questions de déplacements urbains, même si nous nous y restreignons en ce qui concerne les illustrations et la présentation des données généralement disponibles), tout en les resituant par rapport aux besoins de modélisation et en soulignant leurs limites dans cette perspective.

Nous situons tout d'abord les principaux objectifs de la production de données sur les déplacements en milieu urbain (section 1). Nous abordons ensuite quelques éléments de méthodologie de la production de données (section 2). Comme pour la prévision, nous aurons l'occasion d'insister sur la nécessité d'une définition préalable des objectifs de la production des données et des utilisations prévues des données produites. La troisième section est consacrée à la présentation des principales données disponibles en France et tout particulièrement des enquêtes ménages déplacements (section 3). Nous soulignons ensuite quelles sont, selon nous, les principales limites de ce dispositif français pour répondre aux besoins de modélisation (section 4). Ces enquêtes répondent à un besoin de connaissance sur les comportements passés (ou encore comportements révélés, d'où le nom d'enquête de préférences révélées), mais ne sont pas toujours suffisantes pour appréhender les réactions possibles à des situations nouvelles ou pour quantifier certains attributs. Les enquêtes de préférences déclarées, issues des méthodologies utilisées en marketing visent à répondre à cet objectif (section 5). Enfin, nous donnons un exemple de construction de séquence méthodologique relatif à la production des données, pour insister plus particulièrement sur la construction du cadre conceptuel (section 6).

1. POURQUOI A-T-ON BESOIN DE DONNEES ?

Si on se limite aux principaux objectifs affichés par les concepteurs d'enquêtes périodiques sur les déplacements urbains, il est possible de les classer en deux groupes (Bonnell, Le Nir, 1998 ; Daly, (1982) présente une segmentation comparable dans le cadre des analyses portant sur les comportements de mobilité). Dans le premier, nous trouvons l'alimentation en données dans une perspective de planification des déplacements. Le second couvre plutôt l'analyse et la recherche des facteurs explicatifs des comportements de mobilité. On peut classer aisément dans le premier groupe des enquêtes comme l'*enquête régionale origine-destination canadienne* (Montréal et Toronto notamment (Chapleau, 1997 ; Dorland et al., 1997)), l'enquête nationale des Pays Bas (Benlarech et al., 1995) ; alors que dans le second, nous trouvons l'*enquête Transports et Communications* française (Madre, 1997 ; INSEE, INRETS, 1997), la « *Nation-wide Personal Transportation Survey – NPTS* » des Etats-Unis (Benlarech et al., 1995), le *microrecensement* suisse (Junod, 1997), l'enquête "*KONTIV*" en Allemagne (ces dernières enquêtes ne sont pas toutes exclusivement urbaines et couvrent parfois tout le territoire national). Cette segmentation est nécessairement schématique, d'autant plus que certaines enquêtes, comme les enquêtes ménages déplacements françaises, se placent dans les deux groupes. De plus, certaines enquêtes du second groupe servent également à alimenter des modèles de prévision. Avant d'examiner les conséquences de ces objectifs différents sur la production de données, nous précisons ces deux objectifs principaux.

Le but principal de la planification des déplacements est d'assurer les différentes fonctions de l'espace considéré dans des conditions « optimales » sur les plans économique et social. Nous restons

volontairement flou sur ce que peut recouvrir le terme « optimal », les définitions variant fortement selon les objectifs des politiques de déplacement. La planification concrétise le double souci d'une anticipation et d'une réflexion pour l'action. Cette dernière porte sur de nombreux domaines :

- la conception de schémas de transports, que ce soit des schémas nationaux ou dans le cadre des PDU, des DVA.... L'approche qui y est développée n'est plus uniquement modale, mais au contraire de plus en plus inter ou plurimodale ;
- la programmation puis la production d'infrastructures, qu'elles soient routières, de transports collectifs ou d'autres modes ;
- la gestion des infrastructures routières ou des entreprises de transports collectifs et des autres modes ;
- la régulation ou l'orientation de la demande de déplacements tant à court terme (voire en temps réel) qu'à plus long terme ;
- la tarification des déplacements urbains (pour tous les modes) ;
- l'organisation et la gestion des interfaces ou pôles d'échanges entre modes de transport ;
- l'alimentation de la réflexion sur les interactions entre l'aménagement urbain et les déplacements....

Cette liste n'est évidemment pas exhaustive, mais regroupe des préoccupations fréquemment rencontrées. On peut noter par ailleurs que l'on observe un glissement des préoccupations relatives à la production d'infrastructures vers des mesures de régulation au sens large (5 derniers items de la liste ci-dessus) (Stopher, 1998, Axhausen, 1998). Ces réflexions sont menées selon différentes échelles spatiales : du quartier à l'agglomération, du carrefour au réseau routier, de la ligne de bus aux réseaux de transports collectifs.... De même, différentes échelles temporelles sont présentes, du court terme au long terme. Le principal outil disponible pour alimenter ces réflexions et quantifier les effets est, depuis la fin des années soixante, le modèle de prévision de la demande de déplacements (chapitre 1).

Le second groupe d'objectifs vise davantage l'analyse des comportements individuels de déplacements. Il s'agit de rechercher les facteurs explicatifs des comportements de mobilité et de son évolution qu'elle soit saisonnière ou qu'elle constitue une tendance lourde. Il faut donc être en mesure d'étudier l'influence d'un certain nombre de paramètres sur la mobilité. Dans ce domaine, on pourra s'intéresser à des questions comme l'analyse de la mobilité de groupes cibles (personnes âgées, jeunes...) ou de zones particulières comme le péri-urbain, des déterminants du choix modal, l'influence de la tarification.... Au-delà de ces finalités premières, il s'agit bien souvent d'une première étape dans la perspective de la construction d'un modèle de prévision de la demande de déplacements.

Dans le premier groupe, la principale donnée nécessaire est le déplacement repéré notamment par son origine et sa destination (Ampt et al., 1998). L'enquête doit servir à construire la matrice origines-destinations utilisée par le modèle. La taille de l'échantillon sera donc privilégiée dans ce type d'enquête dans le but d'augmenter la représentativité statistique des données. Par contre, la modélisation ne nécessite pas une grande diversité d'informations. A la limite, on peut se contenter des données nécessaires à l'alimentation du modèle et au contrôle de la représentativité des données. En revanche, dans le second groupe, on privilégiera les caractéristiques permettant d'expliquer le comportement de mobilité. On s'attachera donc à formuler de nouvelles hypothèses ou de nouveaux paradigmes relatifs aux comportements de mobilité. A la limite, une seule enquête peut suffire pour formuler des hypothèses, même s'il en faut davantage pour les tester. La richesse de l'information sera donc privilégiée par rapport aux effectifs enquêtés. De même, une attention toute particulière sera portée à la qualité des données produites plutôt qu'à la représentativité statistique des données. Dans certains cas, on utilisera même plutôt une représentativité structurelle (par quota) qu'une représentativité statistique, pour élargir la diversité des situations repérées.

Ces deux groupes d'objectifs ne sont donc pas antagonistes et peuvent se rejoindre. Mais, les quelques éléments mentionnés ci-dessus, tout comme l'état de l'art sur les méthodes d'enquêtes, nous enseignent qu'il n'est pas toujours souhaitable de les associer dans une enquête (les contraintes de

financement et d'organisation imposent toutefois assez souvent de tempérer ce constat). En effet, l'une privilégiera la taille de l'échantillon et la représentativité statistique, tandis que l'autre privilégiera l'étendue de l'information pour chaque individu statistique et la représentativité structurelle. Les contraintes de budget (mais aussi d'efficacité) ne permettent donc pas toujours de conjuguer efficacement ces deux impératifs.

2. METHODOLOGIE DE PRODUCTION DES DONNEES

Tout comme la modélisation, la production de données requiert le respect d'une séquence méthodologique. Nous examinons tout d'abord le cas de la production des données pour l'analyse (section 2.1), puis celui de la production de données pour la construction ou l'alimentation de modèles (section 2.2). Même si notre propos est centré sur la modélisation, l'examen de la méthodologie pour les deux groupes d'enquêtes nous semble pertinent. Tout d'abord parce que l'analyse de la mobilité précède fréquemment la phase de construction du modèle. Mais aussi parce que nous constaterons que les séquences méthodologiques de production des données sont similaires dans les deux cas, même si les méthodes d'enquête mises en œuvre peuvent être différentes. La section suivante nous permet d'aborder l'étape de définition de la méthode d'enquête (section 2.3). Enfin nous abordons la question de la précision statistique des données (section 2.4).

2.1. Méthodologie de la production des données pour l'analyse

Nous reprenons en la complétant la méthodologie (schéma 4.1) proposée par Bonnafous (Bonnafous, Gérardin, 1976).

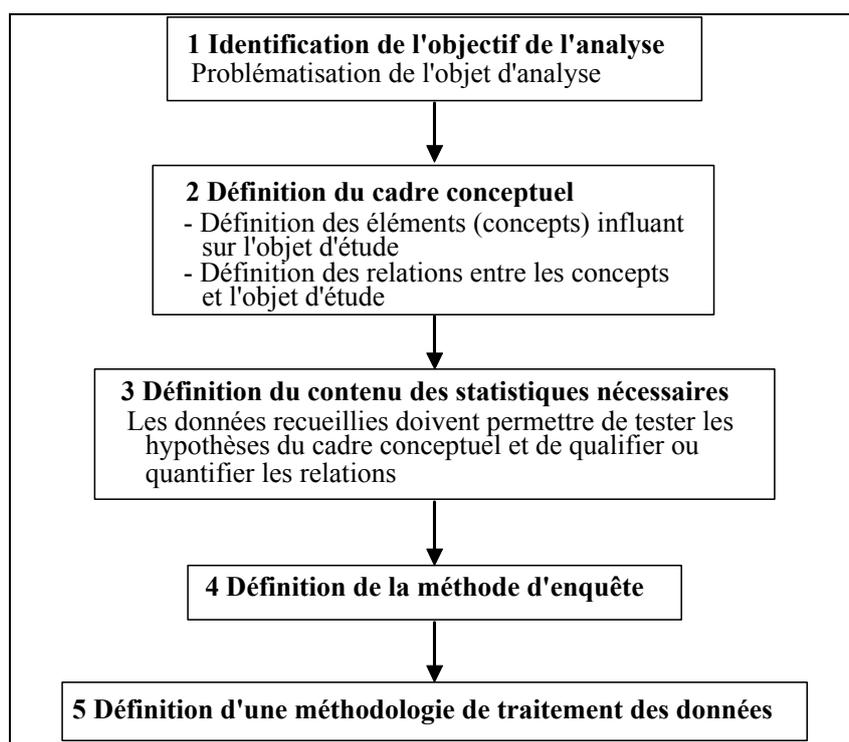


Schéma 4.1 : Méthodologie de la production des données pour l'analyse

La première étape débute par la définition de l'objectif de l'analyse. Les considérations que nous avons développées dans le chapitre 2 à propos de notre schéma de procédure de prévision (schéma 2.1, p. 34) peuvent être reprises ici. Il ne nous semble pas utile d'insister, dans la mesure où en l'absence d'objectif clair, il est évident qu'il ne sera pas possible de préciser le contenu des

statistiques nécessaires. Cette définition passe par la problématisation de l'objet d'analyse. Celle-ci permet de préciser la manière dont on aborde l'objet d'analyse. Confronté à des objets complexes, multidimensionnels, on n'est généralement pas capable de les appréhender dans toutes leurs dimensions. La problématisation éclaire les dimensions ou les angles d'approche qui seront retenus dans l'analyse.

La seconde étape consiste à développer le cadre conceptuel de l'analyse. Cette étape permet de préciser le schéma explicatif de l'objet d'étude. Il s'agit ici de préciser les éléments ou concepts qui influent sur l'objet d'étude et les relations entre ces éléments. Bien évidemment, il s'agit, à ce niveau de formuler des hypothèses sur ces éléments et ces relations. L'enquête doit servir à produire des données sur ces différents éléments, permettant d'une part de mieux les qualifier ou les quantifier et d'autre part de tester ces hypothèses. Elle permet également lorsque ces hypothèses sont validées, de qualifier ou de quantifier la nature des relations entre les différents éléments du cadre conceptuel et l'objet d'étude. Cette réflexion doit bien évidemment être préalable à la réalisation de l'enquête. Dans le cas contraire, l'enquête risquerait fort de ne pas pouvoir répondre à l'ensemble des hypothèses que l'on pourrait vouloir tester.

La troisième étape vise à définir le contenu des statistiques nécessaires. Elle découle directement du cadre conceptuel. Il s'agit de définir les données nécessaires d'une part à la qualification ou à la mesure des différents éléments (ou concepts) du cadre conceptuel et d'autre part aux tests des hypothèses sur les relations entre les facteurs.

La quatrième étape consiste à choisir la méthode d'enquête la plus adaptée pour collecter les données nécessaires. Nous précisons dans la section 2.3 (p. 104) ce que signifie le choix d'une méthode d'enquête.

Enfin, la dernière étape consiste à définir les traitements nécessaires pour répondre à l'objectif d'analyse fixé à la première étape. Cette étape est nécessaire avant la réalisation effective de l'enquête, car elle permet de vérifier que l'on peut effectivement traiter les données pour répondre aux différentes questions posées.

Notre présentation laisse entendre qu'il s'agit d'un processus séquentiel sans rétroaction. La pratique est en fait plus nuancée. Certes l'ordre dans lequel les étapes doivent être abordées ne peut être modifié pour d'évidentes raisons de logique. On imagine mal réfléchir au traitement avant de savoir quelles données seront produites. De même, le choix de l'enquête ne peut être fait avant de savoir quelles données doivent être recueillies (encore que les exemples existent malheureusement...). Ou encore comment pourrait-on définir le contenu des statistiques nécessaires avant de savoir ce que l'on recherche (il paraît que certains y arrivent, si l'on en croît certaines pratiques...). De même, la définition du cadre conceptuel semble pour le moins délicate en l'absence de problématique.

Si l'ordre apparaît pour le moins logique et « *relever de la méthode scientifique* » (Bonnafoos, Gérardin, 1976), il n'exclut pas les interactions entre les phases. En particulier, le choix de la méthode d'enquête peut contraindre le contenu des statistiques nécessaires. Par exemple, certaines questions peuvent sembler plus difficiles à poser selon le choix du mode d'enquête. Le choix d'une méthode d'échantillonnage aléatoire peut rendre l'analyse de groupes spécifiques plus difficiles car l'effectif risque d'être trop faible. A l'inverse, une stratification conduisant à des effectifs suffisants dans chacune des classes pourra s'avérer non optimale quant à la précision des résultats sur l'ensemble de la population.... Le choix d'une méthode de traitement peut imposer des contraintes sur les statistiques produites (par exemple nécessité d'une variable continue et non par classe...). Ainsi la construction de la séquence conduit dans la pratique à un certain nombre d'allers-retours entre les différentes étapes de la méthodologie d'analyse proposée. Par contre, la réalisation sur le terrain de l'enquête ne devrait jamais précéder la définition complète des 5 étapes, même si là encore de nombreuses études ou recherches sont en infraction avec ce mode d'opérer.

Avant de terminer cette section, nous souhaitons prévenir une critique. La justification de cette séquence méthodologique ne signifie pas qu'il ne soit pas possible de formuler de nouvelles hypothèses à partir d'un jeu de données. Bien au contraire, la pratique de recherche de nombreux laboratoires de recherche consiste souvent à exploiter des données d'enquêtes produites pour d'autres objectifs que celui qui est développé. Ces analyses peuvent conduire à formuler de nouvelles hypothèses qui peuvent donner naissance à de nouvelles problématiques, contredisant ainsi notre séquence méthodologique. Toutefois, les données disponibles ne permettent pas forcément de répondre pleinement à ces nouvelles problématiques car... les données n'ont pas été produites dans cet objectif. Prenons un exemple pour illustrer le propos. Nous le tirons du travail de Pochet (1995) sur l'usage de la voiture par les personnes âgées (chapitre 3, p. 88). Ce travail a mis en évidence une forte croissance de l'usage de la voiture chez les personnes âgées en partie observée chez les hommes et en fort devenir chez les femmes à partir de l'analyse de l'usage de la voiture chez les personnes plus jeunes. Pour approfondir l'analyse dans une perspective de modélisation, l'auteur a été confronté à deux questions. La première concerne l'affectation des véhicules aux membres du ménage. La seconde porte une éventuelle démotorisation (notamment pour les ménages bi-motorisés) lors du passage à la retraite ou à un âge plus avancé. Mais les données n'ont pas permis de répondre à ces questions. Pour la première question, l'affectation des véhicules au sein du ménage ne fait pas partie des questions posées lors de l'enquête ménages déplacements. Pour la seconde, aucune donnée de l'enquête ménages déplacements ne permet d'aborder la démotorisation chez les personnes âgées, tout particulièrement pour les ménages qui étaient bi-motorisés avant la retraite, notamment parce que l'information sur la motorisation passée du ménage n'est pas disponible dans l'enquête. (le lecteur qui souhaite approfondir cette question épistémologique pourra se reporter au chapitre 3 de l'ouvrage de Bonnafous, (Bonnafous, 1989)).

2.2. Méthodologie de la production des données pour la modélisation

La séquence méthodologique retenue pour la production des données pour la modélisation est tout à fait comparable à celle présentée pour l'analyse. Nous l'avons distinguée pour mieux mettre en évidence les besoins relatifs à la modélisation.

La première étape traite toujours des objectifs et de la problématisation de l'objet d'analyse que l'on cherche à prévoir ou simuler.

La seconde étape s'apparente en fait à la construction d'un cadre conceptuel. Il s'agit de définir d'une part les éléments qui influent sur l'objet d'analyse et qui devront être intégrés dans le modèle et d'autre part les relations entre ces éléments. A ce niveau, selon la connaissance que l'on a de l'objet d'analyse, il s'agit soit d'hypothèses que la production des données doit venir valider, soit d'une construction que l'on considère comme une approximation acceptable de la réalité compte tenu des objectifs que l'on s'est assignés.

Comme pour le schéma 4.1, la troisième étape concerne la définition des statistiques nécessaires. Les données peuvent être nécessaires à trois niveaux (chapitre 2, p. 38) :

- test des hypothèses formulées à l'étape précédente. Cette production ne sera pas nécessaire si l'on considère le modèle comme acquis ;
- formalisation mathématique des relations entre les variables prises en compte. De nouveau cette production ne sera pas nécessaire si l'on utilise un modèle déjà spécifié ;
- calage du modèle. De nouveau cette production ne sera pas nécessaire si l'on utilise un modèle déjà calé pour une étude précédente. Il faut toutefois que ce calage ait été effectué sur la même zone d'étude, avec le même zonage (ou au moins un zonage plus fin) et avec la même période de référence. Si l'un de ces paramètres doit évoluer, il n'est plus possible de faire l'économie de cette production ne serait-ce que pour décrire le réseau ou collecter des données de trafic ou les données socio-économiques.

Enfin, pour les deux dernières étapes nous pouvons reprendre les commentaires proposés dans la section précédente relative à la méthodologie de la production des données pour l'analyse.

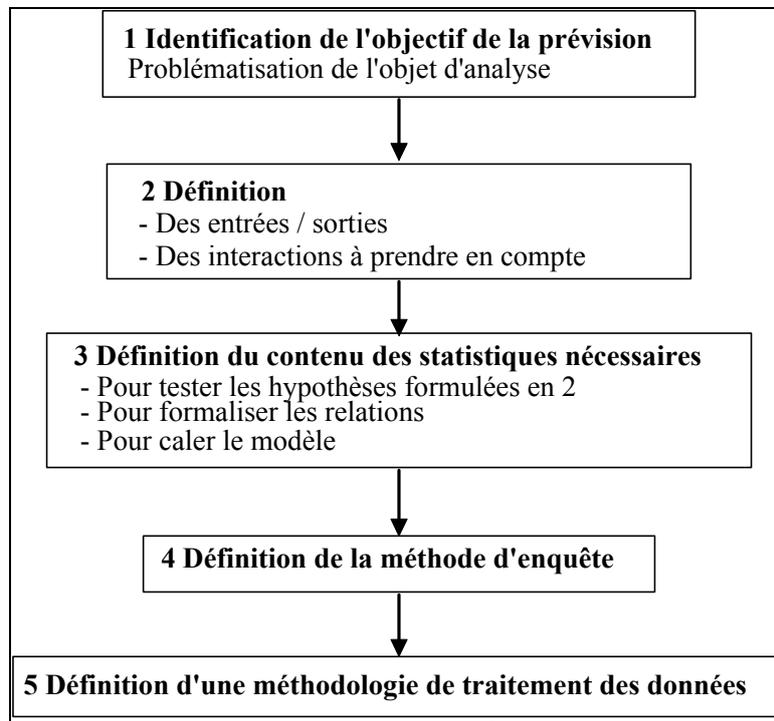


Schéma 4.2 : Méthodologie de la production des données pour la modélisation

On retrouve donc deux séquences méthodologiques qui à la terminologie près sont identiques. Il est en fait possible de généraliser cette séquence à l'ensemble des enquêtes quel qu'en soit l'objectif. Toutefois, nous ne développons pas ce point car il s'éloignerait trop de notre propos.

2.3. Définition de la méthode d'enquête

Faire une enquête ne se limite pas à la rédaction d'un questionnaire suivie de la réalisation d'un terrain. La définition d'une enquête suppose en fait la réponse à un certain nombre de questions. Ces questions peuvent être organisées en grands thèmes interagissant entre eux (schéma 4.3).

Cette présentation ne prétend pas à l'exhaustivité. En particulier, certaines dimensions ne sont pas identifiées en tant que telles, comme le recueil préalable des données ou des informations existantes sur l'objet d'étude, le test du questionnaire, la réalisation préalable d'entretien pour débroussailler le sujet ou recueillir du discours sur l'objet d'étude, l'information autour de l'enquête, le redressement des données... Nous ne voulons pas sous-estimer l'importance de ces tâches, par exemple nous estimons qu'un questionnaire ne devrait jamais être diffusé avant d'avoir été testé au préalable. Toutefois, cette présentation reprend les principales étapes d'une enquête qui doivent être traitées avant la réalisation du terrain en raison de leur interdépendance et de leur importance dans la qualité des données produites. Nous explicitons rapidement le contenu de ces étapes et mettons en lumière cette interdépendance entre les étapes. Toutefois, nous renvoyons aux auteurs cités au début de ce chapitre pour leur approfondissement.

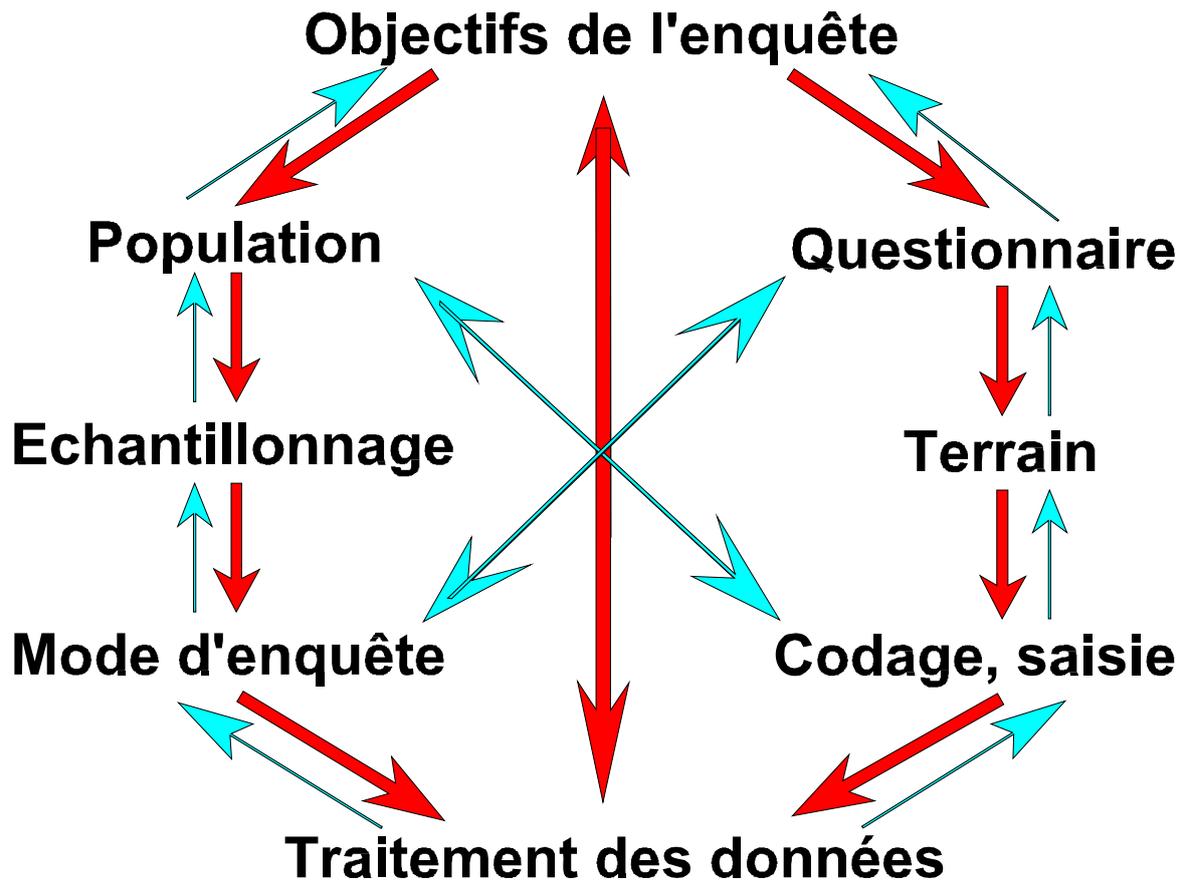


Schéma 4.3 : Schéma de définition d'une enquête par questionnaire

2.3.1. Définition de l'objectif

La définition de l'objectif est le préalable à toute réflexion. Nous en avons donné ici une définition large qui englobe en fait les trois premières étapes des schémas 4.1 et 4.2.

2.3.2. Définition de la population cible

La population cible correspond à l'ensemble de la population pour laquelle on souhaite recueillir des données. Nous préférons utiliser ce terme plutôt que celui de population statistique fréquemment employé. La population statistique fait souvent référence indifféremment à la population cible telle que nous venons de la définir ou à la base de sondage que nous définissons dans la section suivante. Et nous verrons qu'il n'y a pas toujours adéquation parfaite entre population et base de sondage. La population cible peut être une population de ménages, d'individus, de déplacements, de véhicules....

Une définition explicite des objectifs permet en principe de répondre à la question. Un exemple illustrera mieux le propos. Prenons une enquête sur les déplacements urbains. Si l'objectif visé est un recueil de l'ensemble des déplacements réalisés sur un périmètre donné, la population cible sera une population de déplacements réalisés à l'intérieur de ce périmètre. En revanche, si l'on souhaite étudier les comportements de déplacements urbains des résidents du périmètre, les individus résidents constitueront la population cible, voire les ménages si l'on s'intéresse plus particulièrement aux interactions entre individus d'un même ménage. Il convient ensuite de préciser le terme urbain. S'agit-il d'une enquête sur les déplacements urbains en général qui conduirait à retenir l'ensemble des agglomérations dans la population cible, ou d'une enquête sur un certain type d'agglomérations ou sur

une agglomération particulière ? La délimitation de l'urbain mérite ensuite d'être définie. La population cible sera-t-elle restreinte aux limites administratives ou couvrira-t-elle le bassin d'attraction de l'agglomération ?....

Prenons un autre exemple relatif à une enquête de qualité de service sur les transports collectifs. Si l'objectif concerne la construction d'un baromètre de satisfaction des usagers d'une ligne de bus, la population cible sera définie comme étant la population des déplacements ou des individus sur cette ligne. Le choix d'une population de déplacements conduit à donner un poids identique à chacun des déplacements alors qu'une population d'individus fréquentant la ligne conduit au contraire à donner un poids identique à chacun des individus indépendamment de leur fréquentation de la ligne. En revanche, si l'objectif devient la connaissance des paramètres de satisfaction en vue d'action de promotion des transports collectifs pour accroître leur usage, la population cible ne pourra plus être limitée aux seuls usagers des transports collectifs. Ces deux exemples, a priori très simples, illustrent la nécessité d'une clarification préalable des objectifs de la production des données.

2.3.3. Définition de la méthode d'échantillonnage

L'échantillonnage est une méthode qui permet de ne pas enquêter l'ensemble de la population cible, tout en produisant des résultats relatifs à l'ensemble de cette population. Un sondage (enquête sur un échantillon de la population cible) se distingue ainsi d'un recensement qui vise au contraire à recueillir des données pour chacun des individus statistiques constituant la population cible.

La première étape consiste à définir l'unité statistique. Ce choix découle directement de la définition de la population cible. Dans l'exemple de l'enquête déplacements urbains, l'unité statistique peut être le déplacement, l'individu ou le ménage selon l'objectif. Le plus souvent pour les enquêtes transport, l'unité statistique correspond à l'une des catégories suivantes : ménage, individu, déplacement, véhicule, zone, ligne de transports collectifs, section de route (Richardson et al., 1995).

La seconde étape consiste à définir la base de sondage (on l'appelle également base de tirage de l'échantillon). En théorie celle-ci devrait correspondre à la population cible. Malheureusement, il y a rarement adéquation entre les bases de sondage existantes et la définition choisie de la population cible. Si l'on reprend l'exemple des déplacements urbains, il est généralement possible de disposer d'une base de sondage des ménages appartenant à un périmètre urbain (comme le recensement, l'annuaire téléphonique, les abonnés à différents réseaux... même si la base n'est jamais totalement exhaustive). En revanche, il est beaucoup plus délicat de bâtir une base de sondage déplacements qui soit représentative de l'ensemble des déplacements réalisés à l'intérieur d'un périmètre donné. Pour les transports collectifs, la tâche est généralement moins complexe, car il est possible de bâtir des protocoles de tirages à plusieurs degrés couvrant l'ensemble de la population cible (comme nous l'explicitons ci-après). En revanche, pour les déplacements automobiles, un tel protocole est généralement beaucoup plus difficile à mettre en place si la population cible couvre l'ensemble des déplacements quel que soit le réseau viaire utilisé. Il est fréquent dans ce cas d'utiliser une base de ménages ou d'individus ou une base de véhicules. La base de sondage diffère alors de la population cible. Prenons l'exemple d'une base de sondage de ménages pour illustrer le problème. Si la base correspond seulement aux ménages résidant au sein du périmètre, on dispose seulement des déplacements des résidents du périmètre d'étude. Trois solutions s'offrent alors :

- réduire la population cible et donc l'objectif à l'étude des déplacements des seuls résidents ;
- élargir le périmètre d'enquête pour englober des ménages résidents à proximité du périmètre et donc susceptibles de s'y déplacer. Mais dans ce cas, on ne fait que réduire le problème et non le supprimer totalement ;
- compléter le dispositif avec d'autres enquêtes pour enquêter les déplacements des non résidents sur le périmètre d'étude ou les déplacements non repérés par l'enquête (nous l'illustrons plus loin avec l'exemple des enquêtes cordon (section 3.2, p. 121)).

La troisième étape consiste à définir la taille de l'échantillon. De nouveau, elle découle directement des objectifs visés. Si l'on veut mesurer un indicateur, la taille de l'échantillon résultera de la précision recherchée (section 2.4.2, p. 112 ou pour plus de détail sur les méthodes de calcul Focas, 1989 ; Ardilly, 1994 ; Richardson et al., 1995). Si l'on veut comparer des sous-populations par classe d'âge, il faudra disposer d'un effectif suffisant dans chaque classe d'âge pour les discriminer. Une représentativité structurelle (méthode par quota ou sondage à probabilité inégale) conduisant à un poids identique dans chaque classe d'âge sera alors privilégiée par rapport à un tirage aléatoire qui conduirait à un poids de chaque classe d'âge proportionnel au poids de la classe dans l'ensemble de la population cible. Le budget étant toujours limité, le choix de la méthode d'échantillonnage limitera en retour les objectifs que l'on peut assigner à l'enquête. Pour reprendre notre exemple précédent, si les objectifs principaux de l'enquête conduisent à un tirage aléatoire sur l'ensemble de la base de sondage, il faudra peut être renoncer à étudier de manière très fine les classes d'âge les plus âgées dont le poids est plus faible dans l'ensemble de la population. Le choix de la taille de l'échantillon conditionne donc les analyses qui peuvent être menées et donc l'objectif de l'enquête.

Enfin, vient le choix de la méthode de tirage de l'échantillon. On distingue principalement deux méthodes de tirage : tirage aléatoire, méthode par quota. Le tirage aléatoire simple consiste à tirer un échantillon de manière aléatoire à partir d'une base de sondage. La méthode par quota consiste à sélectionner un échantillon répondant à certaines caractéristiques de quotas (par exemple quota par classes d'âge). Le tirage aléatoire peut être subdivisé en plusieurs sous-méthodes (on trouvera des développements dans Ardilly, 1994 et Richardson et al., 1995). Nous nous limitons à celles qui sont le plus souvent utilisées dans les enquêtes déplacements :

- sondage aléatoire simple qui correspond à la méthode générale décrite ci-dessus ;
- sondage stratifié. La méthode consiste à opérer une partition de la population cible pour effectuer un tirage aléatoire au sein de chacune des sous-populations. La stratification répond généralement à l'un des deux objectifs suivants. Le premier correspond à une recherche de précision statistique. Dans ce cas, on privilégie en général une variable de stratification fortement corrélée à l'objet d'analyse (afin que la variance interclasse soit maximale et la variance intraclasse la plus faible possible). Le second correspond à un objectif de précision minimum au sein de chacune des classes. Le choix de la variable de stratification découle alors des objectifs de l'étude. On distingue généralement les sondages stratifiés à probabilité constante (la probabilité de tirage d'un individu statistique est identique dans chacune des strates) des sondages stratifiés à probabilités variables ou inégales (la probabilité n'est pas constante). Par exemple, les enquêtes ménages déplacements françaises appartiennent à ce second groupe de sondage stratifié (section 3.1, p. 117). Pour ces enquêtes, on utilise une stratification géographique garantissant un effectif suffisant dans chacune des zones de tirage. Le tirage est ensuite réalisé aléatoirement dans chacune des zones. La stratification apporte toujours un gain de précision par rapport à un sondage aléatoire simple (Ardilly, 1994). Il faut pour cela connaître la répartition de la population selon les strates. En revanche, un sondage stratifié à probabilité inégale n'offre pas forcément une précision supérieure à celle d'un sondage stratifié à probabilité égale (section 2.4.2.2, p. 114) ;
- sondage à plusieurs degrés. On utilise cette méthode principalement lorsque l'on ne dispose pas d'une base de tirage ou lorsque le coût des déplacements nécessaires pour réaliser les enquêtes est trop important par rapport à l'ensemble du budget d'enquête. Comme pour le sondage stratifié, la première étape consiste à effectuer une partition de la population cible. On effectue ensuite un tirage aléatoire simple (ou une autre méthode d'échantillonnage) au sein de ces groupes pour n'en sélectionner qu'une partie. Enfin, au sein des groupes sélectionnés, on réalise un tirage aléatoire simple (ou tout autre méthode de tirage). Ce mode de tirage correspond à une sélection de l'échantillon opérée en plusieurs étapes. Prenons l'exemple d'une enquête de satisfaction des usagers des transports collectifs pour l'illustrer. Il serait trop coûteux de tirer l'échantillon à partir d'une base englobant tous les ménages d'une agglomération. Par contre, on ne dispose généralement pas de base des individus utilisant les transports collectifs ou de base des déplacements effectués en transports collectifs. Pour

répondre à ce problème, un sondage à trois degrés permet d'obtenir un échantillon représentatif de l'ensemble de la population des déplacements en transports collectifs : tirage sur les lignes de bus, tirage de services pour les lignes sélectionnées, tirage d'individus dans chaque service sélectionné. L'échantillon ainsi constitué est un échantillon représentatif des déplacements effectués en transports collectifs. En revanche, cette base est biaisée si l'on veut produire des résultats au niveau des individus. Il est alors nécessaire de pondérer les déplacements par l'inverse du nombre de déplacements effectués par l'individu enquêté.

Ces méthodes peuvent être combinées pour bâtir des plans d'expérience assez complexes. Cette complexité du plan d'échantillonnage se retrouve bien évidemment dans la détermination de la précision des estimateurs des grandeurs étudiées. Si ces méthodes permettent le plus souvent des économies de coût de terrain d'enquête et de précision pour certains indicateurs (ceux ayant justifié le plan d'échantillonnage choisi), ce n'est pas toujours le cas des autres indicateurs dont la précision peut être réduite.

L'échantillonnage est donc le résultat d'un processus assez complexe méritant une attention particulière. Le choix de la méthode découle directement des objectifs de l'enquête. Mais en retour, le choix d'un plan d'échantillonnage conditionne la précision des données produites et donc les analyses pouvant être menées. L'ajout de nouveaux objectifs à l'issue de l'enquête n'est plus alors toujours possible. De même, en l'absence de base de sondage correspondant pleinement à la population cible, le plan d'échantillonnage retenu peut conduire à déformer cette population. Cette complexité est évidemment source de biais potentiels à chacune des étapes de l'échantillonnage. Il est important de prendre en compte ces biais lors de l'estimation de la précision statistique (section 2.4, p. 110).

2.3.4. Définition du mode d'enquête

Le choix du mode d'enquête nécessite en fait de répondre à plusieurs questions :

- sur le moyen de communication utilisé : en face à face, par téléphone, par voie postale. De nouveaux moyens de communication commencent à être utilisés comme internet (Wermuth et al., 2001). Il est en fait possible de distinguer la distribution du questionnaire, de la passation de l'enquête, voire de la relance des enquêtés ;
- sur le type d'administration de l'enquête : administré (questions formulées oralement par l'enquêteur) ou auto-administré (questions lues et réponses notées par l'enquêté lui-même) ;
- sur le lieu de l'enquête : au domicile, sur le lieu de travail, sur les lieux d'un générateur de trafic particulier, dans un véhicule, sur le bord de la route....

Ces dimensions peuvent se combiner, même si certaines sont incompatibles. De même, au sein de chacune de ces dimensions, il est possible de combiner les techniques : l'envoi postal préalable d'un questionnaire suivi d'un appel téléphonique pour recueillir les réponses ; les relances téléphoniques dans une enquête postale ; le face à face ou le postal pour joindre les ménages sans téléphone ou sur liste rouge dans une enquête téléphonique... (Bonnel, 2001).

De nouveau l'objectif, comme les étapes précédentes, conditionne en partie le choix du mode d'enquête. Par exemple, l'échantillonnage proposé ci-dessus (sondage à plusieurs degrés) pour l'enquête sur la qualité de service dans les transports collectifs ne peut utiliser comme moyen de communication que le face à face (pour le contact initial, même si l'enquête peut ensuite être auto-administrée avec retour à l'enquêteur ou par voie postale) et conditionne le lieu de l'enquête : en véhicule ou en station.

Inversement, le choix d'un mode d'enquête peut conditionner les autres étapes. Par exemple dans une enquête par voie postale, on utilise généralement des questionnaires assez courts pour limiter les refus déjà assez importants par ailleurs. L'étendue des questionnements devra donc être en rapport avec les possibilités de ce type d'enquête. De même, le faible taux de réponse peut interdire certaines

analyses pour lesquelles on craint une corrélation plus forte entre le fait d'accepter de répondre et la question posée. Dans une enquête sur le bord de la route, on ne peut généralement pas se permettre d'arrêter longtemps le véhicule. L'enquête doit alors être très courte ce qui limite l'étendue des objectifs. De même, un tirage aléatoire simple ne sera pas envisageable puisque l'on ne dispose pas d'une base de sondage préalable des déplacements effectués sur l'infrastructure en un point donné pour effectuer le tirage (on dispose tout au plus de comptage, ce qui ne signifie pas qu'il ne soit pas possible de construire un échantillon représentatif de la population cible, par exemple en enquêtant tous les xième véhicules, ou mieux encore en enquêtant tous les véhicules dont l'ordre de passage correspond à une table de nombre aléatoire). Dans une enquête téléphonique, les personnes non abonnées ou sur liste rouge sont exclues de la population enquêtée introduisant un biais par rapport à la population cible (Fréjean et al, 1990)....

2.3.5. Rédaction du questionnaire

Au-delà de la facilité apparente, il s'agit en fait de l'étape la plus délicate. Elle l'est d'autant plus qu'il n'existe pas de recette pour construire le bon questionnaire. Tout au plus, on dispose de quelques recommandations et surtout d'une liste de pièges à éviter. La littérature (Ghiglione, Matalon, 1985 ; Javeau, 1985 ; Richardson et al., 1995...) nous indique donc surtout ce qu'il ne faut pas faire, plus que ce qu'il faut faire.

Bien évidemment, c'est de nouveau l'objectif de l'enquête qui conditionne le contenu du questionnaire, sa structure, le type de questions, la formulation des questions, le type de réponse attendu, le design du questionnaire.... Le choix de la population cible peut également avoir des conséquences sur le questionnaire. Parmi les règles d'art du bon questionnaire, on trouve une formulation des questions et des réponses adaptée aux personnes auxquelles le questionnaire est adressé. Selon la population retenue, on pourra alors faire varier la rédaction des questions. Le mode d'enquête conditionne également la rédaction du questionnaire. Un questionnaire auto-administré ne se présente pas ni se rédige comme un questionnaire administré. L'utilisation de support comme des cartes pour le codage des lieux est possible en face à face. En revanche, il est plus difficile à mettre en œuvre dans une enquête téléphonique....

A l'inverse les choix effectués sur le questionnaire conditionnent les analyses pouvant être menées ultérieurement. Choisir des questions ouvertes ou fermées ; à réponses uniques ou multiples ; quantitatives ou qualitatives... ne permet pas de développer les mêmes analyses par la suite et donc de répondre aux mêmes objectifs.

2.3.6. Organisation du terrain

L'organisation du terrain comprend l'ensemble des tâches logistiques nécessaires au bon déroulement de l'enquête. Ces tâches sont nombreuses et dépendent en partie du mode d'enquête retenu. Nous préférons donc renvoyer le lecteur vers les ouvrages déjà cités qui traitent longuement de ces questions. Nous mentionnons toutefois l'importance de certaines dimensions :

- diffusion de l'information préalable aux enquêtés, directement par le biais de courrier ou indirectement par différents médias ;
- le recrutement et la formation des enquêteurs et des superviseurs. Cette dernière est toujours nécessaire même pour une enquête simple. Elle l'est d'autant plus que l'enquête se complexifie ;
- l'organisation du suivi et du contrôle de la qualité des enquêtes : contrôle du travail des enquêteurs, contrôle de la cohérence des réponses, contrôle du codage, de la saisie... ;
- l'usage des systèmes CATI (« *Computer-Aided Telephone Interview* », enquête téléphonique assistée par ordinateur), CAPI (« *Computer-Aided Personal Interview* », enquête en face à face assistée par ordinateur) permet de simplifier fortement le travail de l'enquêteur ainsi que le contrôle de la qualité de leur travail.

Cette étape découle directement des précédentes. En revanche, elle a moins d'incidence sur les précédentes, si ce n'est que la qualité de l'exécution de cette étape conditionne la qualité des données produites et donc la possibilité de répondre aux objectifs.

2.3.7. Codage, saisie des données

C'est une étape apparemment purement technique. En fait, comme les précédentes, cette étape est importante pour s'assurer de la qualité des données produites.

Le choix du codage des données conditionne également le type d'analyse pouvant être mené ultérieurement. Nous prendrons l'exemple du codage des adresses. L'usage le plus fréquent consiste à coder les adresses à l'aide d'un zonage. Une fois ce zonage établi et le codage effectué, il n'est plus possible de le modifier, sauf à retenir un zonage plus agrégé. Pourtant, les analyses menées ultérieurement nécessitent bien souvent de recourir à des zonages différents. Il serait alors souhaitable de recueillir une information en coordonnée x-y (Bonnel et al., 1994b ; Chapleau, 1991, 1993, 1997 ; Dorland et al., 1997) qui permet ensuite de choisir le zonage adapté selon le type de questions posées (nous revenons sur cette question dans les sections 4.4, p. 127 et 4.6.5, p. 132). L'objectif est donc également conditionné par les choix qui sont faits au niveau du codage.

2.3.8. Définition du traitement des données

Cette dimension a déjà été abordée dans la discussion sur les schémas 4.1 et 4.2, nous n'y reviendrons pas, sauf pour rappeler que cette étape est indispensable pour s'assurer de la capacité effective à pouvoir répondre aux objectifs initiaux. C'est bien souvent à cette étape que l'on se rend compte de certaines difficultés non repérées auparavant. De même, à l'issue de cette étape, il est également bien souvent nécessaire de réduire l'étendue des objectifs que l'on s'était assigné au départ, compte tenu des limites inhérentes aux choix effectués à chacune des étapes.

Cette présentation rapide met bien en évidence l'interdépendance entre les différentes étapes de la définition d'une enquête. Si l'ordre de présentation que nous avons adopté nous semble le plus logique, il n'est pas exclusif d'autres séquences. L'important étant de traiter l'ensemble de ces étapes avant de commencer le terrain.

Avant de clore cette présentation de la séquence de définition d'une enquête, nous souhaitons rappeler l'importance du test du questionnaire. Ce test doit être réalisé dans les conditions les plus proches possibles de la réalisation effective de l'enquête. Il permet bien souvent de se rendre compte de certains problèmes du questionnaire : questions mal comprises, item de réponse oublié... et de certains problèmes de terrain : absence d'éclairage dans un bus interurbain pour des enquêtes matinales ne pouvant dès lors être remplies en auto-administré ; personne ayant oublié ses lunettes pour lire le questionnaire ; montées trop nombreuses à un arrêt... (ces quelques exemples sont choisis parmi une longue liste d'observations issues de tests réalisés dans le cadre des enquêtes menées pour la formation du DESS transports urbains et régionaux de personnes).

2.4. Précision statistique

Tant le producteur d'enquêtes que l'utilisateur de données issues d'enquêtes doivent constamment conserver à l'esprit la précision statistique des données. Trop souvent, on voit des analyses menées sans aucune considération de l'incertitude entourant les données utilisées. Cette situation est d'autant plus présente dans les enquêtes déplacements que les données sont souvent ramenées à la population totale après redressement en multipliant chaque observation par l'inverse du

taux de sondage. De ce fait, l'utilisateur s'abrite derrière l'effectif ainsi affiché et souvent confortable alors que l'effectif enquêté peut ne plus l'être du tout.

Nous présentons les méthodes de calcul de la précision pour les sondages les plus couramment utilisés (nous renvoyons aux ouvrages déjà cités et notamment Ardilly, 1994 pour les cas plus complexes).

Les différentes techniques de sondage permettent de calculer la précision des mesures effectuées pour un niveau de confiance donnée, ce qui n'est pas possible dans les méthodes d'enquête par quotas. Mais il est important de rappeler que les calculs de précision proposés dans les ouvrages statistiques ou sur les méthodes d'enquête, nécessitent le respect strict de ces méthodes d'échantillonnage. S'il est possible de calculer l'erreur d'échantillonnage, l'erreur totale de mesure ne se résume pas à cette erreur. Dans une enquête par sondage, on a l'équation suivante :

$$\ll \text{Erreur totale} = \text{erreur d'échantillonnage} + \text{erreur d'observation} + \text{erreurs dues au défaut de couverture et à la non-réponse} \gg \text{ (Ardilly, 1994).}$$

Il est donc important de prendre en compte l'ensemble des biais potentiels introduits à chacune des étapes de définition de la méthodologie d'enquête.

Dans cette section, nous nous concentrons sur la non-réponse à travers l'analyse du taux de réponse (section 2.4.1) et sur l'erreur d'échantillonnage à travers le calcul de la précision statistique (section 2.4.2).

2.4.1. Non-réponse

Parmi les non-réponses, la non-réponse totale qui correspond à une absence complète de réponse à l'enquête (par opposition à la non-réponse partielle qui ne concerne qu'une partie du questionnaire) est sûrement le problème le plus important qui devrait être documenté dans toutes les enquêtes. Le taux de réponse devrait être fourni systématiquement. Les modes de calcul extrêmement variables que l'on peut trouver dans la littérature sur les rapports d'enquête nécessitent de rappeler le mode de calcul généralement admis (Richardson et al., 1995). La première étape consiste à enlever de l'échantillon à enquêter (issu du plan d'échantillonnage), les individus statistiques hors cibles (par exemple les logements vacants, ou les numéros de téléphone invalides...). On constitue ainsi l'échantillon à enquêter appartenant à la cible. Le taux de réponse est alors le pourcentage de réponses complètes parmi l'échantillon à enquêter appartenant à la cible. Reprenons l'exemple cité dans Richardson et al. (1995) :

« Soit 100 ménages tirés selon le plan d'échantillonnage, 5 logements vacants et 64 réponses complètes, le taux de réponse sera :

<i>échantillon de tirage</i>	<i>100</i>
<i>échantillon hors cible</i>	<i>5</i>
<i>échantillon cible</i>	<i>95</i>
<i>réponses</i>	<i>64</i>
<i>Taux de réponse</i>	<i>64/95 = 67,4 %</i>

Outre le taux de réponse, il est souhaitable de documenter, dans les rapports d'enquête, la source des non réponses : individu non joignable, refus de réponse, questionnaire mal rempli....

Il existe deux sources principales de non réponses : l'individu statistique n'est pas joignable, il refuse de répondre. Pour faire face au premier problème, il faut mettre en place un protocole de contact et de relances (Dijkstra, Smit, 1991 ; Kulka, Weeks, 1988). Dans une enquête téléphonique, il faudra téléphoner plusieurs fois à des heures ou des jours différents, voire envoyer l'enquête par voie postale. Dans une enquête postale, il faudra assurer plusieurs relances de courrier, voire téléphoner ou se

rendre au domicile du ménage.... Pour limiter les refus de réponse, il est souhaitable de recourir à des enquêteurs professionnels et d'assurer une information préalable sur l'enquête à l'aide de différents médias... (pour une analyse plus détaillée des causes de non réponse, nous renvoyons sur les ouvrages cités au début de ce chapitre).

Les non-réponses réduisent la précision des données car il n'est généralement pas possible de postuler à l'indépendance entre la non-réponse et l'objet d'étude. Dans les enquêtes déplacements, la dépendance a d'ailleurs déjà été établie (Ampt, 1997 ; Sammer, 1995 ; Bonnel, 2001). Ces auteurs ont mis en évidence une décroissance du nombre de déplacements quotidiens en fonction du nombre de relances nécessaires pour obtenir un retour dans une enquête postale. Le résultat généralement inverse est obtenu dans une enquête téléphonique (Bonnel, 2001). Enfin, Ampt (1997) a montré dans une analyse des causes de non réponse, qu'une partie des refus de réponse était liée à une mobilité trop faible (le sujet ne s'estime pas intéressant pour l'organisme réalisant l'enquête) ou trop forte (le sujet n'est alors pas assez disponible ou craint la longueur de l'enquête du fait de son « hyper-mobilité »). Cette dépendance peut être réduite lors du redressement. Il faut pour cela disposer de variables corrélées à l'objet d'analyse et présentes à la fois dans la base de tirage (la présence de nombreuses variables dans le recensement en fait une base privilégiée dans cette perspective) et dans l'enquête. Il est alors possible de mener une analyse des non-réponses sur ces variables et de faire un redressement sur les variables « explicatives » de la non-réponse (Dupont, 1995 ; Caron, 1996 ; Armoogum, Madre, 1997 ; Armoogum, 1999 ; Lejeune, 2001). Il est malgré tout nécessaire de faire une hypothèse implicite : au sein du groupe d'individus possédant les mêmes caractéristiques socio-économiques, ceux qui ont répondu à l'enquête ont en moyenne le même comportement que ceux qui n'y ont pas répondu (sous cette hypothèse, il est possible de calculer la perte de précision liée à la non réponse (Ardilly, 1994)). Des travaux récents remettent toutefois en cause cette hypothèse (Richardson, 2000). Il est ainsi souhaitable de mettre en place des études spécifiques concernant les non-répondants (en général des enquêtes plus légères ou à l'aide d'autres modes d'enquête) afin de recueillir de l'information les concernant permettant d'affiner les procédures de redressement. L'enquête déplacements de Melbourne (Australie) constitue un bon exemple de telles procédures (Richardson, Ampt, 1993).

A côté des non réponses totales, on observe souvent des non réponses partielles. Une partie des questions n'a pas reçue de réponse. Si le nombre de questions n'est pas trop élevé, on conserve généralement l'enquête. Il existe alors diverses techniques d'imputation pour « corriger » ces non réponses partielles (Ardilly, 1994 ; Armoogum, Madre, 1997 ; Armoogum, 1999 ; Lejeune, 2001 ; Brownstone, 1998 traite également du problème dans une perspective plus large).

2.4.2. Calcul de la précision

Le calcul de la précision dépend du plan d'échantillonnage. Dans un premier temps, nous présentons le cas le plus simple du tirage aléatoire simple (Ardilly, 1994 ; Richardson et al., 1995) (section 2.4.2.1). Nous abordons ensuite le cas du sondage stratifié (section 2.4.2.2), puis du sondage en grappe (section 2.4.2.3).

Nous n'abordons pas le cas des méthodes par quotas, car ces méthodes de construction de l'échantillon ne permettent pas de définir des intervalles de confiance. Toutefois, il est reconnu que pour de petits échantillons, un sondage par quotas donne des estimations biaisées mais plus précises qu'un sondage aléatoire. Les estimations sans biais fournies par un sondage aléatoire ne deviennent précises que quand l'échantillon atteint une taille suffisante (Ardilly, 1994).

Il est utile de rappeler que les formules de calcul de la précision, que nous donnons ci-dessous, suppose que les hypothèses nécessaires au calcul sont respectées, et notamment un taux de réponse de 100 % ... ce qui n'est évidemment pas le cas (il est possible d'estimer l'incidence du non-respect de

certaines hypothèses. C'est le cas des non-réponses sous certaines conditions particulières (Ardilly, 1994)). Les calculs fournissent donc une sur-estimation de la précision.

2.4.2.1. Sondage aléatoire simple

Les formulations que nous proposons ici s'appuient sur les travaux de Armoogum auxquels nous renvoyons le lecteur pour plus de détail (Armoogum et al., 2001 ; voir aussi Ardilly, 1994 ; Richardson et al., 1995 ; Wonnacott, Wonnacott, 1991). La formulation générale d'un intervalle de confiance pour un paramètre X au seuil de confiance $(1 - \alpha)$ est donnée par :

$$\left[\hat{x} - z_{(1-\alpha/2)} \hat{\sigma}(\hat{x}) \quad ; \quad \hat{x} + z_{(1-\alpha/2)} \hat{\sigma}(\hat{x}) \right] \quad (4.1)$$

où \hat{x} est l'estimateur sur la population enquêtée de X ;

$z_{(1-\alpha/2)}$ représente le quantile d'ordre $1-\alpha/2$ d'une variable aléatoire normale centrée réduite. On fait pour cela l'hypothèse que \hat{x} suit une loi normale (on considère en général que lorsque l'effectif est suffisant (au moins 30, Richardson et al., 1995) \hat{x} suit effectivement une loi normale) ;
 $\hat{\sigma}(\hat{x})$ représente l'estimateur sur la population enquêtée de l'écart-type de X.

Pour une moyenne, la formulation devient (Armoogum et al., 2001) :

$$\left[\hat{\bar{x}} - z_{(1-\alpha/2)} \sqrt{\frac{(N-n)}{N} * \frac{1}{n} \hat{s}_x} \quad ; \quad \hat{\bar{x}} + z_{(1-\alpha/2)} \sqrt{\frac{(N-n)}{N} * \frac{1}{n} \hat{s}_x} \right] \quad (4.2)$$

où : $\hat{\bar{x}}$ est un estimateur sur la population enquêtée de la moyenne de la variable X,

$$\hat{\bar{x}} = \sum_{k=1}^n \frac{x_k}{n} ;$$

\hat{s}_x représente l'écart-type sur la population enquêtée de la variable X,

$$\hat{s}_x = \sqrt{\frac{1}{(n-1)} \sum_{k=1}^n (x_k - \hat{\bar{x}})^2} , \quad x_k \text{ étant la valeur observée de la variable X pour l'individu k}$$

appartenant à l'échantillon enquêté ;

X un paramètre de la population N ;

n le nombre d'individus d'un échantillon de la population N ;

$z_{(1-\alpha/2)}$ représente le quantile d'ordre $1-\alpha/2$ d'une variable aléatoire normale centrée réduite.

Pour un pourcentage P, si l'effectif enquêté est suffisant (dans ce cas, il faut que l'effectif correspondant au paramètre P, soit $\hat{p} * n$ et son complément, soit $(1-\hat{p}) * n$ soit suffisant, en général au moins 20), on peut considérer que la variable dont on cherche à calculer la moyenne suit une loi normale. L'intervalle de confiance de la proportion P s'écrit alors (Armoogum et al., 2001) :

$$\left[\hat{p} - z_{(1-\alpha/2)} * \sqrt{\frac{(N-n)}{N} * \frac{1}{n-1} * \hat{p}(1-\hat{p})} \quad ; \quad \hat{p} + z_{(1-\alpha/2)} * \sqrt{\frac{(N-n)}{N} * \frac{1}{n-1} * \hat{p}(1-\hat{p})} \right] \quad (4.3)$$

\hat{p} est un estimateur sur la population enquêtée de la proportion de la variable P ;

P un paramètre de la population N ;

n le nombre d'individus d'un échantillon de la population N ;

$z_{(1-\alpha/2)}$ représente le quantile d'ordre $1-\alpha/2$ d'une variable aléatoire normale centrée réduite.

La précision ne dépend donc pas du taux de sondage, mais de la taille de l'échantillon (sauf pour des populations de petite taille, car dans ce cas le rapport $\sqrt{\frac{(N-n)}{N}}$ n'est plus proche de 1 et dépend donc du taux de sondage). Elle varie en fonction de la racine carrée de la taille de l'échantillon. Pour doubler la précision (c'est-à-dire diviser par 2 la largeur de l'intervalle de confiance), il faut multiplier par 4 la taille de l'échantillon.

2.4.2.2. Sondage stratifié

Dans le cas des enquêtes ménages déplacements françaises, on ne dispose pas d'un tirage aléatoire simple, mais d'un tirage aléatoire simple après stratification géographique. La stratification retenue conduit à enquêter un nombre identique de ménages dans chaque zone. Elle ne vise pas à une optimisation du plan de sondage pour améliorer la précision sur certains indicateurs, mais à avoir une précision équivalente pour chacune des zones de tirage au sein du périmètre d'étude.

Le calcul de la précision pour une moyenne dans un sondage stratifié nécessite le calcul de l'estimateur sur la population enquêtée de l'écart-type de la moyenne de X, qui est donné par la formule suivante (Ardilly, 1994) :

$$\hat{\sigma}(\hat{x}_{st}) = \sqrt{\sum_{h=1}^H \left(\frac{N_h}{N}\right)^2 * \frac{N_h - n_h}{N_h} * \frac{1}{n_h} * \hat{s}_h^2} \quad (4.4)$$

avec \hat{s}_h l'écart-type de la variable X calculée sur l'échantillon de la strate h,

$$\hat{s}_h = \sqrt{\frac{1}{(n_h - 1)} \sum_{k=1}^{n_h} (x_k - \hat{x}_h)^2} ;$$

n_h l'effectif enquêté de la strate h ;

N_h l'effectif total de la strate h.

Le gain de précision sur la moyenne de X peut être estimé en faisant le rapport des estimateurs :

$$\frac{\hat{\sigma}(\hat{x}_{st})}{\hat{\sigma}(\hat{x})} = \frac{\sqrt{\sum_{h=1}^H \left(\frac{N_h}{N}\right)^2 * \frac{N_h - n_h}{N_h} * \frac{1}{n_h} * \hat{s}_h^2}}{\sqrt{\frac{(N-n)}{N} * \frac{1}{n} * \hat{s}_x^2}} \quad (4.5)$$

Un sondage stratifié est toujours plus précis pour l'estimation de la moyenne qu'un sondage aléatoire simple (Ardilly, 1994). La construction de la stratification nécessite une connaissance préalable de la distribution de la population selon les strates (cette connaissance résulte généralement d'enquêtes antérieures : recensements principalement).

En revanche, la recherche de la stratification optimale nécessite la connaissance de l'écart-type au sein de chacune des strates. Si l'on suppose que le coût unitaire d'une enquête est identique dans chacune des strates, un plan de sondage optimal est défini par la formule suivante (Ardilly, 1994, l'auteur fournit également la formule plus générale lorsque les coûts d'enquête ne sont pas constants, ce qui est généralement le cas pour une stratification géographique) :

$$n_h = n * \frac{N_h * \hat{s}_h}{\sum_h N_h * \hat{s}_h} \quad (4.6)$$

- avec \hat{s}_h écart-type au sein de la strate h ;
- n_h l'effectif enquêté de la strate h ;
- N_h l'effectif total de la strate h ;
- n l'effectif total enquêté qui dépend du budget disponible.

Un sondage stratifié avec un effectif constant dans chacune des strates a donc peu de chance d'être optimal. De plus, un sondage stratifié à effectif constant (donc à probabilité inégale) a de forte chance d'être moins précis qu'un sondage à probabilité égale, sauf si la stratification est effectuée sur une variable explicative de la grandeur estimée (ce qui permet dans ce cas de réduire la variance au sein de chacune des strates). Dans le cas des enquêtes ménages, la stratification géographique ne permet pas forcément de réduire l'écart-type au sein de chacune des zones. De ce fait, la stratification opérée conduit généralement à une perte de précision par rapport à une stratification à probabilité uniforme. Pour les variables habituellement analysées dans les enquêtes ménages déplacements, la perte de précision par rapport à un sondage stratifié à probabilité uniforme est en général inférieure à 15 % (Raux et al., 1996 ; CETUR, 1977).

Dans la pratique, une enquête vise souvent à estimer un grand nombre d'indicateurs. Même si l'on disposait d'enquêtes permettant d'estimer d'une part la répartition de la population entre les différentes strates et d'autre part l'écart-type au sein de chacune des strates pour chacun des indicateurs (ce qui est peu probable), il y a fort à parier que la stratification optimale ne serait pas identique pour chacun des indicateurs. De plus, la répartition par strates ne s'effectue le plus souvent que sur un nombre limité de variables (car le croisement de plusieurs variables conduit rapidement à un nombre de strates très élevé relativement à la taille de l'échantillon enquêté). De ce fait, la précision sur des variables non concernées par la stratification peut s'avérer moindre que dans le cas d'un échantillon aléatoire simple (Ardilly, 1994). Enfin, si la stratification peut conduire à une précision accrue pour un estimateur relatif à l'ensemble de la population, ce n'est plus forcément vrai lorsque l'on analyse des sous-populations (un sondage stratifié à effectif constant dans chaque strate peut alors s'avérer plus pertinent si l'on veut obtenir une précision proche dans chacune des strates, comme dans le cas de l'enquête ménages déplacements). La définition d'une stratégie de sondage optimale s'avère donc un exercice particulièrement complexe (une analyse détaillée est proposée dans Ardilly, 1994 ; on pourra également trouver un exemple d'application relatif à l'enquête transports-communications de 1993-94 dans Armoogum et al., 1997).

2.4.2.3. Effet de grappe

Dans les enquêtes déplacements, la base de sondage est généralement une base ménage. Dans le cas de l'enquête ménages déplacements française, tous les individus du ménage de 5 ans et plus sont enquêtés. De même, tous les déplacements de chaque individu sur une journée sont enquêtés. Pour les analyses menées sur les individus, il faut donc prendre en compte l'effet de grappe à un niveau et pour les analyses sur les déplacements, l'effet de grappe à deux niveaux (l'effet de grappe traduit la perte de précision relative due au fait que le tirage de l'échantillon n'est pas aléatoire au niveau de l'individu ou du déplacement (il l'est seulement au niveau du ménage) comparativement à un échantillon de même taille qui serait obtenu par tirage aléatoire sur les individus ou les déplacements. L'estimation de l'effet de grappe est assez lourde (dans le cas, où la taille de chacune des grappes n'est pas identique (c'est le cas de l'enquête ménages, puisque la taille des ménages est variable et que le nombre de déplacements réalisés par chacun des individus est lui aussi variable), l'estimation de l'effet de grappe est fournie par Ardilly (1994)). Nous donnons ci-dessous une formule approchée à partir de celle fournie par (CETUR, 1977), qui est acceptable lorsque le nombre de grappes est élevé dans la population (c'est le cas dans les enquêtes ménages déplacements avec une population de ménages) :

$$\hat{\sigma}(\hat{x}) = \sqrt{\frac{1}{n_g(n_g - 1)} * \frac{(N_g - n_g)}{N_g} \sum_{k=1}^{n_g} \left(\frac{n_k}{\bar{N}}\right)^2 * (\bar{x}_k - \hat{x})^2} \quad (4.7)$$

où n_g est le nombre de grappes de l'échantillon enquêté et N_g le nombre de grappes dans la population totale ;

n_k est la taille de la grappe k appartenant à l'échantillon enquêté ;

\bar{N} est la taille moyenne d'une grappe dans l'ensemble de la population, soit $\frac{N}{N_g}$ (N taille de la population totale) qui peut être estimée sur l'échantillon enquêté par $\frac{n}{n_g}$ (n taille de population

enquêtée) ;

\bar{x}_k est la moyenne de la variable X sur la grappe k ;

\hat{x} est un estimateur sur la population enquêtée de la moyenne de la variable X ,

$$\hat{x} = \sum_{i=1}^n \frac{x_i}{n}.$$

A l'inverse de la stratification où l'on recherche une homogénéité à l'intérieur d'une strate pour minimiser la variance, dans le cas d'un sondage en grappe, on privilégiera l'hétérogénéité à l'intérieur des grappes. On peut le comprendre aisément à partir d'un exemple caricatural. Si tous les individus à l'intérieur de chacune des grappes ont une moyenne \bar{x}_k identique, on pourrait avoir une précision identique sur l'estimateur de la moyenne de X en n'enquêtant qu'un seul individu au sein de chacune des grappes. Dans ce cas, l'effet de grappe (rapport entre l'écart-type estimé sur l'échantillon issu d'un tirage en grappe et l'écart-type obtenu à partir d'un tirage aléatoire simple, avec deux populations de même taille) serait à peu près égal à la racine carrée de la taille moyenne d'une grappe. Ainsi, dans le cas des enquêtes ménages déplacements françaises, l'effet de grappe maximal au niveau des personnes est proche de la racine carrée de la taille moyenne des ménages $\cong 1,5$, et celui au niveau des déplacements est proche de 3.

On trouvera une estimation de l'effet de grappe pour différents indicateurs de mobilité dans le cas des enquêtes ménages déplacements de Lyon dans Raux et al. (1996) et de Grenoble (1973) dans CETUR (1977). L'effet de grappe est assez limité pour les variables de mobilité au niveau des individus (tout au moins pour les variables étudiées) avec une valeur comprise entre 1 et 1,3, tandis qu'il est beaucoup plus grand pour les variables relatives aux déplacements (telles que les parts modales ou par motifs) entre 1,5 et 2.

3. PRINCIPALES DONNEES DISPONIBLES EN MILIEU URBAIN EN FRANCE

La plupart des agglomérations françaises disposent de plusieurs enquêtes renseignant sur les déplacements urbains. Nous avons réuni les principales dans le tableau 4.1. Cette présentation n'est pas exhaustive, mais reprend les principales données facilement disponibles (par exemple la SNCF produit des statistiques sur les utilisateurs de son réseau, des enquêtes ponctuelles sont fréquemment menées par différents organismes locaux, des résultats peuvent être disponibles à partir d'enquêtes nationales comme l'enquête transport (INSEE, INRETS, 1997)...).

Domaine	Nom de la source	Principales données produites
Données générales	Enquêtes ménages déplacements	Ensemble des déplacements sur une journée des personnes de 5 ans et plus de ménages résidant dans l'agglomération
Réseau routier	Enquêtes cordon ou bord de route	Déplacements routiers (véhicules) en sortie d'agglomération traversant un cordon entourant l'agglomération, ou passant en un point donné de la voirie.
	Comptages	Nombre de véhicules passant en un point donné de la voirie
Réseau transport collectif	Enquêtes O-D ou montée-descente	Déplacements ou voyages sur le réseau de transports collectifs
	Oblitérations Ventes de titres Billetique	Montées avec ticket sur une ligne donnée ou données billettique Indication globale de fréquentation du réseau
Domicile-travail	Recensement INSEE (fichier Mirabel)	Localisation du domicile et du travail, et mode habituel pour se rendre au travail depuis le recensement de 1999

Tableau 4.1 : Principales enquêtes déplacements disponibles en milieu urbain en France

Nous décrivons rapidement ces enquêtes dans cette section.

3.1. Enquêtes ménages déplacements

Les enquêtes ménages déplacements sont la principale source de données en milieu urbain sur les déplacements réalisés au sein d'une agglomération. Les enquêtes réalisées dans les différentes agglomérations reposent sur une méthodologie commune élaborée par le CERTU. Ces enquêtes sont pilotées localement par les autorités organisatrices qui sont maîtres d'ouvrage. Sous réserve du respect de la méthodologie CERTU, elles sont financées à 50 % par l'Etat et 50 % localement. Dans les grandes agglomérations, elles sont le plus souvent réalisées tous les 8 à 15 ans. Le coût de l'enquête est de l'ordre de 1000F par ménage, pour l'élaboration, la préparation, la réalisation et une exploitation standard.

3.1.1. Définition de l'objectif

Le principal objectif des enquêtes ménages déplacements est de fournir une base de données aidant à la réflexion dans le domaine de la planification des déplacements urbains. Elles sont donc conçues pour disposer d'une connaissance fine de la structure des déplacements et pour permettre la recherche de leurs déterminants, tout en assurant une bonne fiabilité des résultats (CERTU, 1998a).

Depuis 1976, cette réflexion est envisagée à un double niveau, national et local.

L'exigence nationale se traduit par une méthodologie d'enquête identique dans toutes les agglomérations et, au niveau du questionnaire, par l'existence d'un tronc commun, fixe quels que soient la date et le lieu de l'enquête. Le respect de ces contraintes est contrôlé par le CERTU. Elles rendent possibles les comparaisons entre les différentes agglomérations françaises ainsi que les analyses sur l'évolution dans le temps de la mobilité. « *On peut ainsi aborder les thèmes suivants :*

- la comparaison entre elles des agglomérations. *Qu'est-ce qui ressort des « tendances lourdes », et qu'est-ce qui est la conséquence des politiques locales de transport ? Ceci est fondamental pour juger de l'efficacité d'une politique de transport ;*
- l'évolution dans le temps d'une agglomération. *L'exemple le plus illustratif est Grenoble, qui est la seule agglomération de province à avoir eu quatre enquêtes ménages (1966, 1973, 1978, 1984 [6 à ce jour avec les enquêtes de 1992 et 2001]). Les enquêtes ménages sont également tout à fait fondamentales pour juger de l'impact d'une infrastructure lourde de transport (rocade, métro, tramway...). » (CETE, CETUR, 1989)*

Les promoteurs locaux peuvent ensuite, en fonction de leurs propres centres d'intérêts, rajouter des questions ou affiner certaines modalités (sous condition de pouvoir retrouver les modalités originelles par réagrégation). Ils complètent souvent l'enquête par un questionnaire dit « complémentaire ». Les enquêtes ménages déplacements contribuent ainsi à alimenter de nombreuses études et réflexions sur le fonctionnement des agglomérations, sur leur planification. A ce niveau, l'objectif de ces enquêtes devient multiple. *« Les techniciens des transports doivent avoir le « réflexe enquête ménages », qui consiste à chercher dans les résultats de l'enquête des éléments ou arguments destinés à enrichir leurs études, dans des domaines très variés :*

- les flux de trafics routiers pour une DDE (Direction Départementale de l'Équipement) ;
- les relations entre urbanisme et mobilité pour une Agence d'Urbanisme ;
- les données sur la clientèle pour un réseau de transports collectifs ;
- les aspirations des habitants en matière de transports pour les élus ;
- les données sur les achats pour une Chambre de Commerce ;
- » (CETE, CETUR, 1989)

Ainsi, sur le plan local, l'enquête ménages déplacements de Lyon réalisée en 1986 avait pour objectif de :

- « - donner la possibilité d'analyses spécifiques dans le cadre de modélisation ;
- constituer une source de données pour tous les partenaires locaux et les diverses études qui en résultent (clientèle des transports collectifs, stationnement, perception du vécu des transports, etc.) ;
- ... ;
- mesurer et analyser les caractéristiques des ménages ;
- mesurer et analyser les caractéristiques individuelles de perception et de fréquentation de divers types d'activités et d'équipements ;
- mesurer et analyser les caractéristiques détaillées des comportements des individus, eu égard à la mobilité et au choix des modes de déplacement ;
- mesurer et analyser les « parts de marché » des différents modes de transport ;
- mesurer la valeur d'indicateurs globaux concernant les déplacements habituels des personnes, afin de permettre, entre autre, une comparaison avec les mêmes indicateurs mesurés lors des enquêtes précédentes ;
- mesurer et comparer ces mêmes indicateurs entre villes ayant réalisé une enquête ménages identique ;
- appréhender et analyser les attitudes, les opinions, les habitudes des ménages intéressés directement ou indirectement par l'usage des différents modes de transport et plus particulièrement par l'usage des transports collectifs depuis la mise en service du métro lyonnais » (CETE de Lyon, 1986).

Cette multiplicité d'objectifs confère toute leur richesse aux bases de données tirées des enquêtes ménages déplacements. Mais elle est aussi source de conflit sur le plan méthodologique, lorsque l'on souhaite remplir tous les objectifs. Elle résulte en partie du montage institutionnel nécessaire pour assurer le financement de l'enquête. Ainsi, la réalisation, au plan local, d'une enquête ménages déplacements nécessite au préalable la mise en place d'un Comité de pilotage regroupant les

financeurs potentiels, et d'un Comité technique chargé d'examiner la mise en œuvre concrète des objectifs.

3.1.2. Définition de la population cible

Compte tenu des objectifs de planification assignés à l'enquête ménages déplacements, la population cible est constituée de l'ensemble des déplacements réalisés au sein du périmètre d'enquête d'une part et d'autre part de l'ensemble des déplacements réalisés par les résidents du même périmètre. De ce fait, nous avons deux populations cibles différentes selon les objectifs :

- population de déplacements réalisés au sein d'un périmètre ;
- population d'individus (voire de ménages) résidant au sein d'un périmètre.

Ces deux cibles se recoupent partiellement, puisqu'une partie des déplacements réalisés au sein du périmètre d'étude le sont par les résidents de ce même périmètre. Toutefois, une partie des déplacements sont réalisés par des non-résidents. Le choix a été fait de privilégier la cible individus pour des raisons évidentes d'accès à des bases de sondage d'une part et pour analyser les comportements individuels d'autre part. De ce fait, au sein du périmètre d'étude, une partie des déplacements n'est pas couverte par l'enquête ménages déplacements :

- déplacements de transit ;
- déplacements d'échange entre le périmètre couvert par l'enquête et l'extérieur de ce périmètre réalisés par des non-résidents de l'agglomération ;
- déplacements professionnels non recueillis durant l'enquête, ce qui biaise l'image des trafics routiers en évacuant les déplacements de ceux dont le métier consiste à se déplacer tels que les chauffeurs de taxis, les livreurs, les représentants... ;
- déplacements réalisés au sein de l'agglomération par les personnes résidant temporairement au sein de l'agglomération (hôtels, lieux de séjour...) ;
- déplacements de marchandises.

C'est pour couvrir ces biais que d'autres enquêtes sont parfois réalisées. C'est le cas des enquêtes cordon qui couvrent les déplacements d'échanges et de transit (section 3.2, p.121). Dans certaines agglomérations, des enquêtes sont réalisées aux lieux de séjour. Enfin, plus récemment des enquêtes ont été lancées pour repérer les déplacements de marchandises en ville (Ambrosini et al., 1999a, b ; Ségalou et al., 2001).

Le périmètre choisi dépend des agglomérations. Il a tendance à s'élargir au fil du temps pour suivre l'élargissement du bassin d'attraction de l'agglomération. Actuellement, il est dans la plupart des cas plus large que le Périmètre des Transports Urbains (PTU), au sein duquel l'Autorité Organisatrice des transports collectifs assure sa compétence. En revanche, dans le passé, ce périmètre était souvent plus restreint que le bassin d'attraction de l'agglomération.

3.1.3. Définition de la méthode d'échantillonnage

A ce jour, la base de sondage utilisée est fréquemment constituée du dernier recensement généralement mis à jour à partir de fichiers des constructions neuves et des habitations détruites. L'unité de base statistique est le logement, donc le ménage qui y réside, c'est-à-dire tous les occupants réguliers du logement, qu'il y ait ou non des liens matrimoniaux ou familiaux.

L'échantillon est constitué par tirage aléatoire de ménages après stratification géographique du périmètre d'enquête. Cette stratification s'opère de manière à avoir un échantillon d'au minimum 75 ménages au sein de chacune des zones de tirage. Le taux de sondage est ainsi différent pour chaque zone. Cet échantillonnage permet de disposer d'une population représentative, au niveau de ces zones dans un premier temps, et au niveau de l'agglomération ensuite.

Tous les individus du ménage de 5 ans et plus sont enquêtés le même jour et l'ensemble de leurs déplacements réalisés la veille du jour de l'enquête est recueilli. Il faut donc tenir compte de l'effet de grappe à un niveau pour les analyses sur les individus et à deux niveaux sur les déplacements (section 2.4.2, p. 112). L'enquête ne porte que sur les jours de semaine.

Malgré l'importance de ces enquêtes, les tailles d'échantillon restent limitées pour d'évidentes raisons de coût. A titre d'illustration, à Lyon, 3 700 ménages ont été enquêtés pour l'enquête de 1976, 4 700 en 1985 et 6 000 en 1995. Dans des agglomérations de plus petite taille, l'effectif est évidemment plus réduit. Le taux de sondage oscille en général autour de 1 %.

La précision des données doit donc être systématiquement prise en compte lors de toute analyse de ces données. Cette précision peut être estimée à partir des formules fournies section 2.4.2 (p. 112). Il faut toutefois tenir compte de la perte de précision relative due à l'effet de grappe à un niveau pour les données sur les personnes et à deux niveaux pour les données relatives aux déplacements. Des exemples de calcul pour un certain nombre de variables sont fournis dans le cas lyonnais par Raux et al. (1996).

3.1.4. Définition du mode d'enquête

Les enquêtes sont réalisées en face à face au domicile des enquêtés par des enquêteurs préalablement formés.

3.1.5. Rédaction du questionnaire

Le tronc commun du questionnaire standard CERTU assurant la comparabilité des enquêtes ménages déplacements entre elles est constitué de trois parties :

- un questionnaire « ménage » qui décrit les caractéristiques du ménage. Il comprend 3 grandes rubriques, composition du ménage, motorisation et divers ;
- un questionnaire « personne » qui comprend des questions sur les déterminants sociologiques de la personne, sur les déplacements domicile-travail ou école, le stationnement ainsi que des questions « d'habitude » d'usage des modes de transport ;
- un questionnaire « déplacements » qui comprend pour chaque personne, la description détaillée de tous ses déplacements de la veille : type de trajet, zone origine et de destination, motif du déplacement à l'origine et à la destination plus éventuellement motif de déplacement des personnes accompagnées, heure de départ et d'arrivée, durée du déplacement, mode de transport et titre de transport s'il s'agit d'un déplacement effectué en transports collectifs....

Lorsqu'un questionnaire complémentaire est ajouté, il est le résultat d'un compromis entre les différents partenaires ayant participé au groupe de travail chargé d'élaborer l'enquête. A des questions sur des sujets d'intérêts locaux s'ajoutent des questions d'opinion, d'attitude, de satisfaction ou des questions ouvertes. En général, une personne de plus de 16 ans du ménage est tirée au sort pour être interrogée.

Compte tenu de la longueur du questionnaire et du fait que toutes les personnes de plus de 5 ans sont enquêtées, l'enquête dure en moyenne une heure et demi par ménage.

3.1.6. Organisation du terrain

Le terrain est généralement assuré sous la direction du réseau des CETE en partenariat fréquent avec l'INSEE.

3.1.7. Codage, saisie des données

Lors de l'élaboration des questionnaires, le maximum de compatibilité est recherché entre le découpage de l'enquête et les autres découpages existants (divers découpages d'études, limites administratives, découpages INSEE...). Toutefois, compte tenu de la multiplicité des découpages et de leurs différences d'objectifs, la compatibilité ne peut être assurée pleinement.

Plusieurs découpages sont utilisés pour l'enquête. Ils s'emboîtent les uns dans les autres. Parmi ceux-ci, deux méritent d'être mentionnés :

- découpage de tirage de l'échantillon. Le nombre de zones dépend de la taille de l'échantillon et de l'effectif retenu par zone en moyenne (autour de 75 ménages en général) ;
- découpage de codage des origines et destinations des déplacements et de la localisation du logement et du travail. Il est généralement beaucoup plus fin que le découpage correspondant à l'échantillonnage. Sur ce découpage, la représentativité n'est pas recherchée à l'échelle de chaque zone.

Le codage est en grande partie réalisé pendant l'enquête. La saisie est, en revanche, réalisée après le terrain.

Chaque enquête donne lieu à plusieurs publications comprenant le document technique de présentation de l'enquête (par exemple pour la dernière enquête de Lyon : (CETE de Lyon et al., 1995)) et des rapports présentant les principaux résultats.

3.2. Enquête Cordon et bord de route

Cette enquête s'intéresse aux déplacements routiers qui sortent de l'agglomération et donne ainsi une image des flux d'échange et de transit. Elle est conduite par les Directions Départementales de l'Équipement dans le cadre de leur compétence sur les voiries nationales. Elle est financée par le Ministère de l'Équipement et par les collectivités locales, car elle est également réalisée sur les voiries de compétence locale pénétrant dans l'agglomération (pour Lyon, CETE de Lyon et al., 1991).

Ces dernières années, les enquêtes cordon et bord de route se sont faites rares du fait du refus de coopération de la part des forces de l'ordre, dont la présence est obligatoire pour arrêter les véhicules.

3.2.1. Définition de l'objectif

L'objectif principal de cette enquête est de contribuer à la réflexion sur les grands investissements routiers autour des agglomérations en fournissant des données sur le trafic sortant dans l'agglomération. Le trafic sortant est privilégié par rapport au trafic entrant afin d'enquêter sur les caractéristiques du déplacement qui vient d'être réalisé, plutôt que sur les caractéristiques prévues pour le déplacement. On obtient ainsi les trafics de transit et d'échanges sur les principales voies d'accès dans l'agglomération.

Réalisée tous les 10 ans environ, elle permet de réactualiser périodiquement les connaissances sur les déplacements routiers. Elle complète les comptages effectués sur le réseau routier et fournit les origines, les destinations et les motifs des déplacements enquêtés. Ces données servent également à compléter les informations fournies par l'enquête ménages déplacements pour caler les modèles de prévision de la demande de déplacements. Ces données sont notamment utilisées dans le cadre des réflexions sur les PDU et les DVA (CETUR, 1993a).

3.2.2. Définition de la population cible

La population cible est constituée des véhicules traversant un cordon situé aux portes de l'agglomération. Seuls les véhicules sortant de l'agglomération sont enquêtés. La localisation de ce cordon dépend de l'agglomération et des possibilités physiques du terrain pour réaliser l'enquête (en général barrière de péage pour les autoroutes).

3.2.3. Définition de la méthode d'échantillonnage

Il n'existe bien évidemment pas de base de sondage exhaustive des déplacements figurant dans la population cible. L'échantillonnage consiste à enquêter toutes les voies permettant de pénétrer dans l'agglomération. L'enquête est en général exhaustive en heure creuse. Par contre, seule une partie des véhicules est enquêtée en heure de pointe. Dans ce cas, un comptage permet le redressement. Les enquêtes se déroulent sur un jour de semaine moyen (en général mardi, mercredi, jeudi).

Le zonage est défini au niveau des communes (arrondissements pour les grandes villes) pour la France et par pays pour l'étranger.

3.2.4. Définition du mode d'enquête

Les données sont recueillies par questionnaire administré aux chauffeurs des véhicules en arrêtant ceux-ci au bord de la chaussée au niveau du cordon.

3.2.5. Rédaction du questionnaire

Les questionnaires sont différents selon le type de véhicules concernés (véhicule léger, poids lourd, car). Compte tenu du mode d'enquête sur voirie, ils sont nécessairement courts. A Lyon en 1990, ils comportaient les renseignements suivants :

- date, heure et lieu d'enquête ;
 - lieu d'immatriculation (département ou pays) ;
 - lieu de résidence ;
 - origine, destination ;
 - motif à l'origine et à la destination ;
 - fréquence du déplacement ;
 - nombre d'occupants du véhicule ;
 - itinéraire dans Lyon (par repérage d'axes ou de nœuds importants) + axe de pénétration dans Lyon pour le transit ;
 - identification de points dangereux et évaluation de la signalisation ;
- et pour les poids lourds :
- nature, poids... des marchandises transportées.

3.2.6. Organisation du terrain

A Lyon, en 1990, les enquêtes se sont déroulées les mardis, mercredis ou jeudis, de 6h00 à 20h00 (et sur 24h00 pour les poids lourds sur autoroute). 45.000 conducteurs ont été interrogés en mai et juin. Le cordon comportait 36 postes d'enquêtes situés en limite d'agglomération, en général un peu au-delà des frontières du Grand Lyon. Pour les autoroutes, les postes d'enquêtes étaient situés aux barrières de péage. La DDE et le CETE estiment que 90 % du trafic sortant de l'agglomération passe par ces postes.

On peut également constituer un sous échantillon de conducteurs acceptant de répondre à un questionnaire complémentaire par téléphone.

Cette enquête est réalisée à peu près tous les 10 ans. A Lyon, les trois dernières l'ont été en 1990, 1979 et 1965.

3.2.7. Codage, saisie des données

Les adresses sont codées à la commune et à l'arrondissement pour les plus grandes villes.

3.3. Comptages routiers

Les comptages routiers permettent de recenser le nombre de véhicules passant dans un sens et dans l'autre à un endroit donné du réseau routier.

En général, plusieurs services effectuent des comptages sur les voiries du territoire de l'agglomération. Sur le territoire du Grand Lyon :

- la Cellule Départementale d'Exploitation et de Sécurité de la DDE du Rhône ;
- la Ville de Lyon ;
- la communauté urbaine de Lyon.

Ces trois organismes en assurent le financement et la maîtrise complète.

3.3.1. Définition de l'objectif

L'objectif des comptages peut être assez variable. Les comptages ponctuels répondent le plus souvent à des besoins d'études ponctuels. Par contre, les comptages permanents répondent à la fois à des besoins d'exploitation et de planification.

3.3.2. Définition de la population cible

C'est la population des véhicules franchissant le poste de comptage. C'est une population de véhicules et non de déplacements. L'unité statistique est donc différente de celle de l'enquête ménages déplacements. Elle est par contre similaire de celle de l'enquête cordon (qui recueille en plus le nombre de personnes dans le véhicule).

3.3.3. Définition de la méthode d'échantillonnage

C'est un recueil exhaustif (sous réserve de la fiabilité et du bon fonctionnement du capteur). Certains capteurs permettent de séparer véhicules légers et lourds et de saisir la vitesse du véhicule. Il est en général possible de conserver l'information par unité de temps assez courte.

C'est une information particulièrement utile dans le calage d'un modèle. Des comptages judicieusement placés permettent d'intervenir sur le calage de la matrice origines-destinations et du réseau. Ils sont parfois également utilisés pour construire une matrice origines-destinations lorsqu'aucune autre enquête n'est disponible (dans de petites agglomérations ou fréquemment dans les agglomérations de pays en développement) ou pour actualiser une ancienne matrice.

Ainsi si l'information est très pauvre (pas d'information sur l'origine ou la destination du véhicule, sur le nombre d'occupants, sur les déplacements correspondants...), elle est néanmoins indispensable dans le calage d'un modèle.

3.4. Bases de données transports collectifs

Les méthodologies étant assez variables selon les réseaux de transports collectifs, nous ne les détaillerons pas. La plupart des réseaux réalisent des enquêtes auprès des usagers. Elles sont principalement de trois types :

- enquête montée-descente. Elle permet de reconstituer le diagramme de charge de la ligne et de connaître le profil de la clientèle sur la ligne ;
- enquête origines-destinations. Elle permet de connaître l'utilisation du réseau. L'information recueillie ne se limite plus à l'arrêt de montée et de descente sur la ligne, mais reprend l'ensemble de l'itinéraire du déplacement de son origine à sa destination. Quelques questions sur l'individu permettent de connaître le profil de la clientèle ;
- des enquêtes plus ponctuelles (enquête qualité de service notamment).

A ces enquêtes s'ajoutent un certain nombre de statistiques disponibles sur les ventes de titres ou sur les oblitérations. La billettique, lorsqu'elle est installée, permet évidemment une information beaucoup plus riche.

Les enquêtes origines-destinations ne sont généralement pas réalisées à la même date sur l'ensemble du réseau, mais de manière plus ou moins périodique ou en fonction des besoins d'étude ou de suivi des lignes après restructuration de service. De ce fait, elles nécessitent un traitement particulier pour reconstituer une matrice origines-destinations. Elles bénéficient généralement de taux de sondage important, voire exhaustif (sauf pour les lignes de métro qui nécessitent en général des dispositifs d'échantillonnage beaucoup plus complexes à mettre en œuvre). Elles permettent donc de reconstituer des matrices avec une précision élevée.

3.5. Recensements de l'INSEE

Périodiquement, le recensement général de la population est effectué au plan national. Ces enquêtes réalisées pour les dernières en 1968, 1975, 1982, 1990 et 1999 sont les seules enquêtes exhaustives (ou plus exactement quasi-exhaustives) dont nous disposons, concernant la population française. Si ces enquêtes ne portent pas à proprement parler sur le problème du transport, elles offrent toutefois des informations sur la localisation du domicile et du travail, et sur le mode habituel de déplacements pour le recensement de 1999.

Ces informations sont mises en forme dans le fichier Mirabel de l'INSEE.

La redistribution des questionnaires, établis relativement au lieu de domicile, selon le lieu de travail offre deux lectures possibles des migrations alternantes :

Selon la commune

Pour chaque commune, on dispose des données suivantes :

- **Résidents** : nombre d'actifs résidant dans la commune ;
- **Stables** : nombre d'actifs résidant et travaillant dans la commune ;
- **Sorties** : nombre d'actifs, résidant mais ne travaillant pas dans la commune ;
- **Emplois** : nombre d'emplois de la commune ;
- **Entrées** : nombre d'actifs travaillant mais ne résidant pas dans la commune ;
- **Solde** : différence entre les entrées et les sorties.

Plusieurs découpages géographiques sont proposés, notamment la commune (ou l'arrondissement pour les plus grandes communes), le canton et le département.

Selon les origines-destinations ou échanges pour le motif domicile-travail

La seconde lecture des questionnaires permet de retracer l'ensemble des échanges entre les communes. Pour chaque commune de résidence est présenté l'ensemble des destinations : la commune de travail des individus concernés et le nombre d'actifs de la commune de résidence travaillant dans la commune dite de travail. Malgré le terme « d'échanges » utilisé, ce fichier ne renseigne pas sur les déplacements domicile-travail, mais seulement sur la localisation du domicile et du lieu de travail habituel. La fréquence des déplacements domicile-travail est inconnue. On utilise fréquemment un ratio proche de 5 allers-retours par semaine pour générer le nombre de déplacements, mais celui-ci dépend bien évidemment des rythmes professionnels, de la distance entre le domicile et le travail...

Le dernier recensement inclut également une question sur le mode de transport habituel pour les déplacements domicile-travail.

3.6. Conclusion

Ces enquêtes fournissent donc une image assez large des déplacements réalisés au sein d'un périmètre donné. Elles constituent une base particulièrement utile pour alimenter les modèles de prévision de la demande de déplacements.

Toutefois, elles ne sont pas exemptes d'un certain nombre de limites dans une perspective de modélisation que nous développons dans la section suivante.

4. LIMITES DU DISPOSITIF FRANÇAIS D'ENQUÊTES SUR LES DÉPLACEMENTS URBAINS POUR LA MODÉLISATION

Le dispositif français d'enquêtes alimente la plupart des études de modélisation. Il permet également de développer de nombreuses études et recherches sur la mobilité urbaine (notamment de nombreuses recherches menées au LET (Laboratoire d'Economie des Transports) s'appuient sur ces données). La similitude de la méthodologie dans le temps permet des analyses en évolution intéressante. Toutefois, malgré l'intérêt de ce dispositif, il n'est pas exempt, selon nous, d'un certain nombre de critiques dans une perspective de modélisation :

- une partie des déplacements est exclue du dispositif d'enquête (section 4.1) ;
- des périmètres d'enquêtes parfois différents (section 4.2) ;
- des dates d'enquêtes parfois très différentes (section 4.3) ;
- des découpages différents et un codage trop grossier des localisations (section 4.4) ;
- une taille d'échantillon trop faible pour l'enquête ménages déplacements (section 4.5).

Cette dernière remarque, nous conduit à formuler une proposition pour alléger l'enquête ménages déplacements et changer de mode d'enquête afin d'accroître sensiblement la taille de l'échantillon pour un coût comparable (section 4.6.).

4.1. Des déplacements exclus du dispositif d'enquête

L'enquête ménages déplacements couvre les déplacements réalisés par les résidents du périmètre d'enquête. De ce fait, une partie des déplacements réalisés au sein de ce périmètre ne sont pas pris en compte par cette enquête (section 3.1.2, p. 119).

L'enquête cordon permet de compléter ce recueil en repérant les déplacements d'échanges et de transit. Par contre, les autres segments restent exclus du dispositif. L'utilisation des autres sources de données est souvent utile pour essayer de les atteindre. C'est notamment le cas, des comptages routiers pour les flux automobiles et des enquêtes réalisées par les réseaux de transports collectifs pour ce mode. Enfin, certaines villes, ont réalisées des enquêtes pour repérer les déplacements de marchandises en ville (Ambrosini et al., 1999a, b).

L'ensemble de ces données a été analysé dans le cadre de la conception du modèle télescopage (Cabanne et al., 2000). Sur cette base, Nicolas et al. (2001) estiment qu'en 1996, seuls 57 % des véhicules kilomètres parcourus un jour moyen de semaine à l'intérieur du périmètre de l'enquête ménages déplacements sont repérés par cette enquête. Ces chiffres sont cohérents avec ceux issus d'autres sources et notamment à partir d'une application du logiciel DAVIS par le CETE de Lyon qui fournit une estimation de 49 % des véhicules kilomètre et de 87 % des déplacements de véhicules repérés en heure de pointe par l'enquête ménages déplacements de Lyon en 1996 (Nicolas et al., 2001).

4.2. Des périmètres d'enquêtes différents

Dans le passé, les périmètres utilisés par l'enquête ménages déplacements et l'enquête cordon différaient fréquemment (en principe aujourd'hui une plus grande coordination est recherchée entre ces deux enquêtes). L'absence de correspondance entre les périmètres se traduit par deux situations aux conséquences fort différentes :

- une commune se retrouve à l'intérieur du périmètre de l'enquête ménages déplacements et à l'extérieur du cordon. Cette commune fait alors partie de la population cible des deux enquêtes. Ce premier cas n'est pas trop gênant, car il est possible d'en tenir compte lors de l'analyse des données et d'exclure les résidents d'une des deux enquêtes ;
- une commune se retrouve à l'extérieur du périmètre de l'enquête ménages déplacements et à l'intérieur du cordon. Cette commune est alors exclue de la population cible des deux enquêtes. Ce second cas est beaucoup plus gênant, car les résidents de la commune échappent aux deux recueils.

Au-delà du problème spécifique de compatibilité des périmètres de ces deux enquêtes, le périmètre de l'enquête-ménages reste encore parfois trop restreint par rapport au bassin d'attraction de l'agglomération. L'étalement urbain et le desserrement des activités conduit à un élargissement de la zone d'attraction de l'agglomération. Les zones aux franges des découpages administratifs (notamment des communautés urbaines) connaissent des développements souvent importants et ne sont pas toujours inclus dans le périmètre de l'enquête ménages déplacements. Elles contribuent pourtant de manière importante aux phénomènes de congestion grandissant aux entrées d'agglomérations (Cabanne et al., 2000).

4.3. Des dates d'enquêtes très différentes

Les enquêtes étudiées ont le plus souvent été réalisées à des dates très différentes. A titre d'exemple, à Lyon :

- enquêtes ménages déplacements : 1976, 1985, 1995 ;
- enquêtes cordon : 1965, 1979, 1990 ;
- recensements : 1968, 1975, 1982, 1990, 1999 ;
- enquêtes transports collectifs : périodicité très variable selon les lignes.

Seuls les comptages « permanents » restent compatibles avec ces enquêtes, car ils sont... permanents.

Pourtant, il est fréquent de devoir combiner ces différentes sources, ne serait-ce que pour construire ou actualiser une matrice origines-destinations. Il est alors nécessaire de produire des interpolations ou des extrapolations de ces données pour construire la matrice de référence. Ces manipulations ne sont bien évidemment pas exemptes de sources d'erreurs potentielles, compte tenu de la multitude des facteurs pouvant influencer sur l'évolution des données correspondantes.

La périodicité des recensements et la non concordance des enquêtes déplacements avec ceux-ci aggravent encore le problème. Les données du recensement sont généralement utilisées pour assurer le redressement de l'échantillon, nécessitant là encore des interpolations ou des extrapolations de certaines données, notamment de population. La modification de la procédure d'enquête pour le recensement qui va devenir permanent devrait résoudre en partie le problème (ou plus exactement poser le problème de manière différente). De plus, aujourd'hui, une plus grande coordination entre les enquêtes ménages déplacements et les enquêtes cordon est assez souvent recherchée. Nous ne pouvons qu'espérer que cette tendance se généralise afin de pouvoir combiner plus facilement ces sources de données très importantes pour la modélisation.

4.4. Des découpages différents et un codage trop grossier des localisations

Les découpages utilisés pour la codification des adresses sont le plus souvent spécifiques à chacune des enquêtes. En général le plus petit dénominateur commun est, dans le meilleur des cas, la commune ou l'arrondissement pour les plus grandes villes, ce qui n'est généralement pas suffisant pour les besoins de modélisation.

L'enquête ménages utilise généralement un découpage différent de celui utilisé pour les enquêtes de transports collectifs, car elles répondent pour la première à une logique de connaissance de la mobilité tous modes et la seconde à une logique d'offre de transports collectifs. Ainsi à Lyon, les découpages de l'enquête ménages déplacements ne sont pas compatibles avec le découpage utilisé pour la construction de la matrice des déplacements en transports collectifs (matrice produite par la SEMALY à partir des enquêtes origines-destinations).

L'enquête cordon, les données d'emplois, ainsi que nombre d'informations relatives aux projections d'urbanisme ne sont disponibles qu'à un niveau communal (ou d'arrondissement pour les plus grandes villes).

Enfin, si les données du recensement sont saisies à un niveau important de désagrégation spatiale, leur exploitation à ce niveau n'est pas toujours possible (les règles de la CNIL (Commission Nationale Informatique et Liberté) sont très strictes concernant la diffusion de l'information à un niveau très fin de désagrégation spatiale).

Cette question de non-compatibilité entre les découpages pourrait être résolue en adoptant un principe de codification en coordonnées x-y. Si ce principe pose d'évidentes difficultés de mises en œuvre, celles-ci ne sont pas insurmontables si l'on en croit l'expérience internationale (Chapleau, 1991, 1993, 1997 ; Chapleau et al., 1999 ; Viégas, 1997) et française dans le domaine des enquêtes d'audience publicitaire (enquêtes réalisées avec le concours du Centre d'Etude des Supports de Publicité dans plusieurs agglomérations françaises). Au-delà des questions de faisabilité sur lesquelles nous ne souhaitons pas nous étendre ici (section 4.6.5, p. 132), cette codification en coordonnées x-y offre plusieurs avantages :

- compatibilité avec tous les découpages quels qu'ils soient ;
- analyse fine de l'accès aux réseaux. Elle permet de connaître de manière précise les conditions d'accès aux différents réseaux. Cette information est particulièrement utile pour les transports collectifs. On sait que la zone de chalandise d'un arrêt de bus est assez limitée. En général, on considère que l'essentiel de la clientèle est captée dans un rayon de moins de 400 mètres autour de l'arrêt. L'utilisation d'un zonage comportant des zones relativement étendues posera donc d'évidents problèmes de détermination des conditions d'accès réel aux transports collectifs pour chacun des individus enquêtés, si des questions spécifiques supplémentaires ne sont pas posées durant l'enquête pour recueillir cette information. La connaissance de cette information offrirait de multiples applications, notamment dans la détermination des lois de partage modal ;

Variable	données d'enquête	coordonnées x-y (*)	zonage (*)
% d'itinéraires égaux (ligne(s) et nœud(s) identiques)		72,4	35,3
% d'itinéraires communs (ligne(s) identique(s), mais au moins 1 nœud différent)		7,6	4,3
% d'itinéraires différents (au moins un nœud et une ligne différents)		20,0	60,4
nombre de lignes empruntées	1,92	1,88	2,06
nombre d'usagers sur le mode métro	161 000	163 000	161 000
proportion des usagers faisant une correspondance métro-métro	0,314	0,262	0,248

Tableau 4.2 : Comparaison des résultats d'affectation à partir d'adresses en coordonnées x-y ou par zone (Chapleau, De Cea, 1983)

(*) Meilleur résultat obtenu selon différentes fonctions d'impédance

- gain de précision dans la codification des réseaux et l'affectation de la demande. Le principe du zonage conduit à considérer que l'ensemble de la demande se trouve concentrée dans le centroïde de la zone (principe commun aux modèles agrégés et désagrégés). Ce principe pose des problèmes lors de la représentation des réseaux, principalement à deux niveaux : codification des liens d'accès aux réseaux (appelés connecteurs de centroïde) et description de l'offre (ces problèmes sont développés dans le chapitre 10). Chapleau et De Cea (Chapleau, De Cea, 1983) ont ainsi montré sur Montréal avec un découpage en 1284 zones (pour le très grand Montréal) que la qualité de reconstitution des itinéraires empruntés par les individus est de piètre qualité à un niveau individuel, même si les charges des principales lignes sont correctement reproduites à un niveau agrégé. En revanche, un traitement complètement désagrégé à partir des informations de localisation en coordonnées x-y permet une bien meilleure qualité de reconstitution sur le plan individuel (tableau 4.2). Certes les résultats seraient probablement meilleurs sur un réseau moins maillé que celui de Montréal (il comporte 1000 nœuds de correspondance dans cette étude). Ils montrent toutefois clairement les limites du zonage. Les résultats globaux sont corrects malgré une qualité de reconstitution très médiocre au niveau du déplacement. Ce résultat ne peut être obtenu que grâce à des compensations d'erreurs. Mais en projection, les erreurs peuvent-elles continuer à s'annuler ?
- des économies de recueil de données. A Montréal et Toronto (Chapleau, 1997, Data Management Group, 1992), le coût du recueil par téléphone des localisations en coordonnées x-y a été similaire à celui de la codification zonale pour la première enquête. En revanche, l'économie est substantielle les enquêtes suivantes ;
- des données disponibles en temps réels, ce qui permet de développer des programmes assez sophistiqués de contrôle de cohérence en temps réels (Bonnel, 1999). Ces contrôles permettent de produire des données de localisation très précises avec un risque d'erreur très limité.

4.5. Une taille d'échantillon trop faible pour l'enquête ménages déplacements

La taille de l'échantillon est généralement trop faible dès lors que l'on souhaite désagréger les données selon certaines dimensions. Nous reprenons certains objectifs de l'enquête ménages déplacements pour illustrer le propos. L'enquête ménages déplacements est conçue notamment pour mesurer des évolutions dans le temps ou entre agglomérations.

Pour mesurer de telles évolutions, il faut que les intervalles de confiance calculés pour chacune des enquêtes soient disjoints. On peut alors affirmer qu'il existe une différence au seuil de précision

retenu (en général 5 %). La précision sera parfois suffisante pour des données relatives à l'ensemble de l'agglomération. En revanche, dès lors que l'on souhaite segmenter les données pour analyser l'influence de certains facteurs, la précision décroît assez rapidement. Nous donnons un exemple sur le cas de l'agglomération lyonnaise qui présente pourtant un effectif important comparativement à la plupart des agglomérations françaises (Raux et al., 1996).

Les auteurs montrent que la précision sur les temps de déplacements déclarés par les enquêtés en transports collectifs entre zones pour un découpage en 5 zones de l'agglomération (sur le périmètre de l'enquête 1976) ne permet de mettre en évidence une évolution que pour 6 flux sur les 25 alors que la différence globale est assez importante (37 minutes en 1976 avec une précision de 2 %, et 32 minutes en moyenne en 1996 avec une précision de 2 %). Le gain de temps en transports collectifs est pourtant probablement général sur l'ensemble de l'agglomération, compte tenu de l'amélioration du système des transports collectifs au cours de ces 20 années. On peut ainsi observer une diminution quasi systématique des temps de parcours en transports collectifs pour chacune des origines-destinations, mais ces différences ne sont pas statistiquement significatives. Ce constat résulte d'intervalles de confiance trop larges pour chacune des enquêtes (supérieurs à 10 % pour 18 des 25 flux en 1976 et pour 14 des 25 flux en 1995). Il faut ainsi des écarts d'au moins 25 % pour la moitié des flux pour que les résultats soient statistiquement significatifs au seuil de confiance de 5 %. Le découpage n'est pourtant pas particulièrement fin. On imagine aisément qu'avec un découpage légèrement plus fin, il n'y a plus aucun flux pour lequel le résultat puisse être significatif.

Pour des calculs de pourcentage, la précision est également bien souvent insuffisante. Ainsi lorsqu'il s'agit de calculer un pourcentage à partir d'un échantillon de 100 observations (Raux et al., 1996), le demi-intervalle de confiance est de 10 % pour un pourcentage de 50 %, soit un intervalle [40 % - 60 %]. Même avec un échantillon de 1.000 observations le demi-intervalle est encore de 3 %, toujours pour une proportion de 50 %. L'un des objectifs de l'enquête ménages déplacements est pouvoir mesurer l'incidence d'une modification importante de l'offre de transports collectifs. Prenons l'exemple de la mesure de l'incidence de la création d'un axe lourd de transports collectifs en site propre. La part de marché des transports collectifs (sur l'ensemble voiture + transports collectifs) est fréquemment de l'ordre de 20 %. Si le nombre de déplacements enquêtée est de l'ordre de 1.000 unités, l'intervalle de confiance (avec une précision de 5 %) est [17,5 % - 22,5 %]. Même avec 10.000 déplacements (soit de l'ordre de 2.000 ménages enquêtés), l'intervalle de confiance est encore de [19,2 % - 20,8 %]. Les intervalles des deux enquêtes s'additionnant, il faut donc que la part de marché des transports collectifs augmente de 5 % (soit 25 %) si l'on a 1.000 déplacements, et de 1,5 % (soit 21,5 %) si l'on a 10.000 déplacements. Lorsque l'on connaît la modération de la croissance de la part de marché des transports collectifs (par exemple à Lyon, sur le périmètre 1976, la part de marché des transports collectifs (sur l'ensemble voiture + transports collectifs) est de 22 % en 1976, 24 % en 1985 et de 21 % en 1995), la probabilité est forte pour que les écarts ne puissent être au mieux que mesurés sur l'ensemble de l'agglomération. En revanche, toute segmentation qu'elle porte sur des découpages géographiques ou des caractéristiques socio-économiques conduira probablement à des résultats statistiquement non significatifs.

Un accroissement de la taille de l'échantillon permettrait de mieux répondre à certains objectifs assignés à l'enquête ménages déplacements.

4.6. Proposition pour une enquête ménages déplacements téléphonique « allégée »

Notre présentation reste sommaire, car ce n'est pas l'objet principal de ce travail. Elle s'appuie sur des travaux menés au sein du LET (Bonnel et al., 1994b, 1994c ; Bonnel, Le Nir, 1998 ; Bonnel, 2001). Nous examinons tour à tour les objectifs (section 4.6.1), le questionnaire (section 4.6.2), le mode d'enquête (section 4.6.3), le périmètre d'enquête (section 4.6.4) et enfin la codification des adresses (section 4.6.5). Mais il est évident qu'avant tout changement, la méthodologie devrait être testée, ne serait-ce que pour s'assurer de la comparabilité des données afin de conserver des

possibilités d'analyse temporelles des différentes enquêtes ménages déplacements réalisées dans le temps (section 4.6.6).

4.6.1. Objectifs : un choix nécessaire

Les objectifs de l'enquête ménages déplacements sont multiples, en raison du nombre de partenaires associés à la définition de cette enquête et surtout à son financement. Toutefois, l'état de l'art en matière d'enquête enseigne qu'il n'est pas souhaitable de mélanger plusieurs objectifs dans une même enquête lorsque les méthodologies les plus adaptées pour répondre à chacun des objectifs divergent trop nettement. C'est pourtant le constat que nous apportons sur les objectifs actuels de l'enquête ménages déplacements. De plus, nous avons pu montrer dans la section précédente que la faiblesse des effectifs ne permettait pas de répondre pleinement à certains des objectifs.

Il nous semble donc nécessaire de réduire le nombre des objectifs assignés à l'enquête. D'autres enquêtes plus légères et donc moins coûteuses sont probablement suffisantes pour traiter certains des objectifs assignés à l'enquête ménages déplacements. L'enquête d'opinion en est probablement l'exemple le plus significatif. Une enquête par sondage auprès d'un échantillon réduit d'un millier de personnes serait sûrement bien suffisante. La pratique courante des instituts de sondage de ce type d'enquête permet d'obtenir des prix relativement modérés. Nous sommes toutefois conscients que la mise en œuvre de telles évolutions risque de se heurter à la multiplicité des objectifs des différents financeurs et des membres du groupe de travail généralement constitué localement à l'occasion de la réalisation d'une enquête ménages déplacements dans une agglomération.

L'analyse des expériences étrangères met d'ailleurs en évidence le faible nombre d'objectifs assignés aux enquêtes sur les déplacements urbains (Bonnel et al., 1994b ; Chapleau, 1997 ; Viégas, 1997 ; Stopher, 1998...).

L'abandon de certains des objectifs nous semble donc nécessaire. L'enquête ménages déplacements devrait se concentrer sur les objectifs principaux sur le plan local qui concernent la planification des déplacements urbains et donc principalement l'alimentation des modèles de prévision de la demande de déplacements. Le respect de cet objectif local ne serait d'ailleurs pas en contradiction avec les objectifs affichés sur le plan national actuellement par le CERTU. Au contraire l'accroissement de l'effectif permis par la simplification du questionnaire renforcerait la significativité statistique des données, même si la comparaison sur le plan temporel et spatial serait évidemment réduite aux seules données recueillies dans le nouveau questionnaire.

4.6.2. Un questionnaire réduit

La diminution du nombre d'objectifs permettrait de réduire fortement la taille du questionnaire et donc la durée de l'enquête. Le choix de l'alimentation des modèles de prévision de la demande de déplacements pour appuyer une réflexion sur la planification des déplacements permettrait de limiter le questionnaire aux données nécessaires à cette alimentation et à l'analyse de la représentativité statistique des données. Il en résulterait un questionnaire réduit à quelques renseignements socio-économiques sur le ménage et les individus le composant et au recueil des déplacements de la veille avec leurs caractéristiques.

La durée de passation d'un tel questionnaire serait fortement allégée permettant d'envisager le recours à d'autres modes d'enquête moins coûteux comme le mode téléphonique.

4.6.3. Une enquête téléphonique

La concentration des objectifs de l'enquête permettra de réduire le volume du questionnaire. Dès lors, le téléphone peut être envisagé comme méthode d'enquête. Cette méthode présente de multiples avantages par rapport à une enquête réalisée à domicile :

- une réduction des coûts d'enquête grâce à la suppression des déplacements pour se rendre chez l'enquêté, à un plus grand nombre d'enquêtes réalisées chaque soir par les enquêtés, une équipe restreinte d'enquêteurs permettant des économies sur leur formation et leur supervision ;
- une meilleure répartition géographique des enquêteurs qui ne sont plus obligés de se spécialiser sur une zone limitée ;
- un meilleur contrôle et suivi des enquêteurs et du déroulement de l'enquête en temps réel lorsque l'on dispose d'un local unique d'enquête. Il est même possible, comme à Montréal (Chapleau, 1997 ; Dorland et al., 1997), de procéder à une saisie directe des données. On peut alors concevoir un programme de validation des données en cours d'enquête et d'aides aux enquêteurs (Chapleau et al., 1999). L'utilisation de plus en plus courante par les instituts de sondage de ce mode d'enquête a conduit au développement de logiciels CATI. Ces derniers facilitent le travail des enquêteurs, notamment pour le codage de questions pouvant être délicates comme la profession ou l'adresse (ces aides sont également possibles pour des enquêtes à domicile avec les systèmes CAPI, mais ils nécessitent l'usage de portables durant l'enquête). Ils offrent également une disponibilité plus rapide des données ;
- un élargissement de l'aire d'enquête sans surcoût important. Celui-ci se résume à l'accroissement de la taille de l'échantillon, alors que le coût unitaire par enquête augmente dans le cas du face à face en lien avec l'allongement des distances de déplacement.

Ces avantages expliquent le développement de cette technique dans la plupart des domaines d'enquête. De nombreuses expérimentations ont été réalisées (Collins, Sykes, 1987 ; Groves et al., 1988 ; Nathan, Eliav, 1988 ; Clayton, Werking, 1991 ; Drew, 1991 ; De Leeuw, 1992 ; De Leeuw, Van der Zouwen, 1992), principalement dans les pays anglo-saxons, pour comparer la qualité des données. Il en ressort une précision légèrement plus faible dans les enquêtes téléphoniques. Mais, généralement, l'écart est trop faible pour justifier le surcoût des enquêtes en face à face. La principale difficulté réside dans l'équipement téléphonique qui n'est pas encore total, dans les listes « rouges » et « orange » et la montée en puissance des abonnés au portable uniquement.

Toutefois, ces expérimentations ont presque toutes été réalisées dans d'autres domaines que les transports. Les expériences françaises dans le domaine des enquêtes déplacements à Grenoble (CETE de Lyon, CETUR, 1986), Le Havre (CETE Normandie et al., 1992), Paris et Chambéry n'ont pas permis de conclure en raison d'échantillons trop faibles et de difficultés dans la constitution des échantillons. De ce fait, certaines critiques subsistent. Elles portent sur :

- la qualité des réponses. La relation de nature différente entre l'enquêteur et l'enquêté du fait de l'absence de face à face peut avoir des incidences sur les résultats. Toutefois, les spécialistes du marketing téléphonique estiment au contraire que la relation est de meilleure qualité car elle se fait sans intrusion au domicile de l'enquêté ;
- la représentativité de l'échantillon. Un taux de refus parfois plus élevé (Smit, Dijkstra, 1991), l'exclusion des ménages non équipés de téléphone et le nombre de ménages sur liste rouge ou liste orange amène des distorsions dans la représentativité des résultats. Ce dernier problème peut être résolu par différentes techniques de génération de numéros (« *Random Digit Dialling* » aux Etats-Unis ou « déclinaison » en France (Tassi, 1992), toutefois ces techniques sont actuellement interdites en France par la C.N.I.L.). Mais dans ce cas, il n'est plus possible d'obtenir l'adresse du ménage enquêté pour l'informer par courrier préalablement à l'enquête. Ces catégories représentent environ 30 % de la population. Enfin, le développement très rapide de la téléphonie mobile pose des problèmes spécifiques (Roy, Vanheuverzwyn, 2000), notamment de disponibilité des individus enquêtés pour répondre à une enquête ;

- le risque d'oubli de "petits" déplacements. Ce point rejoint le premier. Dans le face à face, l'enquêteur a pour consigne de relancer l'enquêté pour s'assurer qu'il n'omette aucun déplacement si « petit » soit-il. Les promoteurs du face à face estiment que cette relance est moins efficace par téléphone. Toutefois, les comparaisons faites en France ne permettent pas de mettre en évidence des différences statistiquement significatives (Bonnel, 1999).

Les premières expériences ne permettent pas de confirmer ou d'infirmer ces craintes. D'autres expérimentations seraient souhaitables dans le domaine des enquêtes déplacements pour pouvoir généraliser les résultats méthodologiques obtenus dans d'autres domaines. Toutefois, l'expérience accumulée au niveau international dans les enquêtes déplacements semble indiquer que ce mode d'enquête produit des résultats d'une fiabilité statistique comparable au face à face pour un coût nettement plus faible. A titre illustratif, l'analyse des enquêtes déplacements étrangères met en évidence une utilisation plus limitée du face à face comparativement aux autres modes d'enquête :

- Allemagne et Autriche : postal (Brög, 2000 ; Sammer, 1995) ;
- Australie : Melbourne : postal, Sydney : face à face (Richardson, Ampt, 1993) ;
- Canada : Montréal et Toronto : téléphone (Chapleau, 1997, Data Management Group, 1992) ;
- Etats-Unis : face à face dans le passé (ce mode a été abandonné principalement pour des raisons de sécurité des enquêteurs), aujourd'hui : principalement envoi postal et recueil par téléphone (Stopher, 1998) ;
- Pays-Bas : jusqu'en 1983 : face à face, de 1984 à 1998 : téléphone, depuis 1999 postal (pour accroître le taux de réponse devenu trop faible) (Karsten, Koning, 1985 ; Moritz, Brög, 1999) ;
- Portugal : Porto et Lisbonne principalement téléphonique (Viégas, 1997) ;
- Suisse : postal, puis téléphonique depuis 1994 (pour recueillir plus finement les déplacements) (Junod, 1997).

4.6.4. Un périmètre élargi

Le périmètre des enquêtes ménages déplacements était souvent trop restreint. Il faudrait privilégier un périmètre évolutif débordant des contours du bassin d'attraction de l'agglomération. La croissance la plus forte en termes d'habitat et d'activité se situe généralement au-delà des frontières administratives de l'agglomération. Les déplacements générés par ces localisations devraient être l'objet d'une plus grande attention, car ils sont des amplificateurs des trois crises que connaît le système urbain. Par l'allongement des distances parcourues, ils accroissent les problèmes de pollution. Ils se situent dans des zones peu denses où la production d'une offre de transports collectifs de qualité est coûteuse. Le report de déplacements auparavant réalisés dans les zones plus centrales vers la périphérie s'accompagne d'un moindre usage des transports collectifs, ce qui pèse sur les recettes de ce mode. Enfin, la dépendance généralement assez forte, au moins pour le travail, avec l'agglomération aggrave les problèmes de congestion aux entrées de villes. La difficulté à mettre en place des modèles sur les bassins d'attraction élargis des agglomérations s'explique en partie par ce manque de données (Cabanne et al., 2000). Toutefois, les toutes dernières enquêtes, notamment Saint-Etienne (printemps 2001) et Grenoble (2001-2002) ont vu leur périmètre s'élargir très fortement pour englober l'ensemble des bassins de vie des deux agglomérations. Nous ne pouvons qu'espérer que ces contre-exemples récents sont le début d'une généralisation de l'élargissement des périmètres.

Le choix d'une enquête téléphonique s'explique aussi par cet objectif. L'élargissement du périmètre d'enquête n'entraîne pas d'accroissement significatif du coût unitaire de l'enquête téléphonique à l'inverse des enquêtes en face à face.

4.6.5. Une codification plus fine des adresses

Nous avons déjà justifié la pertinence d'une codification plus fine des localisations (section 4.4, p. 127). En revanche, la question de faisabilité technique se pose notamment dans le cas du mode téléphonique. En face à face, l'utilisation de cartes permet d'envisager une codification très fine des

adresses (l'enquête globale parisienne assure ainsi une codification à l'aide d'un carroyage de 300 mètres de côté sur l'ensemble de la région Ile de France (DREIF, 1995)).

Par téléphone, l'utilisation de cartes n'est évidemment pas envisageable. Par contre, il est possible de demander des informations aux enquêtés permettant d'assurer une codification des adresses. Il faut pour cela utiliser ou construire des bases de données comprenant diverses informations permettant cette codification. A titre d'information, les enquêtes réalisées à Montréal ou à Toronto s'appuient sur les bases suivantes (Chapleau, 1991, 1993) :

- le nom de l'établissement correspondant (entreprise, administration, commerce...). Des dictionnaires constitués progressivement assurent la correspondance avec une coordonnée x-y ;
- le code postal de l'adresse, qui, au Canada, correspond généralement à un côté d'îlot dans les zones urbanisées ;
- l'adresse, qui grâce à un dictionnaire peut être associée à une coordonnée x-y ;
- l'intersection la plus proche, le dictionnaire d'intersections de rues permet d'identifier la coordonnée x-y de l'intersection. On ne dispose donc plus d'une information exacte, mais d'une information approchée qui reste toutefois suffisamment précise pour la plupart des applications ;
- un lieu public proche (place publique, édifice, établissement...). L'utilisation des dictionnaires permet comme précédemment d'associer à ce lieu une coordonnée x-y ;
- enfin, si aucune de ces informations ne peut être obtenue, l'enquêteur a pour consigne d'obtenir une information même si elle est incomplète. Cette information sera ensuite utilisée pour créer une adresse probable ou pour la répartir entre plusieurs adresses affectées de poids dont la somme est égale à 1. Par exemple, si une adresse ne comporte que le nom de la rue sans le numéro, il est possible de répartir cette adresse au prorata des adresses obtenues dans cette rue au cours de l'enquête. Il est même possible de raffiner cette codification en faisant intervenir le motif du déplacement considéré. La répartition de cette adresse sera obtenue au pro rata des adresses obtenues dans cette rue au cours de l'enquête pour le même motif. On peut également faire intervenir le mode, s'il est collectif, en privilégiant les adresses situées à proximité de la ligne de bus en question.... La procédure utilisée permet de construire une information probable à partir d'une information incomplète, plutôt que de la rejeter en totalité.

La réflexion développée au fil des enquêtes sur ce principe de codification amène les équipes canadiennes à ne plus raisonner en localisation physique, mais plutôt en localisation fonctionnelle. L'individu ne se rend plus dans la zone 36, ni même à la destination x-y, mais dans le supermarché "bidule", dont on sait grâce à un dictionnaire géographique qu'il est situé en x-y. C'est ainsi une nouvelle approche de l'espace qui est en train de voir le jour dans les études de planification (Chapleau, 1993). L'expérience canadienne montre que cette évolution ne pose pas de problème méthodologique. Elle demande un investissement initial dans la constitution des dictionnaires d'adresses. Mais il semble qu'à plus long terme cet investissement se traduise par des coûts de codification plus réduits. Le développement des Systèmes d'Informations Géographiques et de la puissance des ordinateurs permet même d'envisager actuellement une codification en temps réel au cours de l'enquête. Ainsi cette évolution est de celles qui pourraient être facilement introduites dans les enquêtes ménages déplacements, dans le sens où elle ne remet pas en cause les choix actuels. Elle apporte une information beaucoup plus riche et rien n'empêche ceux qui souhaitent continuer à utiliser des découpages d'y revenir.

4.6.6. Mais un dispositif à tester

Le dispositif actuel des enquêtes ménages déplacements avec une méthodologie commune au cours des 30 dernières années et entre les agglomérations françaises offre des possibilités de comparaison tant temporelle que spatiale. Il convient toutefois d'être vigilant lors de ces comparaisons. Une méthodologie commune ne signifie pas toujours une mise en œuvre

rigoureusement identique. Ce peut être le cas des petits déplacements de marche à pied pour lesquels le seuil à partir duquel il y a déplacement est parfois délicat à mettre en œuvre et surtout à faire comprendre aux enquêtés (même si l'enquêteur est en principe garant de l'application de la définition). C'est également, le cas de l'évolution des périmètres d'enquête. Ceux-ci doivent évoluer pour suivre l'évolution de la péri-urbanisation. Mais pour les comparaisons temporelles, faut-il raisonner à périmètre constant ? ou à aire d'attraction constante ?....

Au-delà de ces difficultés, le maintien de la comparabilité temporelle est un objectif fort, d'autant plus que l'analyse de la mobilité et des politiques de déplacements insiste de plus en plus sur l'importance de la prise en compte des dynamiques d'évolution. Nous pensons toutefois que la mise en œuvre de cet objectif ne doit pas s'accompagner d'un refus de toute évolution méthodologique. L'examen des pratiques d'enquêtes déplacements à l'étranger ou de certaines enquêtes (hors déplacements urbains) en France montre que cette évolution permet d'accroître la qualité ou de réduire le coût des données produites.

Toutefois, il est impératif d'accompagner toutes modifications méthodologiques d'une phase de test. Elle doit viser un double objectif, d'une part s'assurer d'un gain effectif, et d'autre part s'assurer d'une possibilité de comparabilité avec les enquêtes précédentes afin de ne pas couper les séries temporelles. L'analyse de ces conditions nous conduirait trop loin par rapport aux objectifs de cet ouvrage. Le lecteur intéressé pourra toutefois consulter les documents suivants (Brög et al., 1983 ; Brög, Meyburg, 1983, Polak, Ampt, 1996 ; Liss, 1999 ; Moritz, Brög, 1999 ; Van Evert, Moritz, 2000 ; Bonnel, 2001).

5. LES ENQUETES DE PREFERENCES DECLAREES

Les enquêtes présentées ci-dessus correspondent au recueil des comportements effectifs des individus. On les regroupe sous le terme générique d'enquêtes de préférences révélées (« *Revealed Preferences* » ou RP en anglais), l'individu révélant à travers l'enquête son comportement passé. Ces enquêtes constituent une source de données indispensables pour l'alimentation des modèles. Toutefois, elles ne permettent pas toujours de répondre pleinement à certains besoins de modélisation. C'est pour répondre à certaines de ces situations que se sont développées progressivement les méthodes d'enquêtes de préférences déclarées (« *Stated Preferences* » ou SP en anglais).

Nous présentons tout d'abord les principaux domaines d'applications (section 5.1). Ensuite, nous décrivons une classification des méthodes regroupées sous ce terme générique (section 5.2) pour développer ensuite celles qui sont principalement utilisées en modélisation (section 5.3). Enfin, nous exposons quelques recommandations concernant ces méthodes (section 5.4). Cette présentation reste volontairement sommaire. Nous engageons le lecteur intéressé par ces enquêtes à consulter la bibliographie abondante sur ces questions (Lee-Gosselin et al., 1998 ; Bonnel et al., 1997 ; Polak, Jones, 1997 ; Hensher, 1994 (l'article prend place dans un numéro spécial de la revue *Transportation* sur les préférences déclarées où l'on trouve d'autres articles intéressants sur le sujet) ; Louvière, 1988, 1992 ; Bates, 1988, 1998 ; Bradley, 1988 ; Gayda, 1993 ; ainsi que de nombreux articles dans des revues anglo-saxonnes portant sur l'économie des transports, les méthodes d'enquêtes... et des contributions aux principaux colloques de la communauté internationale du transport WCTR (« *World Conference on Transport Research* », Conférence Mondiale sur la Recherche dans les Transports), IATBR (« *International Association of Travel Behaviour Research* », Association Internationale de Recherche sur les Comportements de Déplacements), Méthode d'enquêtes...). Les pièges de ce type d'enquête sont effectivement nombreux pouvant conduire à des décalages importants entre les réponses à des stimuli au cours de l'enquête et les comportements effectifs lorsque les individus sont effectivement confrontés à ces stimuli (Bonnel, 1989a, 1989b, 1995b).

5.1. Principaux domaines d'applications : situations nouvelles et estimations de certains indicateurs

Notre analyse des principaux domaines d'application est évidemment orientée vers les perspectives de modélisation. Ces méthodes issues de la recherche en marketing répondent principalement à deux préoccupations :

- la modélisation de contextes nouveaux inexpérimentés par les individus. Cette situation est fréquente, que ce soit lors de l'introduction d'un mode nouveau : tramway ou métro dans une ville qui n'en possède pas, train à grande vitesse, véhicule électrique, voiture en libre service..., ou encore lors de l'introduction de nouveaux modes de régulation de la voiture : voies réservées pour la voiture partagée, restriction d'accès, péage d'accès.... Dans ce cas, l'analyste est souvent confronté à des inconnues quant à la réaction des individus face à ces situations nouvelles. Les enquêtes de préférences révélées ne peuvent pas toujours fournir les éléments pour simuler ces situations. Les enquêtes de réponses déclarées viseront alors à simuler les réactions des individus face à ces situations ou à fournir les données permettant de le faire (Faivre d'Arcier et al., 1998) ;
- la quantification de certains indicateurs. C'est notamment le cas de données de nature plus qualitatives. La prise en compte de certaines données comme le confort, la propreté des véhicules ou des stations, l'accueil... est souvent assez délicate dans un modèle (Ampt, Swanson, 1996), car ces données se prêtent assez mal à une quantification à partir de données issues d'enquêtes de préférences révélées. Les enquêtes de préférences déclarées permettent de caler des fonctions d'utilité prenant en compte ces dimensions. Ces enquêtes sont également utilisées pour quantifier des données a priori plus objectives comme la valeur du temps (Ségonne, 1998) ou la perception relative des différents temps élémentaires composant un déplacement (attente, marche à pied...).

Ces deux domaines principaux ne sont évidemment pas exclusifs d'autres objectifs en modélisation et a fortiori en économie des transports plus généralement, comme le suggère la classification développée ci-dessous.

5.2. Classification des approches de réponses déclarées (SR, « Stated Response »)

Nous empruntons ici la classification proposée par Lee-Gosselin (1997, tableau 4.3). Elle illustre la diversité des méthodes classées habituellement sous ce terme générique (certains auteurs utilisent également de manière abusive le terme de préférences déclarées pour englober l'ensemble de ces méthodes).

En fait la plupart des enquêtes réalisées dans le cadre de l'alimentation des modèles de prévision de la demande de déplacements appartiennent à ce que Lee-Gosselin appelle les préférences déclarées. C'est-à-dire des méthodes dans lesquelles tant les contraintes que les réponses comportementales sont explicitées. Elles ont connu un développement rapide dans les pays anglo-saxons et sont couramment associées aux enquêtes de préférences révélées (Stopher, 1998). En France, leur utilisation est plus récente et reste encore assez limitée. Elles sont le plus souvent utilisées pour estimer des fonctions d'utilité relatives aux alternatives à simuler. Les autres méthodes sont moins souvent utilisées en modélisation. Elles visent davantage à analyser les comportements, notamment sous la perspective des processus de décision, des réactions aux changements, des processus d'apprentissage... (Lee-Gosselin et al., 1998).

CONTRAINTES	
(exprimées en attributs : personnel/foyer/social/spatial/offre,...)	
plutôt explicitées	plutôt élicitées

COMPORTEMENTALES	<i>plutôt explicitées</i>	<p>PREFERENCE DECLAREE</p> <p>(intérêt = compromis, utilité)</p> <p>«Etant donné les niveaux d'attributs dans ces options, laquelle préféreriez-vous : [A]... ? [B]... ?...»</p>	<p>TOLERANCE DECLAREE</p> <p>(intérêt = limites d'acceptation et seuils de changement)</p> <p>«Dans quelles circonstances pourriez-vous vous imaginer faire : [r1]... ? [r2]... ?...»</p>
	REPONSES	<i>plutôt élicitées</i>	<p>ADAPTATION DECLAREE</p> <p>(intérêt = comportement provisoire et réactif ; résolution de problèmes, règles)</p> <p>«Que feriez-vous différemment si vous vous trouviez confronté à ces contraintes spécifiques : [...scénario détaillé] ?»</p>

Tableau 4.3 : Taxinomie des approches d'enquêtes de Réponses Déclarées

Sous le terme générique des approches de réponses déclarées, sont regroupées des enquêtes de nature très différente depuis des entretiens très ouverts s'appuyant sur des jeux de simulation jusqu'à des enquêtes par questionnaire fermé soumises à des procédures rigoureusement standardisées. Mais ce qui réunit ces enquêtes concerne le fait qu'elles cherchent à tirer des conclusions à partir des préférences ou comportements individuels dans des conditions expérimentales ou quasi-expérimentales (Polak, Jones, 1997), ce qui les rapproche de certaines expériences tentées en économies expérimentales.

Nous présentons rapidement ci-dessous les enquêtes de préférences déclarées répondant au première cadran de la typologie de Lee-Gosselin.

5.3. Enquêtes de préférences déclarées

Elles nécessitent d'identifier au préalable les variables déterminantes du choix entre les alternatives étudiées, et pour chacune des variables de définir leur niveau de variation. Par exemple dans une étude sur les trains à longue distance en Suède, Wildlert (Wildlert, 1998) propose les variables suivantes : tarif, temps de déplacement, fréquence, confort du wagon. Pour chacune de ces variables, des niveaux réalistes sont définis (tableau 4.4). En combinant chacun des niveaux de chacune des variables, on peut construire l'ensemble des alternatives possibles. L'enquête consiste alors à soumettre certaines de ces alternatives à l'enquêté. Les modes de réponse peuvent être variés : choix de l'alternative préférée, classement des alternatives, évaluation de chacune des alternatives sur une échelle sont les principaux modes utilisés.

Même avec un petit nombre de variables et un petit nombre de niveaux, on arrive rapidement à un nombre d'alternatives élevé. Dans l'exemple précédent, si chacune des variables présente 3 niveaux, il y a $3^4 = 81$ alternatives. Il est clair que l'on ne peut demander aux enquêtés d'examiner

l'ensemble des alternatives. La construction d'un plan d'expérience permet de réduire fortement le nombre d'alternatives à soumettre à l'enquête.

Variables	Valeurs prises dans le questionnaire
Tarif	-5, +5, +10 % comparé au tarif actuel
Temps de déplacements	±0, -10, -25 % comparé au temps actuel
Fréquence	2 trains par heure, 1 train par heure, 1 train toutes les deux heures
Confort du wagon	rénovation des anciens wagons, wagon actuel de l'intercité, wagon de train à grande vitesse

Tableau 4.4 : Variables étudiées dans l'exercice « *stated preference* », sur les trains à longue distance entre Stockholm, Göteborg et Sundsvall (Wildlert, 1998)

Apparemment, ces enquêtes sont donc assez simples à mettre en œuvre. Toutefois, cette simplicité n'est qu'apparente, car elles nécessitent le respect d'une méthodologie rigoureuse. Bradley (Bradley, 1988) identifie ainsi 10 étapes dans la mise en œuvre d'une étude de préférence déclarée :

- 1- Déterminer ce qui doit être prédit ou évalué ;
- 2- Déterminer le contexte transport (hypothétique) dans lequel les données doivent être collectées ;
- 3- Déterminer la population cible de l'enquête et la méthode d'échantillonnage ;
- 4- Construire les choix ou les alternatives devant être présentés ;
- 5- Décider d'une méthode pour mesurer les préférences ou choix durant l'enquête ;
- 6- Développer la méthode de présentation de l'expérimentation (ensemble des associations d'alternatives) aux enquêtés ;
- 7- Administrer l'enquête ;
- 8- Estimer le modèle d'utilité à partir des données de l'enquête ;
- 9- Evaluer la validité du modèle de prévision ;
- 10- Utiliser le modèle pour l'évaluation ou la prévision.

Bradley recommande donc d'examiner la validité du modèle ainsi construit. Il distingue en fait la validité interne (qui concerne les points 4 à 8), de la validité externe (points 1 à 3, 9 et 10). La validité interne correspond à l'analyse de la cohérence de l'ensemble de la méthodologie développée et à la reproductibilité des réponses elles-mêmes sous des protocoles analogues. La validité externe se réfère au degré de correspondance entre les réponses obtenues au cours de l'enquête et les réponses du « monde réel » (Polak, Jones, 1997). L'examen de ces questions conduit à respecter certaines précautions méthodologiques.

5.4. Précautions méthodologiques

Lorsque ces méthodes commencèrent à être appliquées dans le domaine des transports (au début des années 70 (Polak, Jones, 1997)), elles donnèrent fréquemment des résultats de piètre qualité. Des différences de l'ordre de 100 % entre les prévisions et ce qui était observé par la suite étaient courantes (Ortuzar, Willumsen, 1994). Wildlert dans l'étude citée ci-dessus compare 25 design différents d'enquête. La valeur du temps obtenue par ces enquêtes varie dans un rapport pouvant aller de 1 à 4, malgré des variables et des niveaux de variables identiques et des échantillons comparables. L'auteur souligne 2 raisons principales. Les individus ont tendance à simplifier le travail qui leur est demandé, cela est plus particulièrement visible dans des exercices de classement d'alternatives ou de notation sur des échelles de valeur que dans des exercices de choix entre deux alternatives. Il est important que

l'exercice soit resitué dans le contexte de l'individu. Dans la négative, l'éventail des réponses est beaucoup plus large.

L'expérience acquise depuis les années 70 permet d'éviter certaines erreurs. Nous reprenons ici, les recommandations de Bradley (Bradley, 1988) concernant l'analyse de la validité du modèle produit à partir des données d'enquêtes de préférences déclarées.

En ce qui concerne la validité interne (points 4 à 8 ci-dessus), diverses analyses ont permis de progresser : comparaisons de différentes méthodes de construction de l'ensemble des alternatives devant être présentées ; comparaisons des méthodes de mesures des préférences ou des choix ; comparaison de méthodes de présentation des alternatives ou choix ; comparaison de modèles et de méthodes d'ajustement des modèles. Les critères utilisés pour analyser la validité interne sont les tests statistiques usuels retenus pour analyser les modèles ; l'analyse de la pertinence du modèle et enfin un critère moins souvent utilisé, la reproductibilité c'est-à-dire la probabilité que face à un protocole d'enquête identique les enquêtés fournissent une réponse identique pour une même question. L'analyse de la validité interne a fait l'objet de nombreuses études, publiées principalement dans la littérature marketing, qui permettent de faire face aux principaux problèmes de validité interne sous réserve d'un respect des hypothèses et des conditions d'applications de ces méthodologies. Toutefois, pour Polak et Jones (Polak, Jones, 1997), des recherches complémentaires sont nécessaires sur les dimensions suivantes pour progresser dans la maîtrise des questions de validité interne :

- design statistique : plan d'expérience orthogonal ou non-orthogonal, questions standardisées ou adaptatives (selon les caractéristiques de l'individu)... ;
- mode de présentation des alternatives et mode de réponse : choix, classement, notation de couple d'alternatives ou d'ensemble plus vaste d'alternatives... ;
- mode d'enquête : papier versus assisté par ordinateur, administré versus auto-administré, face à face versus téléphone ; questionnaire : choix des mots, mise en page, utilisation de la couleur... ;
- incidence de l'information fournie durant l'enquête : quantité et qualité de l'information, mode de présentation de l'information... ;
- technique d'estimation des modèles.

En revanche, selon Bradley, la validité externe est plus souvent en cause. La principale difficulté en la matière porte sur le réalisme des alternatives et le contexte dans lequel l'individu est placé. Le réalisme des alternatives concerne en fait plusieurs dimensions. Le niveau des attributs des alternatives doit être conforme à ce que l'individu peut trouver dans la réalité, par exemple le prix d'un produit, ou sur des aspects de confort la garantie d'une place assise (garantir une place assise dans le métro aux heures de pointe peut être considéré par certains comme irréaliste). Le choix des attributs doit être pertinent pour l'individu, par exemple dans une étude de marché ne pas proposer des produits que l'individu n'est pas en mesure d'acquérir (par exemple des voitures de luxe, des billets d'avion...). Enfin les alternatives elles-mêmes doivent être réalistes. C'est une difficulté par exemple d'étude sur le péage urbain. Les individus n'acceptent pas forcément de considérer réellement certaines des alternatives par rejet du principe du péage beaucoup plus que du montant éventuel proposé. Le réalisme implique également une bonne compréhension des alternatives proposées. Ces enquêtes sont souvent utilisées pour tester des produits nouveaux ou des situations nouvelles non encore expérimentées par les individus. Le caractère hypothétique des alternatives nécessite alors un soin particulier dans leurs descriptions afin qu'elles soient correctement comprises par les individus.

Le contexte dans lequel l'individu est placé est également très important. Il est généralement nécessaire de contextualiser l'exercice de préférence déclarée. Et plus encore en transport, il est souvent recommandé d'organiser cette contextualisation autour de l'univers de l'individu. Pour cela il est assez fréquent de combiner une enquête de préférence révélée avec une enquête de préférence déclarée. La première permet d'adapter le questionnaire au contexte de l'individu en s'appuyant par exemple sur les caractéristiques des déplacements qu'il a réalisés. Pour revenir à l'exemple du péage

urbain, les réponses des individus seront très probablement différentes selon que les déplacements de l'individu sont concernés ou non par le péage envisagé.

La complexité de ces enquêtes ne doit pas être sous-estimée. Le respect des quelques précautions évoquées ci-dessus n'étant pas garant de la qualité des données produites. Nous ne serions trop recommander aux personnes souhaitant se lancer dans ce type d'exercice de s'entourer de conseils avisés. Il est fréquent que ces données soient combinées avec des données RP (Morikawa, 1994 ; Stopher, 1998 ; Daly, Rohr, 1998). Cette combinaison permet tout à la fois un enrichissement des données produites, et une validation des données SP.

6. EXEMPLE D'APPLICATION DU SCHEMA METHODOLOGIQUE DE PRODUCTION DES DONNEES : EXEMPLE DE LA CONSTRUCTION DU CADRE CONCEPTUEL

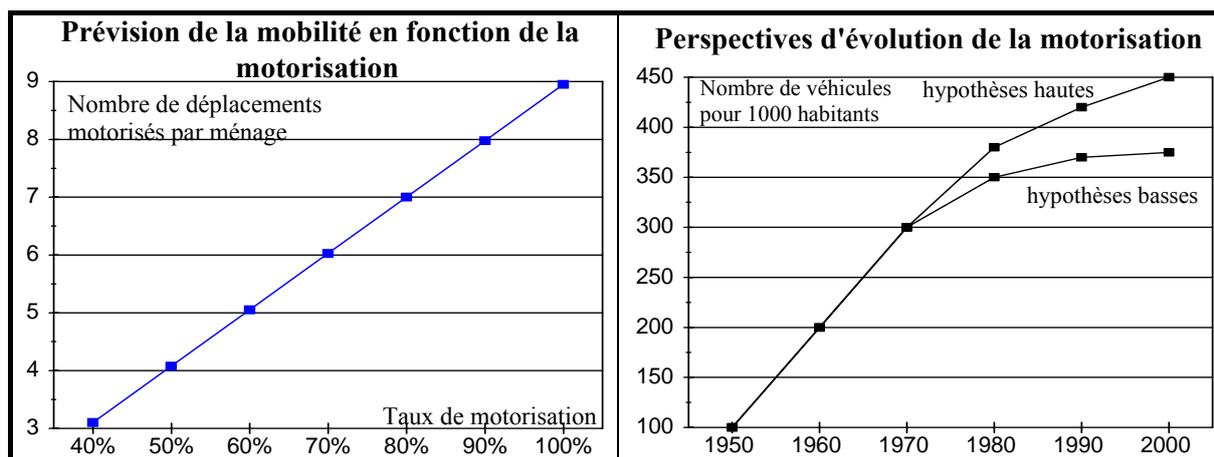
Nous prenons comme exemple la construction du cadre conceptuel d'analyse de la mobilité. Nous illustrons comment ce cadre s'est enrichi progressivement au gré de l'évolution de la définition de la mobilité. Cet exemple nous permet de revenir une fois de plus sur le lien entre la définition des objectifs de la modélisation, du cadre conceptuel d'analyse de l'objet d'étude, du modèle produit et de la production des données.

Il est bien évident que notre propos n'est pas de présenter LE cadre conceptuel de l'analyse de la mobilité. Une telle ambition serait contraire à notre propos. Nous considérons que le choix du cadre conceptuel est le reflet d'une certaine approche du problème qui doit découler des objectifs que l'on s'est fixés. Ortuzar et Willumsen (Ortuzar, Willumsen, 1994) en donnent une illustration très concrète « *il peut être raisonnable d'utiliser un couteau et une fourchette sur une table pour modéliser la position de voitures avant une collision, mais pas pour représenter leurs caractéristiques techniques ou pour représenter les comportements de choix d'itinéraires de leur conducteur* ».

6.1. Premiers modèles dans le cadre des EPIT

La conception de la mobilité dans les premiers modèles de prévision de trafic développés par le SETRA (Services Techniques d'Etudes des Routes et Autoroutes) dans le cadre des EPIT apparaît comme relativement sommaire (graphique 4.1).

La mobilité y est définie comme un nombre de déplacements. De plus, seuls les déplacements motorisés sont considérés. Dans cette approche, quelques caractéristiques socio-économiques des individus ou des ménages expliquent la propension à se déplacer. Dans les étapes suivantes, il n'y a pas d'autres facteurs explicatifs qui soient pris en compte, si ce n'est l'offre de transport. Le cadre conceptuel sous-jacent à cette conception de la mobilité peut être résumé très simplement (schéma 4.4). Dans cette conception de la mobilité, trois concepts influent sur la mobilité : la structure de la ville qui est ici considérée de manière sommaire à travers le nombre de ménages par zone, les caractéristiques socio-économiques du ménage considérées de nouveau de manière sommaire à travers la motorisation et enfin, l'offre de transport considérée uniquement au niveau des modes motorisés voiture particulière et transports collectifs.



Graphique 4.1 : Loi mobilité - motorisation dans le cadre des EPIT
(Sources : SETRA - EPIT - références méthodologiques, (SETRA, 1973))

Les statistiques nécessaires pour alimenter une telle conception de la mobilité et le modèle bâti sur cette conception sont évidemment assez pauvres. Une enquête déplacements (enquête origines-destinations ou enquête ménages déplacements) est nécessaire pour caler le modèle. De même, il convient de recueillir des données sur l'offre de transport, ainsi que les statistiques sur la population et la motorisation de chacune des zones d'études.

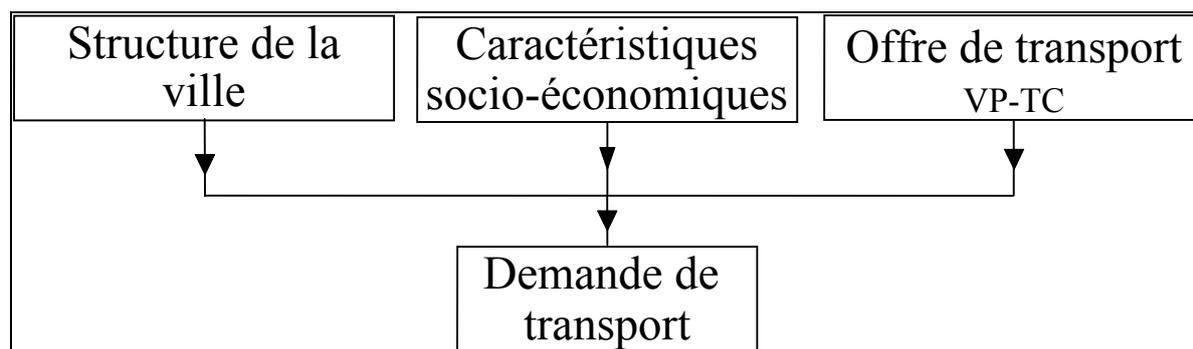


Schéma 4.4 : Cadre conceptuel dans le cadre des EPIT

6.2. Enrichissement du cadre conceptuel

Les nombreuses études sur les déterminants de la mobilité menées dans le courant des années 70 et 80 ont conduit à de nombreux enrichissements de ce cadre conceptuel. Nous n'en ferons pas une présentation exhaustive. Notre propos est plutôt pédagogique pour illustrer la construction d'un cadre conceptuel en fonction de l'évolution de la connaissance de l'objet mobilité, et les conséquences sur la production des données.

Parmi ces acquis, on considère que les individus n'ont pas tous la même perception de l'offre de transport. De plus, les différents temps passés dans les transports ne sont pas tous perçus de la même manière. Enfin, l'hypothèse d'une connaissance exhaustive de l'offre de transport est remise en cause au profit d'une connaissance partielle notamment de l'offre de transports collectifs. Cette conception a conduit à distinguer l'offre de transport telle qu'un analyste transport pourrait la décrire : concept d'« offre de transport - univers objectif de choix », de la perception de cette offre par l'individu : concept d'« offre de transport - univers subjectif de choix ». Le passage de l'un à l'autre apparaît comme un processus de filtre déformant, filtre car l'individu n'a qu'une connaissance partielle et

déformant car l'individu intègre cette offre par son propre prisme. Il est apparu que ce processus dépendait notamment des caractéristiques socio-économiques de l'individu, mais aussi de ses habitudes de déplacement. Cette évolution a conduit au cadre conceptuel suivant (schéma 4.5).

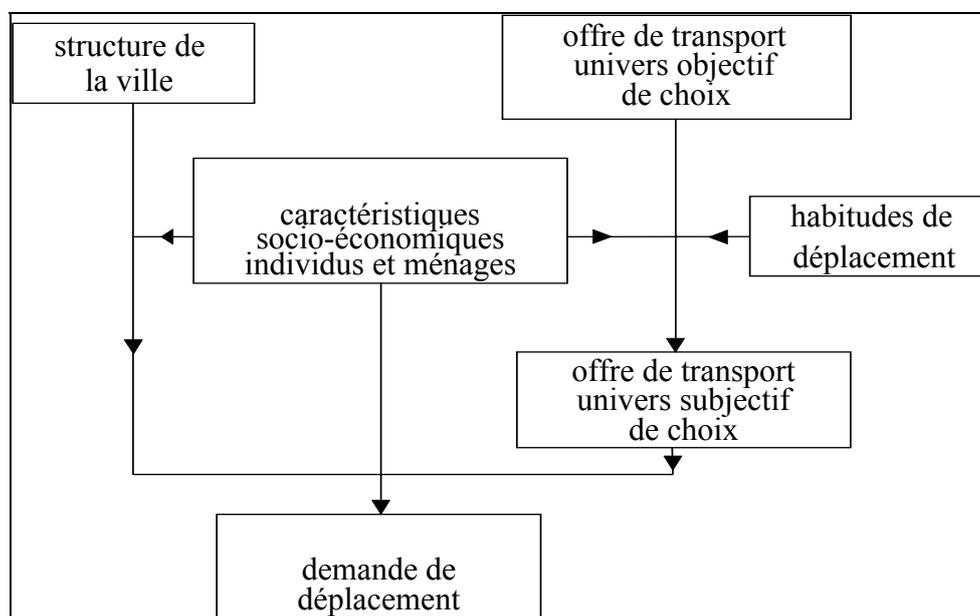


Schéma 4.5 : Cadre conceptuel prenant en compte la perception individuelle de l'offre

Les conséquences de cette nouvelle approche sont importantes sur le contenu de la production des données. Dans le cas précédent, il n'était pas nécessaire de relier les caractéristiques socio-économiques des individus ou du ménage à la mobilité. En revanche, cette fois, ces données doivent être recueillies dans une même enquête afin d'évaluer d'une part le processus menant de l'univers objectif de choix à l'univers subjectif de choix en fonction des caractéristiques individuelles ou du ménage et de ses habitudes de déplacements, et d'autre part l'incidence de cet univers subjectif de choix sur la mobilité. De ce fait, seule une enquête ménages déplacements est à même de répondre à ce nouvel objectif.

Dans cette enquête, il est également nécessaire de recueillir des données sur la connaissance et la perception de l'offre de transport, ainsi que sur les habitudes de déplacement. Ces données sont nécessaires pour formaliser puis caler les relations permettant le passage de l'univers objectif de choix à l'univers subjectif de choix.

Le modèle retenu pourra être assez proche du précédent à ceci près que la formulation de l'offre de transport doit être modifiée. Toutefois, la disponibilité de données socio-économiques recueillies dans le même temps que les déplacements permet d'envisager une formulation faisant intervenir ces caractéristiques.

La seconde évolution que nous présentons, concerne plus directement l'évolution de la définition de la mobilité. Le déplacement n'est généralement pas un but en soi mais permet la réalisation d'activités. La mobilité n'est donc pas qu'un nombre de déplacements, mais plutôt un moyen, de réaliser un programme d'activités qui se concrétise à travers le schéma des activités réalisées (chapitre 3, p. 67). Ce programme d'activités que l'individu souhaite réaliser résulte notamment de son mode de vie, de la répartition des rôles au sein du ménage, de ses contraintes personnelles ou interpersonnelles. Par exemple, la répartition des rôles au sein du ménage conduit à une certaine répartition des tâches d'accompagnement des enfants, ou des tâches ménagères.... Les caractéristiques du mode de vie pourront expliquer la plus ou moins grande propension à la réalisation

de certaines activités (sportives, culturelles...). Enfin, les contraintes personnelles ou interpersonnelles telles le travail ou les ressources du ménage viennent limiter les possibilités de réalisation (Bonnel, 1992). La confrontation de ce programme avec la structure urbaine conduit aux activités souhaitées et aux besoins de déplacement. La confrontation avec l'offre disponible conduit enfin aux activités réalisées et la demande de déplacements que l'on peut observer à travers le schéma d'activités. Cette évolution est traduite dans le cadre conceptuel du schéma 4.6.

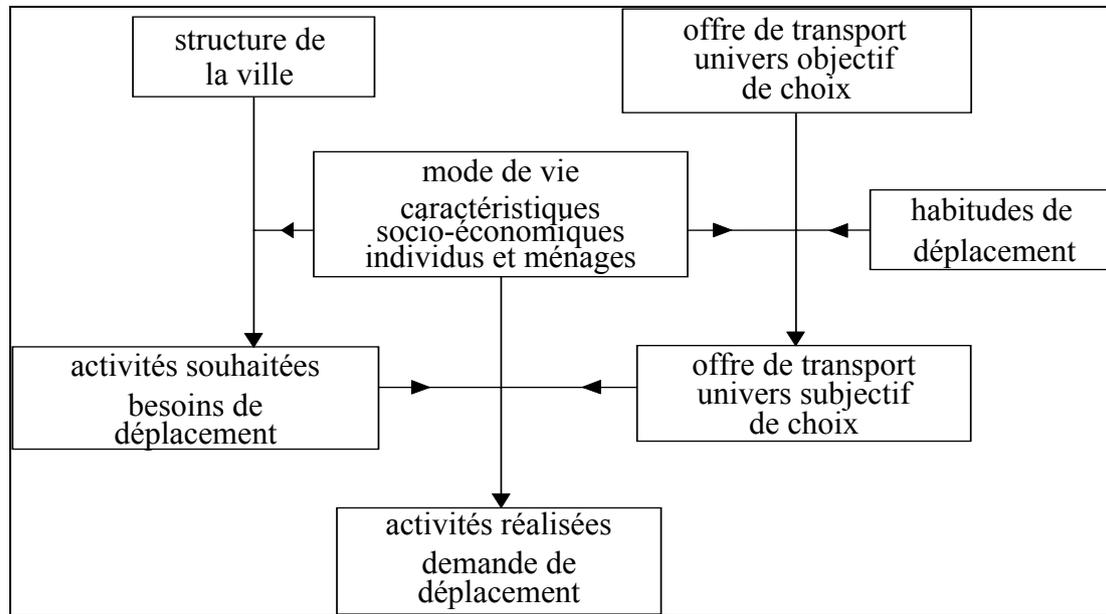


Schéma 4.6 : Cadre conceptuel considérant les déplacements comme demande dérivée des activités

Cette nouvelle approche a bien évidemment des conséquences sur la production des données. Il ne s'agit plus uniquement de collecter des déplacements, mais de recueillir les activités réalisées dans le temps et l'espace. Les déplacements se déduisent alors des changements de lieux. De nouveaux outils ont été construits pour ce recueil du schéma d'activités (schéma 3.1, p. 68). Ce nouveau mode d'approche de la mobilité a débouché d'une part sur une prise en compte des activités à travers les motifs de déplacements dans le cadre des outils plus traditionnels de modélisation et d'autre part sur les modèles basés sur les activités qui restent encore principalement du domaine de la recherche (Ettema, Timmermans, 1997 ; Timmermans et al., 1997).

7. CONCLUSION

Cette présentation générale ne saurait évidemment couvrir la question de la production des données pour la modélisation. Son but était d'ouvrir sur un certain nombre de dimensions indispensables lorsqu'il s'agit de s'interroger d'une part sur le choix d'une méthodologie de production de données dans une perspective modélisatrice et d'autre part sur la qualité des données que l'on utilise. Une dimension relativement importante a toutefois été omise dans cette présentation. Il s'agit des enquêtes longitudinales (enquêtes déplacements périodiques sur des panels de population). Son importance en modélisation nous conduit à proposer deux références pour les lecteurs souhaitant approfondir cette question (Golob et al., 1997 ; Kitamura, 2000). Nous ne souhaitons toutefois pas aborder cette méthode spécifique pour ne pas alourdir encore ce chapitre (de plus, les enquêtes de déplacements urbains par panel ne sont pas très fréquentes).

Tout exercice de modélisation nécessite l'utilisation de données. Le second chapitre nous a permis d'identifier les étapes du processus de prévision où celles-ci sont nécessaires : pour (in)valider le modèle théorique proposé, pour formaliser mathématiquement le modèle, pour le caler et le valider

et enfin pour l'utiliser en prévision. Que les données soient produites par le modélisateur ou qu'il ait recours à des données existantes, il ne peut faire l'impasse d'une réflexion sur la qualité des données produites qui conditionne logiquement la qualité du modèle et des prévisions, c'est-à-dire son opérationnalité. Pour cela, il devra analyser la méthode de production des données pour porter un diagnostic sur leur qualité. Nous espérons que ce chapitre fournit des éléments permettant de l'aider dans cette démarche.

Au terme de cette analyse des différentes dimensions du processus de prévision, il nous reste à examiner l'outil le plus souvent utilisé, en France, pour cela : le modèle à quatre étapes.

Chapitre 5 - Fondements théoriques des modèles de prévision de la demande de déplacements

Nous englobons ici l'ensemble des modèles que l'on inclut généralement sous le terme générique de modèles à quatre étapes (c'est-à-dire l'ensemble des modèles répondant au groupe B de notre classification du chapitre 1 (p. 30)). La portée de ce chapitre déborde toutefois ce groupe et peut s'appliquer plus largement.

Les fondements théoriques s'appuient sur la théorie micro-économique néoclassique du choix du consommateur (section 1). L'application à la sphère des transports a été développée dès les années soixante, notamment par Ben-Akiva (Ben-Akiva, Lerman, 1985), McFadden (Prix Nobel 2000 pour ses travaux sur les choix discrets, Domencich, McFadden, 1975) ou Manheim (1984) souvent cités, si ce n'est comme pères fondateurs au moins comme principaux formalisateurs des fondements théoriques de la transposition de l'approche micro-économique des choix du consommateur dans la sphère des transports. Nous nous appuyons sur ces auteurs pour notre présentation et notamment sur le dernier en ce qui concerne l'analyse de l'équilibre offre-demande (section 2). Nous explicitons ensuite les termes de cet équilibre à travers l'analyse des fonctions d'offre et de demande qui nous renvoient aux notions de coût généralisé et d'utilité. Enfin pour ordonner l'ensemble des modèles englobés sous le terme générique de modèle à quatre étapes, nous présentons différentes dimensions permettant de classer ces modèles (section 3).

1. QUELQUES ELEMENTS DE THEORIE MICRO-ECONOMIQUE NEO-CLASSIQUE DU CHOIX DU CONSOMMATEUR

Notre ambition n'est évidemment pas de développer cette théorie, mais plutôt de donner quelques éléments pour introduire l'application au domaine des transports. Notre présentation se veut donc volontairement simple et rapide (pour aller au-delà de ces quelques éléments, voir notamment : Brémond, 1990 ; Renaud, Tabourin, 1998, auteurs auxquels nous empruntons l'essentiel de notre présentation, ou pour une présentation plus formalisée sur le plan mathématique : Henderson, Quandt, 1980).

Le modèle néoclassique micro-économique traditionnel, s'appuie sur l'analyse du comportement du consommateur (section 1.1 à 1.3) d'une part et sur l'analyse du comportement de l'entreprise (section 1.4) d'autre part.

1.1. Analyse du comportement du consommateur

« Le consommateur type est supposé rationnel. L'unique motivation du consommateur, lorsqu'il acquiert ou vend des biens et services sur le marché, est de maximiser sa satisfaction ou, en d'autres termes, son utilité. » (Brémond, 1990).

« Le consommateur de la théorie néoclassique doit ordonner ses préférences. Dans ce but, il réalise des arbitrages à deux niveaux.

- Le premier concerne les produits eux-mêmes. Il existe un grand nombre de produits, de nature différente. Le consommateur doit, dans un premier temps, retenir les produits qu'il désire acquérir. Parmi l'ensemble des produits disponibles sur le marché »... il retiendra un certain assortiment.

« - Le second est relatif aux quantités. Le consommateur doit en effet, dans un deuxième temps, faire des choix quant aux quantités de chaque bien qu'il désire consommer.

Le croisement des types de produits et des quantités constitue différents paniers de consommation. A chacun de ces paniers sera associé un niveau d'utilité, c'est-à-dire un niveau de satisfaction.

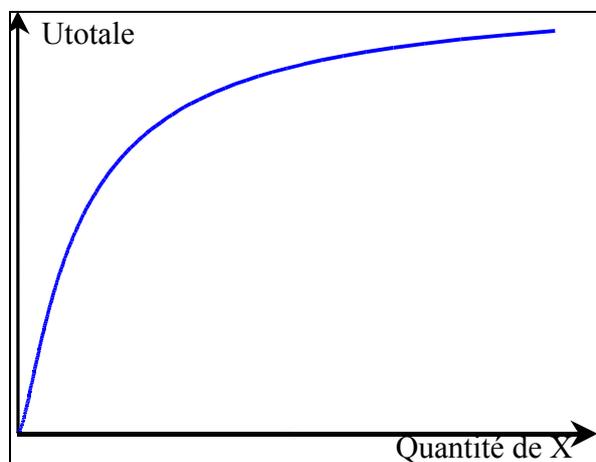
Supposant les produits choisis, la théorie néoclassique s'attache à comprendre comment se mesure l'utilité des différents paniers de biens en fonction des quantités consommées. Pour comprendre l'approche néoclassique, il est indispensable d'approfondir la notion d'utilité marginale... »

«... Les auteurs néoclassiques utilisent une analyse marginaliste. L'utilité marginale est l'utilité supplémentaire procurée par la consommation d'une unité additionnelle d'un bien. Cette utilité marginale est une fonction décroissante de la consommation de ce bien. » (Renaud, Tabourin, 1998).

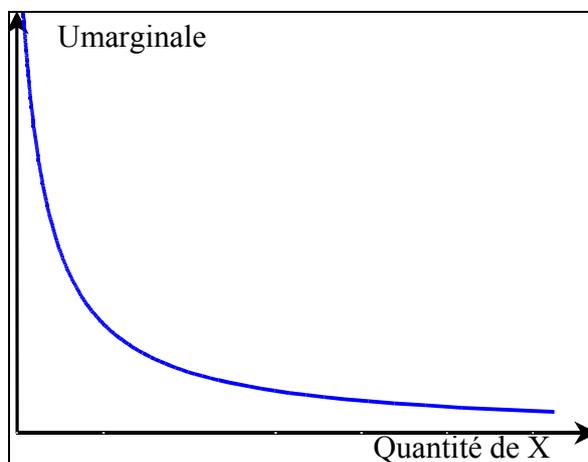
Cette décroissance de la fonction d'utilité marginale est assez intuitive. Plus la consommation d'un bien s'accroît plus la satisfaction que l'on retire de cette consommation décroît. L'utilité marginale tend ainsi vers 0 lorsque la consommation atteint des niveaux de saturation pour l'individu. De nombreux exemples peuvent être donnés, par exemple au niveau alimentaire, pour illustrer la forme de la courbe d'utilité marginale.

En revanche, l'utilité totale est une fonction croissante de la quantité consommée, car l'utilité marginale est une fonction toujours positive.

Si nous considérons un bien divisible, la fonction d'utilité U sera une fonction continue. Bien souvent, même si le bien n'est pas infiniment divisible, on se ramènera à fonction continue par lissage de la courbe d'utilité. Les graphiques 5.1 et 5.2 fournissent la forme générale des fonctions d'utilité totale U_t et d'utilité marginale.



Graphique 5.1 : Forme générale de la fonction d'utilité totale



Graphique 5.2 : Forme générale de la fonction d'utilité marginale

« Comme l'utilité marginale d'un bien est décroissante, le consommateur cherche une allocation de ses ressources qui couvre des consommations variées dont les utilités marginales des premières unités consommées sont les plus fortes. Cela correspond assez bien à la réalité où nous avons tous une multitude de besoins à satisfaire. La situation d'équilibre du consommateur est celle où il a choisi une allocation de ses ressources entre les différents biens, allocation qui le satisfait plus que toute autre.

En situation d'équilibre du consommateur, le dernier franc dépensé a la même utilité, quel que soit le bien acheté. S'il a une utilité supérieure pour l'achat d'un bien A, cela signifie que le consommateur n'a pas acheté assez de A et trop des autres biens.

On notera que cette allocation des ressources est personnelle et marquée temporellement.

- Personnelle, dans la mesure où chaque individu a des besoins subjectifs spécifiques.

- Marquée temporellement, car ces besoins subjectifs évoluent dans le temps.

... Pour passer des préférences individuelles à la consommation elle-même, il faut tenir compte du fait que le choix du consommateur est contraint par un budget restreint dans un environnement de marché. » (Renaud, Tabourin, 1998).

1.2. Optimum du consommateur

La détermination de l'optimum du consommateur amène donc à se poser la question de la quantité consommée de chacun des biens :

$$Q = \{q_1, \dots, q_I\} \quad (5.1)$$

Où q_1, \dots, q_I sont les quantités consommées des biens 1, ..., I.

« Les choix [du consommateur] sont cependant contraints par le niveau de son revenu et par les prix qui se forment sur le marché des biens. » (Renaud, Tabourin, 1998).

La contrainte de revenu (R) s'exprime à partir du prix des produits p_1, \dots, p_I :

$$\sum_i p_i q_i \leq R \quad (5.2)$$

L'individu cherchant à maximiser son utilité, la solution du problème nécessite de poser l'équation de l'utilité totale de l'individu en fonction de la quantité consommée de chacun des biens, soit :

$$U = U(q_1, \dots, q_I) \quad (5.3)$$

On obtient donc un système d'équations :

$$\max_{q_i} (U(q_1, \dots, q_I)) \quad (5.4)$$

sous la contrainte $\sum_i p_i q_i \leq R$

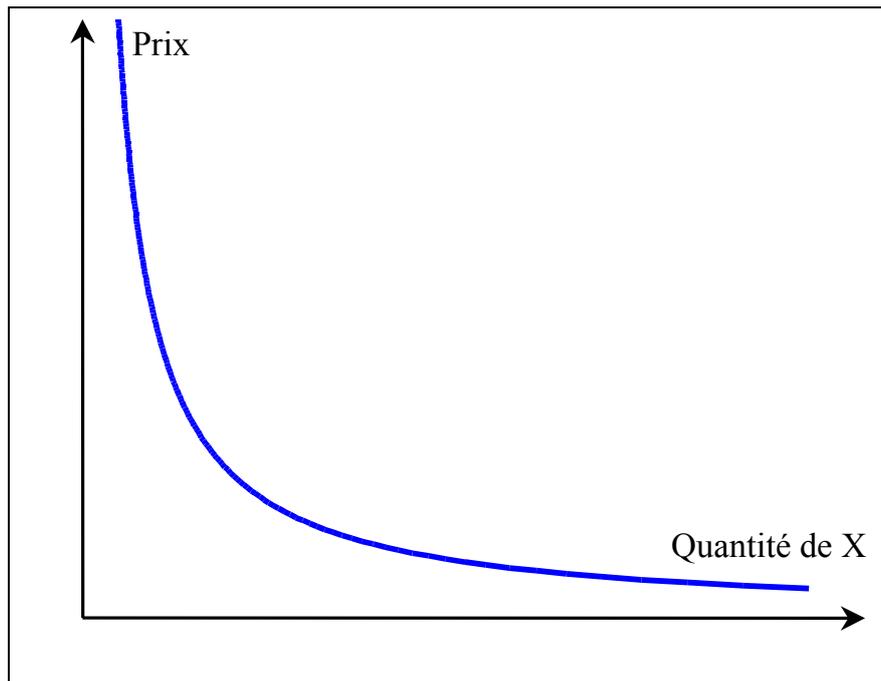
que l'on résout généralement en posant le lagrangien.

1.3. Construction de la courbe de demande

Compte tenu de ce système de contraintes, il est possible de construire pour chacun des produits une courbe de demande en fonction du prix du produit (graphique 5.3). *« Par l'ajustement de sa demande au prix, le consommateur cherche à maintenir un niveau optimal de satisfaction. Cette recherche de l'optimum par le consommateur est le fondement de la configuration de la fonction de demande, à savoir sa décroissance par rapport au prix. » (Renaud, Tabourin, 1998).*

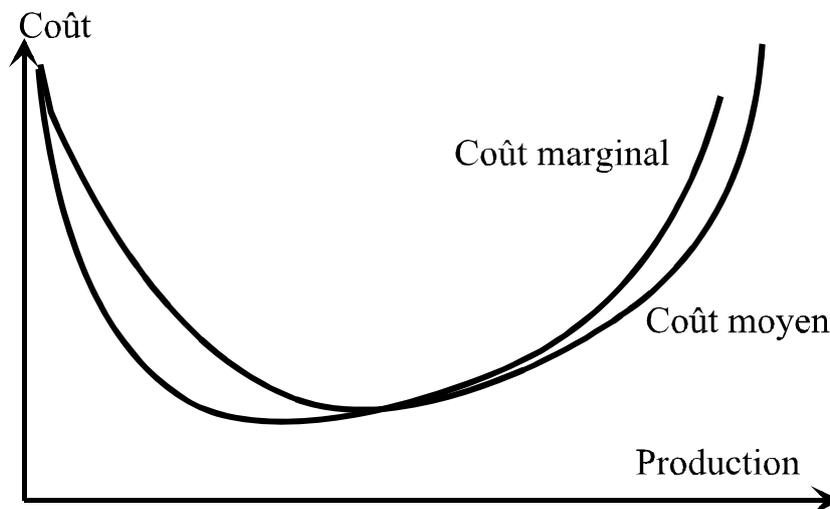
La courbe de demande correspond donc à la maximisation de l'utilité pour le consommateur en fonction du prix de X.

La détermination du niveau de consommation dépend donc du niveau de prix, ce qui nous conduit vers l'analyse du producteur.



Graphique 5.3 : Illustration graphique de la courbe de demande en fonction du prix de X

1.4. Analyse du producteur



Graphique 5.4 : Forme générale des courbes de coûts

Nous ne développons pas l'analyse du producteur, qui apporte peu de plus-value à notre propos. Nous nous contentons d'évoquer les courbes de coûts qui permettent de déterminer les prix des produits. Compte tenu des facteurs de production, on retient généralement une forme en U pour les fonctions de coûts marginal et moyen (graphique 5.4, Renaud, Tabourin, 1998. L'application à la sphère des transports conduit souvent à une courbe de coût marginal différente pour les niveaux de production faible avec un coût marginal très faible et quasiment constant, qui progresse ensuite lentement comme sur le graphique 5.4. Cette forme différente est liée aux coûts fixes généralement très importants). Dans un premier temps, les rendements sont croissants, le coût marginal ainsi que le coût moyen sont décroissants. L'accroissement continu de la production conduit à une seconde phase de rendement décroissant (tout au moins dans une analyse de court terme). Le coût marginal progresse.

Au début, de cette seconde phase, le coût moyen continue à décroître, puis à son tour entame une croissance.

1.5. Généralisation de l'analyse : l'équilibre

L'analyse du consommateur peut être généralisée à l'ensemble des produits. Le consommateur cherchera toujours à maximiser son utilité totale résultant de la consommation des différents produits sous la contrainte de revenu. Cette approche est formalisée par le système d'équations 5.2 à 5.4 (p. 147).

L'analyse peut être généralisée à l'ensemble des consommateurs. De même, la généralisation de l'analyse du producteur à l'ensemble des producteurs permet de déterminer les prix en fonction de la quantité produite. La solution du problème consiste à trouver le point d'équilibre entre ces deux systèmes d'équation. C'est le point d'équilibre entre l'offre et la demande. Cette notion d'équilibre est importante. Elle est largement reprise dans la transposition faite pour l'analyse de la demande de transport, même si elle se pose en des termes un peu différents.

La difficulté de la résolution de ce système tient dans la détermination de l'ensemble des fonctions d'utilité du consommateur pour chacun des biens i et chacun des consommateurs et de l'ensemble des fonctions de production pour chacun des producteurs et pour chacun des produits. De ce fait, l'analyse est souvent simplifiée pour être appliquée à quelques produits appartenant à un domaine particulier. Il est alors souvent fait abstraction (tout au moins de manière directe) de la contrainte de revenu du consommateur.

2. APPLICATION A LA SPHERE DES TRANSPORTS

La résolution de l'équilibre offre-demande généralisé à l'ensemble des produits pose évidemment des problèmes redoutables, ne serait-ce que de calage des coefficients. C'est pour les contourner que l'on considère, par simplification, le domaine d'étude comme un isolat. L'analyse développée ici repose assez largement sur les travaux de Manheim (1984) qui s'appuie, dans sa présentation des modèles transports, sur la formalisation de l'équilibre offre demande appliqué à la sphère des transports (section 2.1). Notre présentation s'en écartera toutefois un peu afin de tirer parti des développements exposés dans les deux premiers chapitres (section 2.2). La transposition au domaine des transports de l'approche néoclassique nécessite de préciser l'interprétation que l'on donne des fonctions d'offre et de demande (sections 2.3 et 2.4). En conclusion, nous revenons sur l'interprétation de cette transposition à la modélisation de la demande de transport (section 2.5).

2.1. Equilibre offre-demande de Manheim

Nous situons notre présentation, comme dans les chapitres précédents, sur le système urbain. L'analyse est toutefois aisément transposable pour d'autres systèmes. Nous nous appuyons sur la formulation des fonctions d'offre et de demande proposée par Manheim (1984).

Il définit la fonction d'offre J par l'équation suivante :

$$S = J(T, V) \quad (5.5)$$

où, S est le niveau de service de l'offre de transport ;

T correspond au système de transport ;

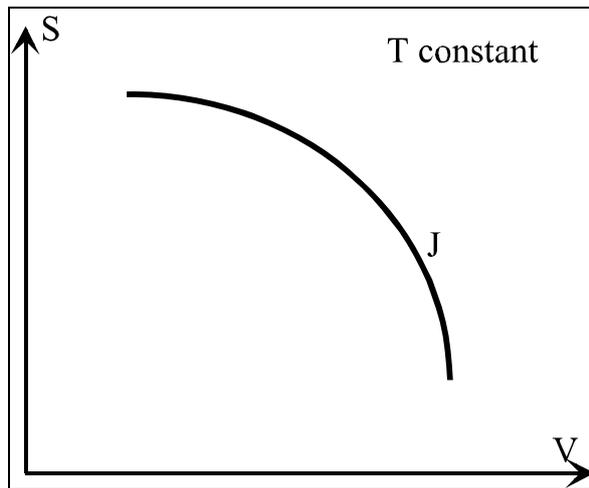
V est le volume de trafic ou de demande observé sur le système de transport.

La fonction de demande est définie par :

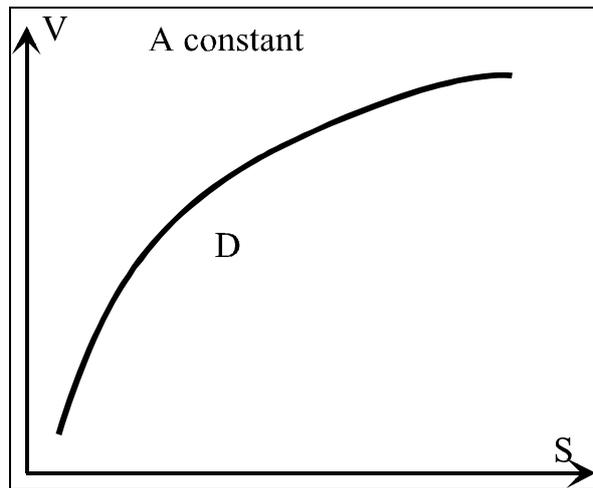
$$V = D(A, S) \quad (5.6)$$

où, A correspond au système d'activités.

Pour un système de transport et un système d'activités donnés, on peut représenter les fonctions d'offre et de demande (graphiques 5.5, 5.6). La forme des courbes se justifie aisément. Pour un système de transport donné, plus la demande sur le système de transport est forte, plus la qualité de service décroît. Une illustration peut en être donnée à travers la courbe débit-coût qui sera présentée dans le chapitre sur l'affectation de la demande (chapitre 10, p. 321). Plus le trafic s'accroît en s'approchant de la capacité de l'axe, plus la vitesse sur l'axe diminue. Pour un système d'activités donné, plus la qualité du service s'accroît plus la demande est forte. C'est ce que l'on peut observer lorsque l'on introduit un système de transport plus performant.



Graphique 5.5 : Fonction d'offre J
(Manheim, 1984)

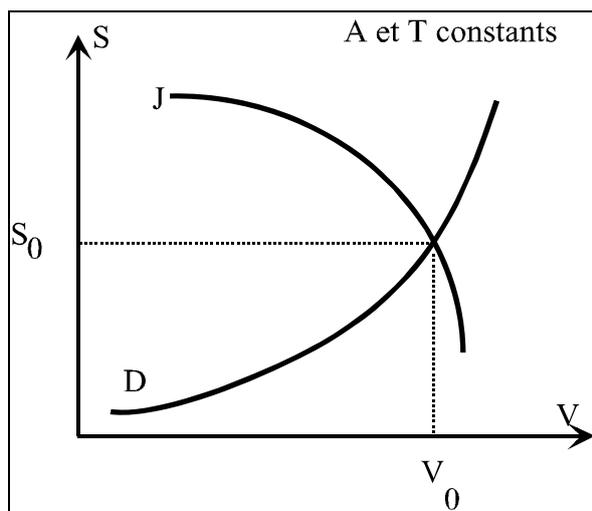


Graphique 5.6 : Fonction de demande D
(Manheim, 1984)

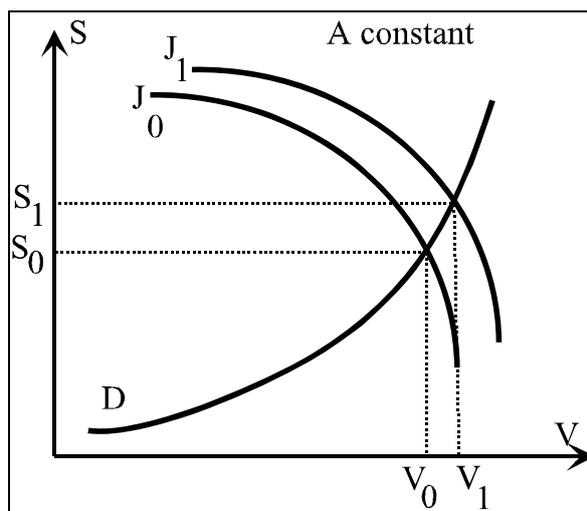
La confrontation des deux courbes d'offre et de demande sur un même graphe permet de trouver le point d'équilibre entre l'offre et la demande pour des systèmes de transport et d'activités donnés (graphique 5.7).

En prévision, la modification du système de transport ou du système d'activités conduit à une nouvelle situation d'équilibre (graphique 5.8). Dans l'illustration qui est donnée, nous considérons une amélioration du système de transport (et un système d'activités figé, sinon, il faudrait également faire évoluer la fonction de demande). La fonction d'offre est alors décalée vers le haut. Par exemple, la construction d'une voie supplémentaire sur une route permettra de repousser les phénomènes de saturation sur des niveaux de trafic plus élevés. Il s'établit donc un nouvel équilibre correspondant à un niveau de trafic plus élevé et à une meilleure qualité de service.

On pourrait de même faire évoluer le système d'activités d'un état 0 vers un état 1, ce qui aurait pour incidence de modifier la courbe de demande et de conduire à un nouvel équilibre. L'enjeu de la modélisation consiste donc à définir les relations en jeu dans la définition de cet équilibre afin de les formaliser.



Graphique 5.7 : Equilibre offre-demande
(Manheim, 1984)



Graphique 5.8 : Incidence d'une amélioration du
système de transport
(Manheim, 1984)

2.2. Interprétation au sein du système urbain

Nous souhaitons reprendre la présentation de Manheim mais en la resituant dans le cadre de l'approche développée dans le premier chapitre (section 3.1, p. 26) du système urbain décomposé en trois sous-systèmes :

- sous-système de transport ;
- sous-système de localisation des activités ;
- sous-système des relations sociales.

Nous avons défini la demande de transport exprimée comme étant la résultante de la « confrontation » des trois sous-systèmes. Le modèle doit donc traduire ces interactions entre les trois sous-systèmes. Toutefois, pour préciser l'interprétation que nous souhaitons donner des fonctions développées par Manheim, il est nécessaire de spécifier la nature des interactions prises en compte par le modèle.

En effet, dans une analyse de long terme (qui correspondrait au groupe C défini dans le premier chapitre, p. 31), la demande de l'individu n'est pas qu'une demande de transport, mais elle est aussi une demande de localisation. De même, les fonctions d'offre prises en compte doivent intégrer tant l'offre de transport que l'offre de localisation. Et surtout, l'état de ces fonctions n'est plus statique, mais fait partie entière de la dynamique du système, de la recherche de l'équilibre. En revanche, dans le groupe B, que nous privilégions dans notre analyse, l'état des sous-systèmes de transport et de localisation des activités est considéré comme statique, ou tout au moins leur évolution est exogène à la dynamique du système modélisé et correspond donc à un état 1 (en situation de référence) et à un état 2 (à l'horizon de la prévision) du système tous deux considérés comme figés. La recherche de l'équilibre, qui est un élément central de l'approche développée par Manheim, n'est donc pas de même nature dans les deux groupes. L'interprétation des fonctions sera alors bien évidemment différente, même si l'architecture générale de la théorie subsiste.

En nous plaçant maintenant dans le groupe B des interactions entre les trois sous-systèmes, nous pouvons préciser l'interprétation des graphiques précédents. La fonction d'offre permet de définir le niveau de service du sous-système des transports en fonction des caractéristiques de ce sous-système et du volume de demande. La fonction d'offre ne dépend donc pas directement des deux autres sous-systèmes de localisation des activités et des relations sociales. Elle en dépend par contre indirectement

dans la mesure où le volume de demande est défini par les caractéristiques des deux autres sous-systèmes.

La fonction de demande de déplacements permet de définir le volume de déplacements. Dans l'approche que nous avons développée dans le chapitre 1, nous avons considéré que la demande exprimée de déplacements résulte de la confrontation des trois sous-systèmes. Nous devons donc amender la formulation proposée par Manheim. Nous définirons donc la fonction de demande D par la relation suivante :

$$V = D((A, R), S) \quad (5.7)$$

où, A correspond au sous-système de localisation des activités ;

R correspond au sous-système des relations sociales ;

S est le niveau de service de l'offre de transport ;

V est le volume de trafic ou de demande observé sur le système de transport.

Cette nouvelle équation ne remet pas en cause la présentation de Manheim. Elle traduit par contre la difficulté de la formalisation du système des relations sociales qui conduit souvent à considérer la réunion des deux sous-systèmes (que nous illustrons en les réunissant sous une même parenthèse). C'est l'interprétation du système d'activités que Manheim propose dans son ouvrage (il ne faut donc pas confondre la définition que nous avons donnée pour le sous-système de localisation des activités qui correspond à l'inscription spatiale des équipements où sont réalisées les activités avec la définition retenue par Manheim du système d'activités qui correspond tout à la fois à cette inscription spatiale, mais aussi à la pratique spatialisée de ces activités par les acteurs du système urbain).

Pour déterminer l'équilibre, il est nécessaire de formaliser les fonctions d'offre et de demande.

2.3. Formalisation de la fonction d'offre : le coût généralisé

Manheim a défini la fonction d'offre J par l'équation suivante (équation 5.5, p. 149) :

$$S = J(T, V)$$

où, S est le niveau de service de l'offre de transport ;

T correspond au système de transport ;

V est le volume de trafic ou de demande observé sur le système de transport.

La fonction d'offre permet donc de déterminer le niveau de service de l'offre de transport. Ce niveau de service dépend des caractéristiques du système de transport et du volume de déplacements utilisant le système de transport. La formalisation de la fonction d'offre s'effectue presque systématiquement par le coût généralisé.

Le coût généralisé est une fonction permettant de traduire dans un indicateur unique des éléments de natures différentes. La formule générique du coût généralisé la plus souvent utilisée est la suivante :

$$C_g = P + \left(\sum_i \alpha_i T_i \right) * V_t \quad (5.8)$$

où C_g est le coût généralisé ;

P est le prix du déplacement ;

T_i est une partie élémentaire du temps nécessaire pour effectuer le déplacement ;

V_t est la valeur que l'individu attribue à son temps ;

α_i est un coefficient pondérateur du temps élémentaire T_i.

Le prix est le plus souvent un prix ressenti, qui peut être différent du prix réel. Par exemple pour un trajet automobile, l'usager ne prend bien souvent en compte que le prix de l'essence et des péages éventuels. En revanche, les autres coûts d'usage de la voiture sont généralement négligés. La détermination du prix dépendra donc du contexte et de l'objet d'étude.

Le temps de déplacement est décomposé en temps élémentaires pour tenir compte de la perception différente par les individus de ces temps. Par exemple, pour un déplacement effectué en transports collectifs, le temps de marche à pied n'est pas ressenti de manière analogue à celui de l'attente du véhicule ou encore à celui du temps de circulation dans le véhicule. Pour illustrer le propos, nous donnons à titre d'exemple la formule de coût généralisé retenu par le LET et la SEMALY pour le modèle stratégique développé sur l'agglomération lyonnaise (Lichère, Raux, 1997a, b) :

Temps généralisé = temps en véhicule + 2 * temps d'accès (origine + destination) + 1,8 * (coefficient de régularité) * intervalle + (pénalité d'accès)

Dans cet exemple, le temps généralisé correspond à la somme des temps élémentaires pondérés par les coefficients α_i . Le temps en véhicule n'est pas pondéré et constitue en quelque sorte la référence par rapport à laquelle les autres temps sont mesurés. Le temps d'accès qui correspond au temps de marche à pied pour accéder à l'arrêt est pondéré par un facteur 2. Il est donc ressenti comme deux fois plus pénible que le temps de circulation dans le véhicule par les individus dans leur processus de choix entre plusieurs itinéraires. Le temps d'attente du véhicule est obtenu par l'intervalle de passage entre deux véhicules multiplié par un coefficient de régularité de la ligne qui permet de tenir compte du plus ou moins grand respect des horaires. Le temps d'attente est pondéré par un facteur 1,8. Enfin, les auteurs ajoutent une pénalité d'accès qui permet de distinguer les lignes selon leur attractivité.

La valeur des coefficients dépend bien évidemment de la formulation retenue pour le coût généralisé, mais aussi du contexte d'étude. Il est donc souhaitable que les coefficients soient calés pour l'étude considérée. Ce calage peut être effectué sur des données issues d'enquêtes de préférences révélées (enquête ménages déplacements ou enquêtes transports collectifs par exemple) ou de préférences déclarées.

La valeur du temps permet de sommer des données monétaires et des données temporelles. Le coût généralisé étant censé représenter le niveau du service offert à l'usager, il est souhaitable de retenir une valeur comportementale plutôt qu'une valeur tutélaire. La détermination de cette valeur est toutefois assez délicate et fait l'objet d'une littérature abondante (Ortuzar, Willumsen, 1994 ; Ségonne, 1998 ; Jara-Diaz, 2000 ; Gunn, 2000...)

Cette formule générique ne prend en compte que le temps et le prix du déplacement. Elle peut toutefois être étendue pour prendre en compte d'autres dimensions relatives aux caractéristiques du système de transport. Nous donnons deux exemples pour l'illustrer.

Dans nombre d'études de la SNCF, pour la modélisation des trafics intercités, la formule suivante est utilisée (Bonnel, Klein, 1999) :

$$C_g = p + h * (t + 0,5 * I + 0,5 * \left(\frac{t}{f}\right)^2 + c) \quad (5.9)$$

où p est le prix du billet ;

t est le temps de trajet ;

f est la fréquence des trains sur la relation ;

I est l'intervalle moyen entre deux trains, obtenu par :

$$I = \frac{a-t}{f-1}$$

a étant l'amplitude horaire de la journée ;
r est le nombre de changements sur l'ensemble de la desserte ;
c est une constante représentant le temps des trajets terminaux ;
h est la valeur du temps moyenne des voyageurs.

Dans une étude pour le « *London Transport Buses* » (Ampt, Swanson, 1996), une enquête de préférences déclarées a été réalisée à l'aide de dessins représentant différents niveaux de qualité pour certains attributs de qualité de service des transports collectifs. Cette enquête permet de quantifier la perception de ces différents éléments de la qualité du service. Les résultats permettent ensuite d'intégrer ces attributs dans une fonction de coûts généralisés.

Les différents exemples donnés jusqu'ici ne semblent pas correspondre totalement à la formulation proposée par Manheim. Celui-ci prend en compte les caractéristiques du système de transport, mais aussi le volume de la demande utilisant ce système dans la fonction d'offre. Or, jusqu'ici, les formulations proposées pour les coûts généralisés (équations 5.8 ou 5.9) ne tiennent apparemment pas compte des volumes de déplacements. En fait, les volumes de déplacements sont généralement pris en compte dans le calcul du temps de parcours en véhicule, tout au moins en ce qui concerne l'automobile.

Les temps de parcours sont calculés à partir de lois appelées « lois débit-vitesse » ou « lois débit-coûts » qui donnent la vitesse des véhicules ou le coût (ou le temps) du déplacement en fonction de la capacité de la voie et du trafic sur cette voie. Nous ne développons pas ces lois qui seront détaillées dans le chapitre sur l'affectation du trafic (chapitre 10, p. 321). Elles permettent de tenir compte des volumes de déplacements sur le système de transport. Dans certains modèles, les temps de parcours des bus tiennent également compte de la vitesse de circulation des automobiles sur voirie (Lichère, Raux, 1997a, b).

Il est important de souligner les limites de cette formalisation. Même si des éléments de confort peuvent être intégrés dans la fonction de coût généralisé, leur estimation reste délicate. De plus, il semble difficile de prendre en compte l'ensemble des éléments entrant dans la perception de l'offre par les individus. Pour certains individus, l'attrait esthétique, touristique ou fonctionnel d'un itinéraire pourra être déterminant dans la perception de l'itinéraire. Pour d'autres ce sera la possibilité de faire d'autres activités comme lire le journal... La fonction de coût généralisé ne traduit donc qu'une partie (que l'on estime toutefois prépondérante) des éléments entrant dans la perception de l'offre.

L'habitude et l'information sont également des aspects importants. L'habitude de certains itinéraires conduit généralement les individus à sous-estimer le coût généralisé des itinéraires correspondants. Inversement, pour les itinéraires inconnus, les coûts généralisés sont généralement surestimés. Par ailleurs, on suppose généralement que les individus connaissent l'ensemble de l'offre de transport dans toutes ses composantes. Cette hypothèse est évidemment forte. Dans la pratique, il est fréquent que certaines composantes de l'offre ne soient pas du tout considérées par les individus, ce qui revient de fait à leur attribuer un coût généralisé quasi-infini.

La formalisation mathématique de la fonction de coût généralisé est également problématique. L'additivité de l'ensemble des éléments entrant dans la fonction de coût n'est évidemment pas évident, car il n'est pas sur que les individus disposent ainsi d'un référent unique permettant d'assurer la conversion de l'ensemble des composantes du coût généralisé.

2.4. Formalisation de la fonction de demande

L'approche néoclassique du comportement du consommateur introduit le concept d'utilité. Cette notion est centrale dans cette analyse. Elle découle des hypothèses de cette théorie, comme nous l'avons mentionné au début de ce chapitre : « *Le consommateur type est supposé rationnel. L'unique*

motivation du consommateur, lorsqu'il acquiert ou vend des biens et services sur le marché, est de maximiser sa satisfaction ou, en d'autres termes, son utilité. » (Brémond, 1990).

La détermination de la fonction de demande nécessite de connaître la fonction d'utilité du consommateur (graphique 5.3). En effet, nous avons montré que la courbe de demande correspond à la valeur maximum de l'utilité pour le consommateur en fonction des paramètres du bien consommé. L'utilité pour un individu donné dépend donc des caractéristiques de cet individu et des caractéristiques relatives au bien considéré. L'utilité s'écrit donc :

$$V_{in} = f(X_{in}) \quad (5.10)$$

où V_{in} est l'utilité du bien i considéré pour l'individu n ;

X_{in} correspond au vecteur des variables permettant d'estimer l'utilité du bien i considéré pour l'individu n .

Le problème est que l'utilité du consommateur pour chacun des biens n'est pas connue. Elle constitue ce que Bonnafous (1989) appelle une chimère ou un concept pseudo-opérateur. Nous avons déjà indiqué que par simplification, nous nous limitons aux biens considérés dans le processus de choix. Ne connaissant pas la forme précise de la fonction d'utilité, il est donc nécessaire de faire des hypothèses sur sa forme fonctionnelle, pour chacun des biens pris en compte dans l'analyse. Le recours à des données d'enquêtes permet ensuite de caler les coefficients inconnus des fonctions. On utilise généralement une forme additive (éventuellement après transformation de variables) qui permet de simplifier les procédures d'estimation des coefficients. La forme générique est alors :

$$V_{in} = \sum_k \beta_{kin} X_{kin} \quad (5.11)$$

où V_{in} est l'utilité du bien i considéré pour l'individu n ;

X_{kin} correspond aux différentes variables k permettant d'estimer l'utilité du bien i considéré pour l'individu n ;

β_{kin} est le coefficient de la variable X_{kin} .

Il est donc nécessaire de préciser la nature des variables X_{kin} . Comme nous l'avons indiqué ci-dessus, ces variables correspondent à certains attributs du déplacement (notamment le coût généralisé ou bien les variables entrant dans le calcul du coût généralisé) et à certaines caractéristiques socio-économiques des individus ou des ménages auxquels ils appartiennent. Le choix des variables dépend bien évidemment de la nature du choix qui doit être modélisé. Il dépend également du modèle comportemental choisi. Le choix des variables dépend donc du modèle explicatif théorique retenu par l'analyste (chapitre 2, p. 33).

Un exemple permettra d'illustrer le propos. Nous l'empruntons à Clément (CERTU, 1998b). Dans le cadre d'une étude visant à construire un modèle désagrégé de choix combinés destination-mode de transport pour Lyon, diverses fonctions d'utilité ont été construites relatives à différents segments de demande (l'exemple choisit ne correspond pas une formulation déterministe de l'utilité, nous nous limitons donc ici à la partie déterministe de l'utilité correspondant à l'équation 5.10 (section 2.4.1 ci-dessous)). A titre d'exemple, l'expression de l'utilité du mode VP vers une zone du centre de l'agglomération pour un ménage possédant 2 voitures ou plus et pour un temps en voiture supérieur à 10 minutes s'exprime ainsi pour le motif domicile-travail en heure de pointe du soir :

$$V = 1,541 - 0,321 - (0,07099 * \text{temps de rabatement en VP}) - (0,1528 * 10 \text{ mn}) - (0,08447 * (\text{temps de trajet VP} - 10 \text{ mn})) + 0,9085 * \ln[\text{population} + (\exp(3,385) * \text{emplois})]$$

Cette formule mérite quelques explications :

- le premier coefficient 1,541 est une constante liée à l'équipement automobile du ménage ;
- le coefficient -0,321 est une constante liée au choix d'une destination dans le centre de l'agglomération ;

- le coefficient 0,07099 pondère le temps de rabattement en VP c'est-à-dire le temps de marche nécessaire pour accéder à la voiture (on a généralement une valeur supérieure pour ce coefficient comparativement à ceux en véhicule, ce que n'a pas confirmé le calage sur les données disponibles) ;
- le coefficient 0,1528 pondère les 10 premières minutes passées dans le véhicule ;
- le coefficient 0,08447 pondère le temps en véhicule au-delà de ces 10 minutes. Cette distinction a été introduite pour tenir compte de la perception relativement plus négative des temps de déplacements courts en voiture comparativement aux temps plus longs. On peut interpréter cette différence par une sorte de coût fixe lié au fait de prendre sa voiture ;
- les paramètres suivants tiennent compte de la population et des emplois de la zone de destination. Plus la population et surtout les emplois sont importants, plus la zone est attractive et donc l'utilité est grande.

D'autres fonctions d'utilité ont été construites pour les zones non centrales, pour les ménages possédant moins de voiture et pour les autres motifs de déplacements. Dans cet exemple, les variables socio-économiques caractérisant l'individu n'apparaissent pas explicitement. En fait, il n'y a qu'une variable prise en compte directement qui concerne le taux d'équipement automobile. Elle n'est toutefois pas présente dans la fonction d'utilité car elle est utilisée pour segmenter la population. La fonction d'utilité est différente selon le nombre de voitures à disposition du ménage. Les autres caractéristiques de l'individu interviennent implicitement à travers les motifs. Comme pour l'équipement automobile, la fonction d'utilité diffère selon le motif du déplacement.

Il est probable que les déterminants retenus dans l'exemple ci-dessus ne recouvrent pas la totalité des déterminants « expliquant » le choix combiné du mode et de la destination par l'ensemble des individus. C'est le cas de tout modèle qui par définition n'est qu'une approximation de la réalité (chapitre 2). C'est notamment pour cette raison, que l'utilité n'est pas toujours définie uniquement dans sa composante déterministe, mais également dans sa composante probabiliste.

2.4.1. Approche déterministe ou stochastique

Dans les sections précédentes, nous avons considéré l'utilité comme une fonction déterminée en totalité par les variables et les coefficients la composant. Cette approche est qualifiée de déterministe. Elle a toutefois été remise en cause pour faire face aux problèmes posés dans la confrontation de la théorie aux données empiriques. Ben-Akiva (Ben-Akiva, Lerman, 1985) et McFadden (Domencich, McFadden, 1975) situe les prémisses de cette remise en cause dans les travaux développés en psychologie (Luce, Suppes, 1965). « *Le développement de ces théories provient de la nécessité d'expliquer les observations de laboratoire mettant en évidence l'inconsistance et la non transitivité des choix* » (Ben-Akiva, Lerman, 1985). Les auteurs (Luce et Suppes) ont alors proposé un mécanisme de choix probabiliste.

« On peut argumenter que le comportement humain est intrinsèquement probabiliste. On peut aussi argumenter, cependant, que l'hypothèse du comportement probabiliste traduit l'acceptation par l'analyste d'une connaissance insuffisante du processus de décision de l'individu. S'il était possible de connaître les causes de cette inconsistance, la théorie du choix déterministe développée dans la section précédente pourrait être utilisée. Ces facteurs sont cependant en général inconnus ou connus mais non mesurables » (Ben-Akiva, Lerman, 1985).

« Un autre type d'inconsistance apparaît dans les travaux empiriques lorsque l'on observe les choix d'un échantillon d'individus. On peut observer que deux individus ou plus placés dans des situations identiques, ayant des caractéristiques socio-économiques identiques effectuent des choix différents. ... Le mécanisme probabiliste peut être utilisé pour expliquer les effets de variations non observées entre les décideurs ou entre les attributs des alternatives proposées. Il permet de prendre en compte un comportement effectivement probabiliste tout autant qu'une perception incorrecte de certains attributs de l'alternative choisie. » (Ben-Akiva, Lerman, 1985).

On se trouve donc face à deux hypothèses dichotomiques correspondant à un comportement intrinsèquement probabiliste ou bien à une apparence probabiliste relative à notre incapacité à identifier ou mesurer les facteurs pertinents expliquant le comportement humain. Pour répondre à ces deux hypothèses Luce et Suppes (1965) ont distingué deux approches permettant d'introduire un mécanisme de choix probabiliste : l'utilité constante et l'utilité aléatoire.

2.4.1.1. Approche probabiliste avec utilité constante

Dans cette approche l'utilité reste constante. On retient donc l'équation 5.10 (p. 155) pour exprimer l'utilité V . En revanche, l'individu a un comportement probabiliste que l'on exprime en indiquant non pas un choix déterministe de l'alternative conduisant à la maximisation de son utilité, mais en indiquant un choix probabiliste.

Dans l'approche déterministe, le choix découle directement de la valeur des utilités V_{in} des alternatives appartenant à l'ensemble C_n des alternatives disponibles pour l'individu n :

$$\begin{aligned} & \text{L'alternative } i \text{ est choisie si et seulement si :} \\ & \forall j \neq i \in C_n, V_{jn} \leq V_{in} \end{aligned} \quad (5.12)$$

Par contre dans l'approche probabiliste, la connaissance des utilités respectives des alternatives disponibles ne suffit plus à déterminer le choix de l'alternative. Il faut faire des hypothèses supplémentaires pour pouvoir calculer la probabilité $P_n(i)$ de choix de l'alternative i dans l'ensemble C_n des alternatives disponibles pour l'individu n (dans la suite, nous noterons également cette probabilité $P(i|C_n)$). Nous revenons sur cette approche dans le prochain chapitre (chapitre 8, p. 229).

2.4.1.2. Approche probabiliste avec utilité aléatoire

La formalisation de cette approche est attribuée à Manski (1977) d'après Ben-Akiva et Lerman (1985). Il justifie son approche par une analyse des facteurs expliquant le caractère aléatoire de l'utilité (section 2.4.2 ci-dessous).

$$\begin{aligned} & \text{L'utilité est alors définie par une partie déterministe et une partie aléatoire :} \\ & U_{in} = V_{in} + \varepsilon_{in} = f(X_{in}) + \varepsilon_{in} \end{aligned} \quad (5.13)$$

En général, on utilise une fonction f du vecteur X_{in} additive qui conduit à la formulation suivante :

$$U_{in} = V_{in} + \varepsilon_{in} = \left(\sum_k \beta_{kin} X_{kin} \right) + \varepsilon_{in} \quad (5.14)$$

où ε_{in} est un vecteur aléatoire correspondant au terme d'erreurs.

Par convention, la notation V représente la partie déterministe de l'utilité, tandis que U correspond à la somme de la partie déterministe de l'utilité et du terme aléatoire.

Considérons un contexte de choix entre deux alternatives 1 et 2. Dans l'approche stochastique avec utilité aléatoire, nous n'avons plus un choix exclusif, mais une probabilité de choix (pour simplifier les notations, nous abandonnons l'indice n de l'individu) :

$$P(1) = P(U_1 \geq U_2) = P(V_1 + \varepsilon_1 \geq V_2 + \varepsilon_2) = P(V_1 - V_2 \geq \varepsilon_2 - \varepsilon_1) \quad (5.15)$$

où $P(1)$ correspond à la probabilité que l'alternative 1 soit choisie.

La connaissance de la partie déterministe de l'utilité ne suffit plus à la résolution de l'équation. Il est également nécessaire de faire des hypothèses sur le terme aléatoire ε . Le choix d'une distribution statistique pour le terme d'erreur détermine alors la solution (nous abordons ces distributions et les hypothèses correspondantes dans le chapitre 9).

2.4.2. Interprétation du terme aléatoire ε

Le terme aléatoire ε est souvent appelé le terme d'erreurs. Il est important de savoir à quoi il correspond. Il est possible de classer les erreurs de la manière suivante (nous reprenons la classification présentée au chapitre 2, p. 50) :

- erreurs relatives à la construction théorique. Nous avons tout d'abord les différentes hypothèses sous-jacentes à la théorie néoclassique du choix du consommateur. Nous avons ensuite le choix des variables entrant dans la fonction d'utilité. Il est évident qu'il n'est pas possible, ne serait-ce que pour des raisons d'estimation, de prendre en compte l'ensemble des variables pouvant avoir un rôle dans le processus de choix de l'individu. L'estimation de l'utilité est donc entachée d'« erreurs » ;
- erreurs relatives à la spécification des relations. Nous retenons généralement une forme additive (éventuellement après transformations de variables) pour l'utilité. Il n'est pourtant pas certain que la contribution effective des variables soit toujours additive. Quelle que soit la forme choisie, la connaissance de l'objet d'étude est généralement insuffisante pour valider complètement les choix effectués, au-delà d'une concordance plus ou moins grande avec les données disponibles. La seconde dimension a trait au choix de la fonction de distribution pour les résidus et aux hypothèses nécessaires sur les résidus (nous développons cette seconde dimension dans le chapitre 9, p. 248) ;
- erreurs relatives à la production des données. La mesure réalisée lors d'une enquête est toujours entachée d'erreurs. L'erreur d'échantillonnage peut être calculée lorsque les hypothèses nécessaires à son calcul sont respectées. En fait, à ces erreurs évaluables, il faut en ajouter d'autres (base de sondage incomplète, refus de réponse, individus non contactés, questions mal comprises, qualité insuffisante du questionnaire, biais liés à l'enquêteur ou à l'enquêté...) (Bonnel et al., 1997, chapitre 4, p. 112) ;
- erreurs de calage. L'estimation des coefficients inconnus s'effectue à l'aide de procédures statistiques nécessitant le respect d'un certain nombre d'hypothèses statistiques qui n'est pas toujours le cas. Elles constituent ce que l'on appelle les biais d'estimation.

Certaines de ces erreurs (modèle comportemental, production des données) ont déjà été évoquées ou le serons dans les chapitres suivants (spécification des fonctions, calage). Nous abordons ci-dessous certaines des hypothèses nécessaires à la construction théorique néoclassique du choix du consommateur. Nous évoquons ici les principales hypothèses (relativement à notre préoccupation de modélisation), qui découle de la construction de l'homo oeconomicus doté de la rationalité économique :

- l'homo oeconomicus est cohérent. Cette cohérence se traduit notamment par le respect de la transitivité: si $U_a > U_b$ et si $U_b > U_c$ alors $U_a > U_c$;
- l'homo oeconomicus est consistant. Placé face à une situation identique, il fait toujours le même choix ;
- l'homo oeconomicus est informé. Pour être capable d'évaluer toutes les options qui s'offrent à lui, il doit les connaître de manière détaillée, ou tout au moins dans les composantes qui sont prises en compte dans les fonctions d'utilité ;
- l'homo oeconomicus est doté d'un référent unique. L'individu doit être capable de traduire l'ensemble des dimensions d'une alternative selon un même référent unique pour pouvoir comparer les alternatives entre elles ;
- l'homo oeconomicus est calculateur. Il doit savoir réévaluer en permanence l'ensemble des alternatives qui se présentent à lui chaque fois que l'une d'elles évolue ;

- l'homo oeconomicus réagit continûment. Il doit être capable de réagir de manière continue face à une modification marginale de l'une des alternatives. Cette condition est nécessaire pour pouvoir dériver les fonctions d'utilité.

Il est facile de montrer que cette rationalité supposée des individus est prise en défaut (c'est le cas de chacune des dimensions évoquées ci-dessus, certaines expériences menées en économie expérimentale permettent ainsi de rejeter une à une les hypothèses ci-dessus). Ce n'est pas que les individus ne soient pas rationnels, mais leur rationalité n'est pas forcément celle de l'homo oeconomicus telle que l'on est conduit à la définir pour la modélisation.

Si nous insistons sur ces limites, ce n'est pas pour rejeter le modèle proposé (la simplification est inhérente au modèle, comme nous l'avons maintes fois mentionné), mais pour attirer l'attention sur les conditions nécessaires à l'application de la théorie. Il existe toutefois des solutions pour contourner certaines de ces limites lorsqu'elles apparaissent trop fortes au regard de l'objet d'étude. Il n'est pas possible de les énumérer toutes. Nous citons juste un exemple à titre d'illustration.

L'homo oeconomicus réagit continûment. Il est évident que cette hypothèse n'est pas vérifiée face à des modifications de l'offre de transport aussi importante soit-elle. La mise en service d'une nouvelle infrastructure s'accompagne généralement d'une montée en charge progressive pouvant même s'étaler sur une durée très longue (Goodwin, 1984). Il semble en fait que les changements loin d'être continus se réalisent de manière privilégiée à l'occasion de changements dans le cycle de vie (Bourgin, 1978). Pour faire face à cette infraction à notre hypothèse, la prévision de la demande sur une nouvelle infrastructure s'accompagne fréquemment d'une hypothèse de montée en charge progressive. Le problème n'est donc pas résolu, mais simplement contourné.

2.5. Réflexions sur l'équilibre offre-demande appliqué à la modélisation de la demande

Dans la théorie néoclassique du choix du consommateur, la notion d'équilibre est centrale. C'est l'équilibre qui permet de définir le niveau optimal de production pour le producteur et le niveau optimal de consommation qui maximise l'utilité des consommateurs. Cet équilibre dépend des fonctions d'offre des producteurs et des fonctions de demande des consommateurs.

Dans la transposition que nous avons exposée, la notion d'équilibre est également fondamentale. Il ne s'agit toutefois pas exactement du même équilibre. En effet, la fonction d'offre est figée. Plus précisément, c'est l'offre qui est figée à chacun des horizons d'étude. Un retour sur le schéma 1.2 (p. 31) permet d'explicitier ce constat. Notre analyse des interactions entre les trois sous-systèmes, nous conduit à ne retenir que les actions des sous-systèmes des localisations et des transports sur le troisième sous-système. De plus, nous avons précisé que l'évolution des deux premiers sous-systèmes est exogène à l'exercice de prévision. L'offre de transport est donc figée en situation de référence, tout comme en prévision (pour chacun des scénarios d'offre de transport étudiés). Il en est de même du sous-système des localisations.

En revanche, si nous nous étions placé dans le groupe C des interactions (schéma 1.3, p. 32), il aurait été nécessaire de considérer une offre de transport évolutive, comme élément endogène de notre modèle, de même pour les localisations. La fonction d'offre de transport serait donc différente de celle présentée dans cette section. Le choix des interactions prises en compte et de la nature de ces interactions conditionne donc le choix des fonctions de demande et d'offre.

Si l'offre de transport est figée, cela ne signifie pas que le niveau de service est constant. La fonction d'offre que nous avons définie à travers la notion de coût généralisé peut tenir compte du niveau de la demande. C'est généralement le cas des fonctions d'offre relatives à l'usage de la voiture pour tenir compte des phénomènes de congestion.

La notion d'équilibre reste donc fondamentale dans cette approche. D'un point de vue opérationnel en modélisation, on considère que la situation de référence (c'est-à-dire celle issue des enquêtes) correspond à une situation d'équilibre. La connaissance de cet équilibre nous permet de caler le modèle. L'utilisation du modèle en prévision conduit à modifier certains paramètres d'entrée du modèle (relatifs à l'un des trois sous-systèmes). On suppose que l'on débouche alors sur une nouvelle situation d'équilibre.

Au-delà des problèmes de formalisation et d'estimation des modèles qui font l'objet des chapitres suivants (chapitres 6 à 10), certains auteurs critiquent cette notion d'équilibre (Goodwin, 1998). Pour eux, la mesure de la situation de référence ne correspond pas à une situation d'équilibre mais à certain état transitoire d'un système en évolution constante. En d'autres termes, la mesure qui est faite ne correspond pas à un état stabilisé du système. Le système est en évolution constante, car la réponse des consommateurs n'est pas immédiate face à une modification des paramètres du système. Goodwin (1998) donne ainsi une fourchette de temps de réponse face à différents éléments d'une politique urbaine allant de quelques jours à 50 ans ou plus....

Il n'en reste pas moins que la plupart des modèles présupposent cet équilibre. Il est possible de classer les modèles selon la manière dont ils posent les termes de cet équilibre.

3. CLASSIFICATION DES « MODELES A QUATRE ETAPES »

Nous limitons notre classification aux grandes catégories génériques de modèles. Dans les chapitres suivants, nous approfondirons la classification en nous appuyant sur les principales formes fonctionnelles de modèles.

La première distinction qu'il est coutume d'opérer concerne la structure de choix de l'individu. On distingue alors les modèles séquentiels qui décomposent le choix des caractéristiques du déplacement en une séquence de choix, des modèles directs qui considèrent ces caractéristiques comme étant intimement liées et donc ne modélisent qu'un seul choix.

La seconde distinction concerne le niveau d'agrégation. Dans l'approche désagrégée, chaque individu est considéré isolément à travers ses caractéristiques individuelles. A l'inverse, dans l'approche agrégée, on définit un individu moyen représentatif d'un groupe ou d'une classe d'individu.

La dernière distinction communément admise concerne l'approche de l'utilité : déterministe ou probabiliste.

3.1. *Modèle direct - séquentiel*

L'approche séquentielle décompose le processus de choix de l'individu en séquences. Le plus souvent, on identifie 4 séquences, d'où le nom de « modèle à quatre étapes » :

- génération. Cette étape correspond au choix de l'individu de se déplacer ou de ne pas se déplacer. L'étape de génération permet donc de définir le volume de déplacement réalisé par les individus. Le plus souvent, le résultat du modèle de génération consiste à déterminer le nombre de déplacements émis ou attirés par chacune des zones de l'aire d'étude ;
- distribution. Cette étape correspond au choix de la destination par l'individu. A l'étape précédente, il a décidé de se déplacer à celle-ci, il décide de sa destination. Le résultat de cette étape consiste à construire la matrice origines-destinations des déplacements (alors que la précédente avait permis de calculer les marges de cette matrice) ;
- répartition modale. Cette étape correspond au choix du mode de transport par l'individu pour réaliser son déplacement entre une origine et une destination. Le résultat de cette étape

consiste à construire autant de matrices origines-destinations des déplacements qu'il y a de modes de transport pris en considération ;

- affectation. Cette étape correspond au choix de l'itinéraire par l'individu pour réaliser son déplacement entre une origine et une destination avec un mode donné. Le résultat de cette étape fournit la charge de trafic sur chacun des tronçons du réseau ou sur chacune des lignes de transports collectifs.

Malgré leur dénomination de modèle à quatre étapes, il y a en fait souvent d'autres étapes qui s'ajoutent à celles-ci. Les principales sont :

- la conversion véhiculaire. Elle permet de passer pour le mode voiture d'une matrice de déplacements d'individus à une matrice de déplacements de véhicules avant la phase d'affectation. La conversion correspond simplement à la prise en compte du taux d'occupation des véhicules. Cette phase prend place entre la répartition modale et l'affectation (chapitre 10, p. 327) ;
- le passage à la pointe. L'estimation de la demande est souvent faite pour la journée. Pourtant, ce sont les données de la période de pointe qui servent au dimensionnement des infrastructures. Dans ce cas, une phase de passage à la pointe peut être définie. Elle prend place le plus souvent après la distribution ou la répartition modale. A l'inverse, l'estimation se limite parfois à la période de pointe, alors que l'on peut avoir besoin de données à la journée ou à l'année pour l'évaluation économique. Dans ce cas les données doivent être projetées à la journée ou à l'année, le plus souvent après l'affectation. Dans certains modèles, le choix de l'heure de départ est explicitement modélisé (Mahmassani, 2000 ; De Palma, Fontan, 2001) (chapitre 10, p. 328) ;
- l'estimation du trafic d'échange et de transit. Le plus souvent les quatre étapes présentées ci-dessus sont réalisées à l'intérieur d'un périmètre d'étude. Toutefois, ce périmètre ne peut constituer un isolat complet. Il entretient des échanges avec les zones externes ou il peut être simplement traversé par des flux de déplacements (transit). L'estimation du trafic d'échange et de transit fait souvent l'objet d'une phase spécifique (on fait alors le plus souvent appel aux données des enquêtes cordon, chapitre 7, p. 194).

A l'inverse, dans certains modèles, certaines phases sont regroupées. C'est le cas, par exemple, de modèles combinant le choix du mode et de la destination (CERTU, 1998b). Par ailleurs, en France, les modèles utilisés ne comportent pas toujours toutes les étapes. C'est le cas des modèles appliqués dans les logiciels DAVIS, développés par l'INRETS, ou TERESE, développé par la SEMALY qui mettent en œuvre des modèles d'affectation de la demande. Ces logiciels étaient fréquemment utilisés en France jusqu'à une date récente pour les études de prévision de trafic. Aujourd'hui, l'utilisation de logiciels multimodaux se multiplie (DAVIS est aujourd'hui remplacé par DAVISUM, TERESE continue à être utilisé mais est parfois combiné à d'autres modules réalisant les autres étapes).

Il est évident que cette approche séquentielle pose des problèmes de pertinence. Il est peu probable que l'individu décompose effectivement son processus de choix en étapes telles qu'elles sont modélisées. Il est probable au contraire qu'il considère simultanément l'ensemble des dimensions. Ces modèles posent également quelques problèmes de cohérence interne (nous y reviendrons dans les chapitres suivants). Pour faire face à ces problèmes, le modèle direct considère l'ensemble des éléments du choix simultanément en une seule équation. Conformément à la problématique des modèles (chapitre 2, p. 48), le modèle direct pose de redoutables problèmes de mesurabilité pour l'estimation de ces coefficients. On trouve peu d'exemples d'applications de modèles directs du fait de la complexité de la procédure d'estimation. La plupart des modèles développés s'appuie sur la structure générale des modèles à quatre étapes.

3.2. Modèle agrégé - désagrégé

Dans l'approche agrégée, l'individu pris en compte est un individu moyen. Il s'agit donc d'un individu construit et fictif doté de caractéristiques moyennes, généralement les moyennes des caractéristiques des individus de sa zone. A l'inverse dans l'approche désagrégée, l'individu est considéré isolément à travers ses propres caractéristiques. A priori, l'approche agrégée apparaît donc moins pertinente. Elle peut être perçue comme une perte délibérée d'informations.

Un petit schéma illustre simplement ce constat (schéma 5.1).

L'individu moyen ainsi défini traduit la réalité moyenne, mais bien peu les situations individuelles. Il n'est donc pas certain que les choix qu'il effectuera seront identiques à la moyenne des choix effectués par chacun des individus. C'est pour tenter de réduire un peu la dispersion autour de l'individu moyen, que l'on introduit l'approche agrégée catégorielle (schéma 5.2). Dans ce cas, on ne définit plus un seul individu moyen, mais on détermine des classes les plus homogènes possibles au regard de l'objet d'étude. Ensuite, pour chacune de ces classes, on définit un individu moyen. Le problème subsiste donc, mais son intensité est réduite.

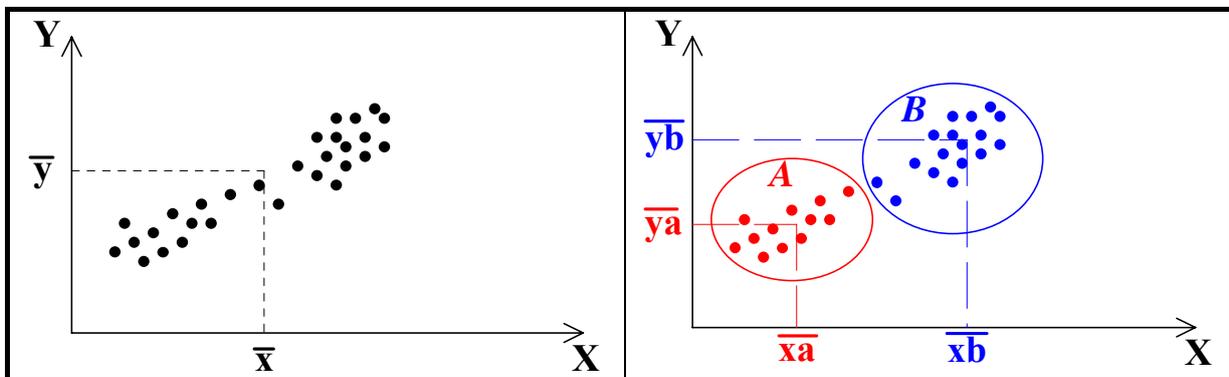


Schéma 5.1 : Approche agrégée

Schéma 5.2 : Approche agrégée catégorielle

En revanche, dans l'approche désagrégée, tous les individus sont considérés isolément. L'approche agrégée s'apparente alors à une perte d'informations « délibérée » lorsque l'on dispose de données d'enquête individuelle (schéma 5.3).

Apparemment l'approche désagrégée est à privilégier dans la mesure où elle est censée assurer une plus grande pertinence de la modélisation. Dans les faits, la réalité n'est pas toujours aussi simple pour trois raisons principalement.

La première découle des hypothèses accompagnant les modèles désagrégés et notamment l'hypothèse sur la distribution du terme aléatoire dans la détermination de l'utilité. La confrontation de ces hypothèses avec les données empiriques impose bien souvent des structures de modèles plus complexes (chapitre 9, p. 248). L'estimation de ces modèles peut alors devenir très complexe et l'interprétation des résultats peut nécessiter des compétences assez pointues, notamment en statistiques. Le risque est alors grand d'avoir une validation et surtout une maîtrise insuffisante du modèle, qui peuvent avoir de lourdes conséquences en prévision. Pour certains objectifs, l'utilisation d'outils relativement simples comme des modèles agrégés unimodaux peut s'avérer tout à fait satisfaisant sur le plan de l'opérationnalité (avec des prévisions de trafic assez proches de l'observé (GESMAD, 2000)).

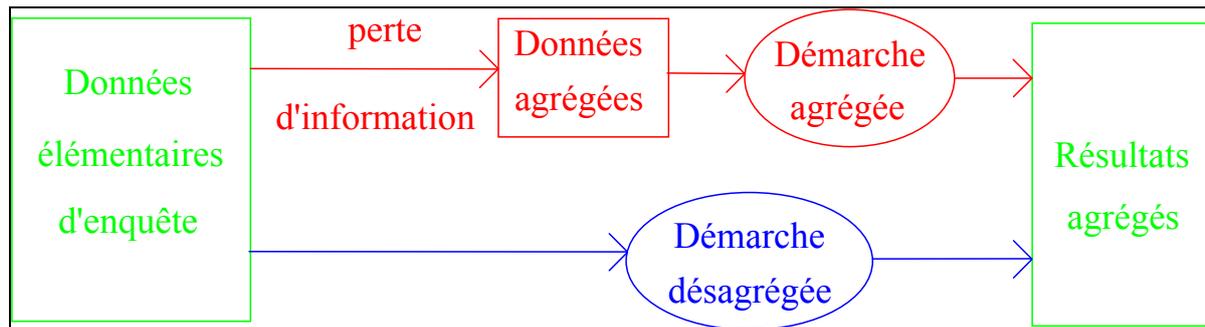


Schéma 5.3 : Démarche agrégée / désagrégée
(Source : CETE de Lyon et al., 1986)

La seconde tient aux limites du zonage que l'on a déjà mentionnées (chapitre 4, p. 127 et 132). Lorsque l'on utilise un zonage, la localisation des extrémités du déplacement dans la zone est inconnue. De ce fait, on définit un centroïde de zone, point « fictif » qui regroupe l'ensemble des localisations de la zone. Les données de localisation sont donc tout autant agrégées dans les modèles agrégés que désagrégés (Supernak, 1983). Cette codification par rapport à une connaissance des localisations en coordonnées x-y peut aussi être considérée comme une perte d'informations délibérée. Elle est commune aux approches agrégées et désagrégées (seule l'approche totalement désagrégée développée par Chapleau à partir des enquêtes de Montréal s'en démarque (Chapleau, 1983, 1991, 1997, 1999 ; Dorland et al., 1997)). De ce fait, des informations comme la distance d'accès aux différentes lignes de transports collectifs sont définies de façon moyenne pour les individus de la zone. Il n'est pas exclu que cette approximation soit au moins aussi importante que l'agrégation des caractéristiques socio-économiques des individus, au moins pour certaines étapes comme la répartition modale.

La troisième concerne la prévision avec un modèle désagrégé. Si les données sont disponibles à un niveau désagrégé pour l'estimation du modèle, il est rarement possible d'utiliser le même niveau de désagrégation pour la prévision. L'analyste doit procéder à un certain niveau d'agrégation (chapitre 9, p. 293).

Nous revenons sur cette discussion dans la conclusion du chapitre 9 (p. 297) après avoir présenté les approches agrégées (chapitres 6 à 8) et désagrégées (chapitre 9). Dans la pratique, l'approche désagrégée s'est développée tout d'abord au niveau de l'étape de répartition modale. Cette étape reste le domaine d'application privilégié de l'approche désagrégée, même si des modèles désagrégés ont déjà été développés pour toutes les étapes.

3.3. *Modèle déterministe - stochastique*

Nous avons déjà évoqué cette distinction (section 2.4.1, p. 156). Dans l'approche déterministe le choix est effectué de manière déterministe (à partir des caractéristiques de l'individu, du déplacement...). Par contre dans l'approche stochastique, on définit une probabilité de choix pour l'individu. L'approche stochastique nécessite de choisir une loi de distribution du terme aléatoire, encore appelé le terme d'erreurs.

3.4. *Combinaison des dimensions*

Le croisement de ces combinaisons donne en théorie 8 classes de modèles. Dans les faits, on en trouve beaucoup moins. Il y a peu d'exemple de modèles directs combinant les 4 étapes utilisés de manière opérationnelle. On trouve, en revanche, des modèles combinant certaines étapes (chapitre 8, p. 236 et 237). Parmi les modèles séquentiels, les modèles agrégés sont le plus souvent déterministes,

même si la phase de choix du mode peut être interprétée comme résultant d'une approche probabiliste. Enfin, les modèles désagrégés sont le plus souvent séquentiels et stochastiques.

En France, l'utilisation dominante reste le modèle agrégé à quatre étapes, historiquement restreint à la phase d'affectation. En revanche, dans les pays anglo-saxons le recours aux modèles désagrégés est largement répandu. Les pratiques sont toutefois en train d'évoluer, en France, pour répondre à l'évolution des problématiques de transport de plus en plus multimodales.

4. CONCLUSION

Si nous avons principalement deux classes de modèles de prévision de la demande de déplacements aux fondements théoriques proches, cela n'empêche pas une grande diversité dans les modèles développés. Nous analysons cette diversité dans les chapitres suivants en présentant les principales familles de modèles pour chacune des étapes des modèles à quatre étapes :

- génération de la demande de déplacements : chapitre 6 ;
- distribution des déplacements : chapitre 7 ;
- répartition modale – approche agrégée : chapitre 8 ;
- modélisation désagrégée : chapitre 9 ;
- affectation des déplacements : chapitre 10.

Il peut apparaître surprenant de voir figurer la modélisation désagrégée au chapitre 9 au milieu des étapes de répartition modale (chapitre 8) et d'affectation (chapitre 10). Le choix correspond à l'historique du développement et de l'usage dominant des modèles désagrégés lors de l'étape de choix du mode. L'approche désagrégée peut toutefois être utilisée pour les 4 étapes de la modélisation et des modèles ont déjà été développés pour chacune de ces étapes.

Chapitre 6 – Génération de la demande de déplacements

La génération constitue la première étape des modèles à 4 étapes. Elle vise à déterminer le nombre de déplacements réalisés par les résidents du secteur d'étude. Dans la séquence des choix de l'individu relatif à réalisation et à la détermination des caractéristiques d'un déplacement, cette étape correspond au choix de se déplacer ou non de l'individu.

Dans son principe, cette étape est très simple. Elle consiste à déterminer le nombre de déplacements émis et reçus par chacune des zones du terrain d'étude en fonction d'un certain nombre de caractéristiques de la zone. La mise en œuvre l'est parfois un peu moins selon les objectifs recherchés et les données disponibles. Nous commençons par préciser certaines définitions (section 1), avant de présenter les grandes familles de modèles disponibles : la méthode normative, l'analyse par régression, le modèle catégoriel (section 2), et enfin les principales limites de cette étape (section 3).

1. DEFINITIONS ET CHOIX DE L'UNITE DE MESURE

Les modèles à quatre étapes s'appuient sur un découpage de l'agglomération en zones (nous reviendrons au chapitre 10, p. 308 sur la définition de ce zonage). La définition du zonage dépend des objectifs de l'étude, des données disponibles et de la taille de l'agglomération. D'une manière générale, le découpage comprend de quelques dizaines de zones pour des modélisations de très long terme à faible spatialisation jusqu'à plusieurs milliers de zones pour les plus grosses agglomérations (1277 zones pour le modèle MODUS de la DREIF (Direction Régionale de l'Équipement de l'Île de France, DREIF, 1998), 1500 zones pour le modèle développé sur EMME2 par le Ministère des transports sur Montréal).

1.1. Emission, Attraction

La finalité du modèle de génération est de produire les émissions et attractions par zone. Il convient donc de définir ces termes :

- les émissions correspondent à l'ensemble des déplacements émis par une zone (schéma 6.1). C'est-à-dire l'ensemble des déplacements dont l'origine se situe dans la zone. Les émissions sont également parfois appelées les productions (probablement par « traduction » du terme anglais « *trip production* » qui signifie en fait émission) ;
- les attractions correspondent à l'ensemble des déplacements reçus par une zone (schéma 6.1). C'est-à-dire l'ensemble des déplacements dont la destination se situe dans la zone.

L'étape de génération permet de constituer les marges de la matrice origines-destinations qui sera produite par l'étape suivante de distribution. Cette dernière utilise (chapitre 7) comme données d'entrée les marges de la matrice origines-destinations des déplacements qui sont produites par le modèle de génération.

Cette définition générale est en fait parfois amendée selon les objectifs de la modélisation et les principes retenus pour cette modélisation comme nous le mentionnons section 1.5 (p. 169).

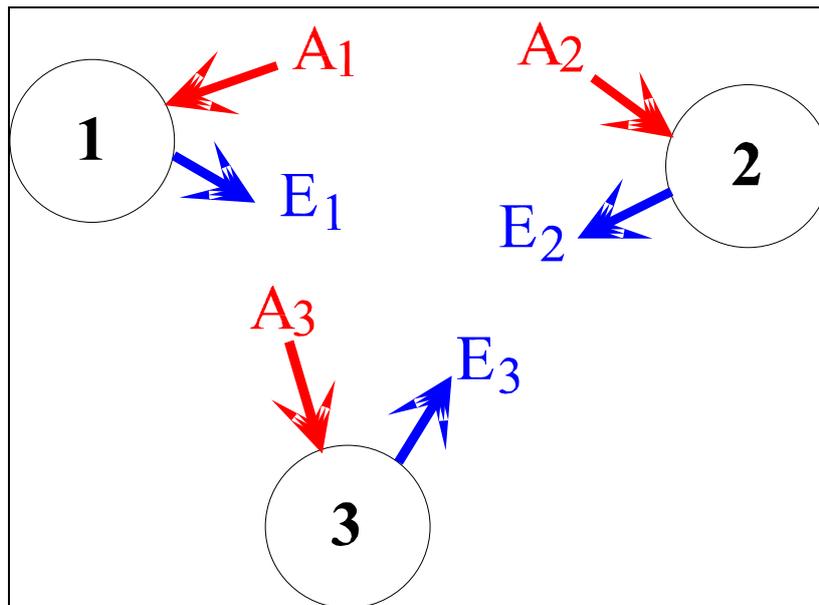


Schéma 6.1 : Génération des déplacements : émission et attraction par zone

1.2. Déplacements, activités, chaînes, sorties, schéma d'activités

La sortie habituelle d'un modèle de génération est la production d'émissions et attractions qui sont généralement mesurées en nombre de déplacements par zone. Il est donc important de préciser la définition qui est donnée des déplacements. Généralement, on est conduit à retenir la définition qui a été utilisée lors de la production des données que l'on utilise. En France, on retient donc le plus souvent la définition des enquêtes ménages déplacements : « *Le déplacement est le mouvement d'une personne, effectué pour un certain motif, sur une voie publique, entre une origine et une destination, selon une heure de départ et une heure d'arrivée, à l'aide d'un ou plusieurs moyens de transport* » (CERTU, 1998a, manuel de l'enquêteur). Si en théorie cette définition est opératoire, l'application sur le terrain n'est pas toujours très facile (ce constat n'est pas spécifique à la définition retenue par le CERTU, mais peut s'appliquer à la plupart des définitions). A partir de quelle distance ou durée considère-t-on que l'on a un déplacement ? Quelles sont les activités réalisées (ou motifs) sur un trajet principal que l'on doit prendre en compte ? Un individu qui va au travail, s'arrête en route pour déposer ses enfants, puis mettre une lettre à la boîte aux lettres, puis acheter du pain et enfin de l'essence avant d'arriver au travail, effectue un ou bien 5 déplacements ? L'attitude à adopter face à ces différentes situations est précisée dans le manuel de l'enquêteur (CERTU, 1998a). L'application stricte de la définition conduit à 5 déplacements, puisqu'il y a chaque fois activité et chaque fois mouvement sur la voie publique. Toutefois, quelle que soit la définition retenue par le concepteur de l'enquête, il y a fort à parier que les enquêtés ne retiendront pas tous la même définition, quelles que soient les qualités de relance de l'enquêteur. Les choix effectués ont quant à eux des conséquences évidentes sur le nombre de déplacements réalisés par un individu mais aussi sur les caractéristiques de ces déplacements. Prenons un autre exemple pour illustrer le problème. Un individu va à la gare en voiture, y achète le journal, puis prend le train. Si on retient un seul déplacement, il aura un mode combiné voiture + transports collectifs. Si on retient deux déplacements chaque déplacement a un mode unique et l'intermodalité au cours du déplacement disparaît.

C'est notamment pour cette raison que certains auteurs utilisent la notion de sortie dans les analyses de mobilité. L'utilisation en matière de modélisation est par contre moins fréquente, même si elle a tendance à se développer en France et surtout à l'étranger (Zhang et al., 2001 ; Lichère, Raux, 1997a). La sortie est définie par la succession des activités et déplacements inclus entre un départ et un retour successifs au domicile. Selon les auteurs, la sortie est parfois appelée boucle ou chaîne de

déplacements. Certains auteurs définissent également la chaîne de déplacement comme étant la succession des déplacements entre le domicile et le motif principal de la sortie ou entre ce motif principal et le retour au domicile. La sortie ou boucle est ainsi composée de deux chaînes de déplacements (schéma 6.2, Lichère, Raux, 1997a). L'avantage de cette définition est évidemment de ne pas être soumise aux aléas de définition du déplacement et surtout de compréhension de la définition par les enquêtés (en ce qui concerne le nombre de boucles, sorties ou chaînes). Cette approche présente également des avantages au niveau du choix du mode de transport dans la mesure où l'unité modale est fréquente pour l'ensemble des déplacements dans une sortie (par exemple, sauf cas particulier, il n'est généralement pas possible d'aller au travail en vélo et d'en revenir en voiture). Enfin et surtout, cette notion est plus stable et permet en général un meilleur calage des modèles, notamment de génération. En revanche, la modélisation de la génération des sorties ou des chaînes pose le problème du passage aux déplacements qui reste nécessaire pour l'étape d'affectation sur les réseaux.

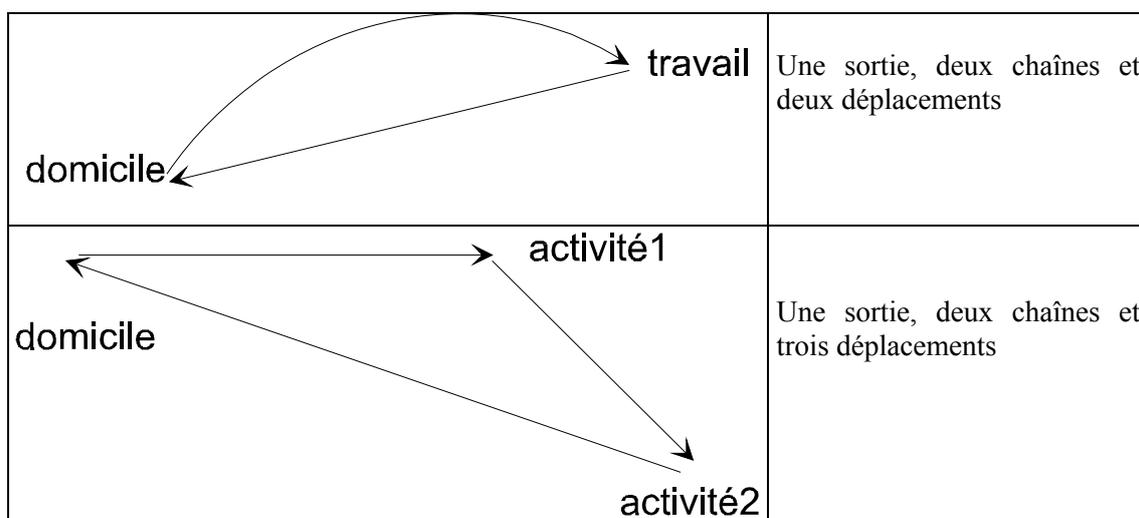


Schéma 6.2 : Déplacement, chaîne, sortie

C'est pour tirer les conséquences de ces évolutions, que la production des données a également évolué, notamment aux Etats-Unis (Bonnafous et al., 1981 ; Bonnel, 1985 ; Axhausen, 1995 ; Stopher, 1998). Dans les approches en termes de sorties, la réalisation des activités en dehors du domicile est privilégiée. La connaissance du lieu de réalisation, permet d'en déduire les déplacements lorsqu'il y a un changement de lieu. L'expérience montre que cette approche ne conduit pas aux mêmes résultats en termes de mobilité. On peut observer des différences pouvant aller jusqu'à 15 à 20 % de déplacements recueillis en plus dans les enquêtes d'activités par rapport aux enquêtes déplacements (Jones et al., 1980). C'est pour ces raisons que les enquêtes déplacements auprès des ménages aux Etats-Unis sont en général centrées sur la réalisation des activités avec des questions du genre qu'avez-vous fait ensuite ? plutôt que où êtes-vous allé ensuite ? Pour certaines enquêtes, le recueil des activités ne se limite plus à celles réalisées hors du domicile, mais intègre également celles qui se déroulent au domicile. Cette évolution répond à plusieurs objectifs et notamment :

- meilleure compréhension de l'organisation de la journée et des interactions au sein du ménage ;
- analyse de la complémentarité/concurrence entre les activités à domicile et hors domicile notamment pour les activités de loisirs ;
- prise en compte des nouvelles technologies, et notamment d'internet dans la complémentarité/substitution entre les activités au/hors domicile, notamment pour les achats avec le développement de l'e-commerce ;
- déroulement pouvant apparaître plus naturel de l'ensemble de la journée qui facilite donc la description de la journée grâce à une mémorisation plus aisée.

Compte tenu de cette information nouvelle, l'unité de mesure n'est plus la sortie du domicile composée de déplacements auxquels sont associés des motifs de déplacements, mais davantage la boucle d'activités à laquelle est associée des déplacements lors des changements de localisation des activités. La modélisation porte alors souvent sur la modélisation de séquences d'activités au sein des boucles. Ce type de modélisation conduit à des outils opérationnels comme à Portland aux Etats-Unis.

Ces quelques exemples qui pourraient facilement être développés mettent en évidence l'importance, mais aussi la difficulté de la définition de l'unité de mesure de la mobilité (ou des activités) qui sert de base à la quantification des émissions et attractions. Lors de toute étude, il est donc important de s'assurer de la définition retenue lors de la production des données, tout particulièrement lorsqu'il s'agit de combiner des données provenant de sources différentes.

1.3. Jour ou période de pointe

Dans le passé, les études de prévision de trafic avaient souvent pour finalité première le dimensionnement de nouvelles infrastructures de transport. Il était donc nécessaire de produire une matrice origines-destinations correspondant à la période de pointe de la demande de transport. Pour les transports en commun, l'intensité maximale de la pointe du matin étant proche de l'intensité maximale de la pointe du soir, celle du matin était assez souvent privilégiée, car les trafics y sont plus homogènes en termes de motifs de déplacements. La part des déplacements domicile-travail/études y est effectivement fortement prépondérante, avec de plus une durée de pointe évoluant peu, tout au moins dans les agglomérations moins congestionnées. A l'inverse, la pointe du soir s'étale de plus en plus et la part des déplacements travail/études - domicile diminue fortement au profit de déplacements d'achats/service ou de loisirs. De ce fait, la période de pointe du soir était souvent privilégiée pour la voiture (DREIF, 1998).

L'évolution des préoccupations de planification et d'évaluation socio-économiques nécessite de plus en plus des modélisations concernant l'ensemble de la journée (GESMAD, 2000), voire des périodes plus longues prenant en compte les déplacements de fin de semaine.

Au-delà des objectifs de la modélisation, le choix de la période de pointe du matin ou du soir ou de la journée a également d'autres implications en matière de modélisation. Un calage du modèle sur les déplacements de l'ensemble de la journée permet d'utiliser la totalité des données disponibles. A l'inverse un calage sur la période de pointe réduit la quantité d'informations utiles et conduit donc à une moindre précision de l'estimation. Par contre, pour les étapes suivantes, nécessitant le calcul du coût généralisé des déplacements, il est évident que les valeurs de coûts généralisés ne sont pas identiques tout au long de la journée. De plus, le sens de l'évolution est inverse pour les deux modes motorisés avec un coût généralisé accru pour la voiture en période de pointe du fait de la congestion et inversement d'un coût souvent plus faible pour les transports collectif du fait de meilleure fréquence.

1.4. Déplacements des individus ou des ménages

Les premières études de modélisation ont généralement retenu un nombre de déplacements par ménage (SETRA, 1973), tandis qu'actuellement l'unité de mesure est plutôt l'individu.

Deux raisons principales expliquaient le choix du ménage. Une partie des déplacements concerne l'entretien et le fonctionnement du ménage (achats, certains déplacements de service, les déplacements liés aux enfants). C'est donc au sein de cette composante que ce gère la répartition des déplacements. C'est également au sein du ménage que se répartissent les moyens en termes de revenus ou de motorisation dont l'influence est forte sur le nombre de déplacements motorisés. La mobilité motorisée et surtout voiture particulière étant privilégiée, la relation entre la motorisation du ménage et

le nombre de déplacements automobiles était de meilleure qualité statistique au niveau du ménage que de l'individu (SETRA, 1973, graphique 4.1, p. 140).

Si en théorie ce choix apparaît pertinent compte tenu de l'importance des interactions entre les membres du ménage et de la cellule familiale comme instance de décision ou de répartition des activités et des moyens (Jones et al., 1980 ; Andan et al., 1988), ce niveau pose également plusieurs difficultés en termes de prévision. Au cours des trente dernières années, la composition des ménages a fortement évolué avec une forte diminution de la taille des ménages. Il est relativement aisé de prévoir l'évolution de la population (tout au moins à un niveau spatial assez agrégé), l'incertitude majeure se situant au niveau de l'influence des flux migratoires. En revanche, la prévision du nombre de ménages est plus complexe car elle nécessite des hypothèses sur l'évolution dans le temps de la taille et de la composition des ménages.

D'autres limites ont également été avancées, notamment par Supernak (Supernak, 1983 ; Supernak et al., 1983) :

- la taille de l'échantillon à enquêter est plus importante pour une génération au niveau du ménage par rapport à un choix de l'individu pour une même précision statistique ;
- certains changements démographiques sont plus aisément modélisés au niveau de l'individu que du ménage, comme le vieillissement de la population par exemple ;
- une modélisation au niveau du ménage nécessite souvent une classification des ménages selon leur composition et leur taille. Si l'on souhaite introduire d'autres dimensions, le nombre de classes devient rapidement difficile à gérer.

L'ensemble de ces raisons a conduit progressivement à un abandon du ménage au profit de l'individu comme unité de mesure de la génération des déplacements.

1.5. Segmentation par motifs

Afin de renforcer le caractère explicatif des modèles de génération, les déplacements ont été assez rapidement segmentés par motif. Cette segmentation permet de relier plus aisément des caractéristiques des individus ou des zones avec le nombre de déplacements par motifs (par exemple actifs/non actifs ou nombre d'emplois pour les déplacements domicile-travail...).

La segmentation suivante est assez fréquente :

- domicile – travail ;
- domicile – école, avec parfois une distinction selon le niveau d'études (primaire, collège, lycée, université) ;
- domicile – achats/services ;
- domicile – loisirs ;
- domicile – autres motifs ;
- secondaire.

Cette décomposition met en évidence la séparation entre les déplacements ayant une extrémité au domicile et les autres déplacements qui sont appelés des déplacements secondaires. Cette segmentation permet d'affiner la génération au niveau des motifs contraints dont le volume est directement relié à la population active et à la population scolaire ou étudiante. Le gain est par contre plus limité pour les autres motifs. Les données de l'enquête ménages déplacements de Lyon montre une évolution assez nette de la répartition des motifs au cours des 20 dernières années avec la réduction du poids des déplacements contraints et l'accroissement des déplacements secondaires (tableau 6.1).

	Enquête ménages déplacements 1976 (déplacements)		Enquête ménages déplacements 1985 (déplacements)		Enquête ménages déplacements 1995 (déplacements)		Enquête ménages déplacements 1995 (chaînes)	
	domicile vers vers domicile	domicile vers vers domicile	domicile vers vers domicile	domicile vers vers domicile
travail	12,8 %	11,4 %	11,9 %	11,1 %	9,8 %	8,9 %	11,3 %	11,2 %
école	8,9 %	8,4 %	7,7 %	7,2 %	7,2 %	6,5 %	7,4 %	7,4 %
achats/ services	7,3 %	9,2 %	9,0 %	8,2 %	8,3 %	9,5 %	8,3 %	8,4 %
loisirs	5,0 %	5,9 %	7,0 %	7,8 %	7,1 %	8,1 %	6,3 %	6,2 %
autres	7,8 %	7,1 %	5,7 %	7,0 %	6,9 %	6,3 %	5,8 %	5,9 %
secondaires	16,3 %		17,3 %		21,3 %		21,3 %	

Tableau 6.1 : Répartition des déplacements par motifs
(Sources : enquêtes ménages déplacements Lyon 1976, 1985, 1995)

Le poids croissant des déplacements secondaires pose le problème de la définition des motifs de déplacements. Si nous reprenons l'exemple de la personne conduisant ses enfants à l'école pour aller ensuite au travail, une approche par motif de déplacement ne retiendra pas le motif domicile travail, puisque le premier motif hors domicile est l'accompagnement. A l'inverse, une approche par sortie permet de définir un motif principal pour la sortie (ce motif est alors affecté aux deux chaînes composant la sortie (schéma 6.2, p. 167)). Pour les données présentées dans le tableau 6.1, nous avons retenu la définition suivante, le motif de la sortie est travail, si ce motif est présent parmi les déplacements de la sortie, sinon le motif est école si le motif est présent ... en retenant la hiérarchie suivante travail, école, accompagnement, achats/service, loisir, autre qui s'appuie sur un degré de contrainte globalement décroissant pour les activités correspondantes. Cette définition conduit à une répartition par motif différente de la définition par motif de déplacement. Le poids du travail et dans une moindre mesure de l'école s'accroît au détriment des autres motifs. Cette définition présente deux avantages majeurs :

- la définition est plus conforme au motif principal de la sortie du domicile ;
- la définition permet une symétrisation de la matrice des déplacements par motif principal qui n'est pas sans intérêt comme nous le mentionnons ci-dessous pour l'estimation des émissions et attractions. Il apparaît en effet que la présence d'activités comme les achats/services ou les loisirs est beaucoup plus fréquente lors du retour du travail vers le domicile que sur le trajet pour aller travailler, à l'inverse des accompagnements (inclus dans le motif autres) qui sont plus fréquents sur le trajet aller.

Lorsque l'on retient une définition par motif de déplacement, on peut rencontrer deux méthodes de calcul des émissions et attractions par motif. Pour l'illustrer, nous prenons l'exemple du domicile-travail :

- dans la première, les émissions pour le motif domicile-travail sont estimées sur les déplacements réalisés pour le motif orienté domicile vers travail (le retour travail vers domicile n'est pas pris en compte) en considérant les zones émettrices, c'est-à-dire la zone u domicile. Les attractions sont estimées également sur les déplacements réalisés pour le motif orienté domicile vers travail, mais en considérant les zones réceptrices, c'est-à-dire la zone du travail ;
- dans la seconde, la définition des émissions et attractions que nous avons donnée section 1.1 est modifiée. L'esprit de la définition est conservé, mais sa traduction devient très différente. Elle s'appuie sur une décomposition en motifs en distinguant les motifs pour lesquels le domicile est à l'origine ou à la destination du déplacement (déplacements basés au domicile ou

home-based trips (HB) en anglais) des déplacements secondaires (*non home-based trips* (NHB)). Généralement seul les déplacements basés au domicile sont pris en compte :

- l'émission d'une zone (A) correspond alors à l'ensemble des déplacements basés au domicile des résidents de la zone A dont la destination est le travail quelle que soit la zone de destination (le schéma 6.3 permet de préciser la définition (Ortuzar, Willumsen, 1994)) ;
- l'attraction d'une zone A correspond à l'ensemble des déplacements basés au domicile (quelle que soit la zone du domicile) ayant l'autre extrémité dans la zone A, l'activité réalisée dans la zone A étant le travail (schéma 6.3).

Dans certains cas, les déplacements secondaires ne sont pas distingués. Dans ce cas, la définition des émissions pour le motif travail correspond à la définition précédente à laquelle on ajoute les déplacements secondaires dont l'origine est située dans la zone A lorsque le motif à l'origine est le travail (schéma 6.3). Pour les attractions, on ajoute à la définition précédente les déplacements attirés par la zone A dont le motif à destination est le travail (schéma 6.3).

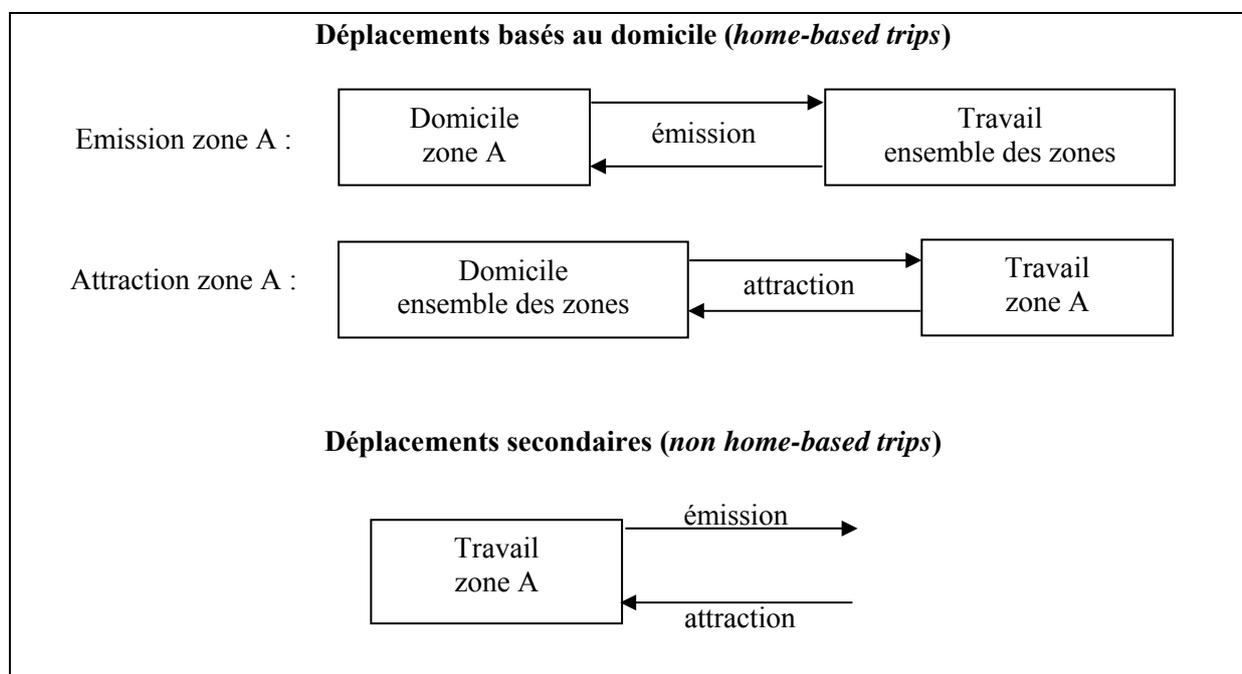


Schéma 6.3 : Emissions et attractions, définition pour le motif travail

Toutefois, ces définitions ne permettent pas de définir sans hypothèses supplémentaires le nombre de déplacements sortant ou entrant d'une zone pour les motifs non secondaires, pourtant nécessaire pour l'étape suivante de distribution. La solution du problème doit passer par une hypothèse de symétrisation. Le schéma suivant pour le motif domicile-travail permet d'illustrer le problème (schéma 6.4).

Il est clair qu'aucune des deux définitions ne permet d'accéder directement au nombre de déplacements émis par une zone, c'est-à-dire sortant de la zone (déplacements des flux 1 et 4), et au nombre de déplacements reçus, c'est-à-dire entrant dans la zone (flux 2 et 3) :

- dans la première méthode, seuls les flux 1 et 3 sont calculés. Pour estimer les flux émis, on additionne les flux 1 + 3 ce qui nécessite l'hypothèse d'égalité entre les flux 3 et 4. Pour estimer les flux reçus, on additionne les flux 3 et 1 ce qui nécessite l'hypothèse d'égalité entre les flux 1 et 2 ;

- dans la seconde méthode, on calcule d'une part les flux 1+2 (émissions) et les flux 3+4 (attractions). Pour estimer les flux émis, on additionne alors les flux (1+2) et (3+4) en divisant l'ensemble par deux. On calcule de même les flux reçus, ce qui nécessite de faire la même l'hypothèse d'égalité entre les flux 1 et 2 d'une part et 3 et 4 d'autre part.

L'hypothèse identique dans les deux cas, signifie que lorsque l'on somme sur l'ensemble des zones, le nombre de déplacements domicile vers travail est égal au nombre de déplacements travail vers domicile. La lecture du tableau 6.1 contredit immédiatement cette hypothèse pourtant indispensable. En revanche, en retenant une définition des motifs s'appuyant sur une appréhension des sorties du domicile, la symétrie est garantie (à quelques rares exceptions près des individus qui sont hors du domicile au début ou à la fin de la période de recueil (soit à 4 heures du matin dans le cas de l'enquête ménages déplacements)). Si l'on retient malgré tout une définition par motif de déplacement, la seconde méthode est en général à privilégier car elle permet d'estimer les régressions sur des effectifs enquêtés doubles de la première méthode. Pour résoudre le problème, on pourrait estimer quatre régressions relatives à chacun des flux, mais on se heurte alors à d'autres problèmes de cohérence en prévision.

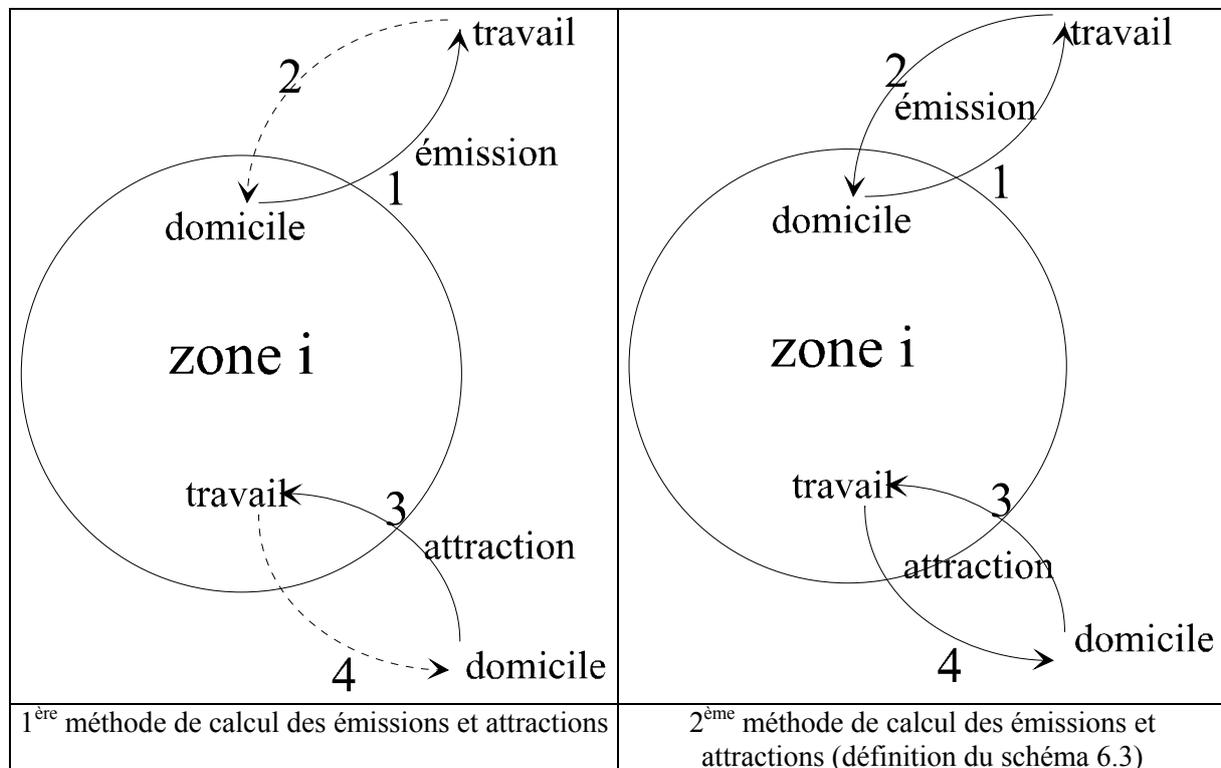


Schéma 6.4 : Représentation du calcul des émissions et attractions pour les déplacements basés au domicile selon la méthode de calcul

1.6. Principales variables explicatives de la génération des déplacements

Pour l'émission, les variables qui sont prises en compte sont le plus souvent relatives aux caractéristiques de la population résidant dans la zone considérée. Les variables que l'on retrouve le plus souvent dans les modèles de génération sont les suivantes :

- revenu du ménage (parfois de l'individu) ;
- motorisation du ménage généralement exprimé en nombre de voitures disponibles ou à disposition du ménage bien qu'il soit préférable de retenir une définition en nombre de

- voitures par individu du ménage ou par individu de plus de 18 ans du ménage (chapitre 3, p. 80) ;
- composition du ménage lorsque la génération s'effectue au niveau du ménage ;
 - nombre de personnes ou de ménages de la zone ;
 - nombre d'enfants en âge d'aller à l'école en distinguant souvent selon l'âge afin de retrouver les niveaux primaire, collège, lycée, université ;
 - l'âge des individus, ou plus généralement une répartition par classe d'âge.

Pour l'attraction, on utilise davantage les variables qui caractérisent les activités disponibles dans la zone (y compris la population comme source d'attraction par exemple pour les loisirs (visites à la famille, aux amis...)) :

- nombre d'emplois de la zone, souvent segmentés en emplois tertiaires et non tertiaires et parfois en isolant les emplois commerciaux parmi les emplois tertiaires ;
- nombre de places dans les écoles, collèges, lycées et universités ;
- nombre de ménages ou d'individus de la zone.

Lorsque qu'une décomposition par motif est utilisée isolant le motif achats/services, des données sur les emplois commerciaux ou les surfaces commerciales sont parfois utilisées.

Ces données ne sont pas exclusives d'autres sources de données lorsqu'elles sont disponibles sur une agglomération. Toutefois, la difficulté d'utiliser d'autres données réside dans la disponibilité de l'information en situation de référence, mais aussi et surtout de la possibilité de simuler l'information correspondante en situation de prévision.

2. PRINCIPALES FAMILLES DE MODELES DE GENERATION

Les premiers modèles utilisés ont souvent été des modèles normatifs faute de données disponibles permettant de caler un modèle de génération sur le terrain d'étude (section 2.1). La mise en place progressive du dispositif d'enquêtes a permis de caler les modèles sur les données locales par régression linéaire (section 2.2). Toutefois, afin de tenir compte de la forte variabilité des comportements des individus ou ménages, les modèles catégoriels ont été introduits. Ils consistent à définir des classes au comportement plus homogène et à développer ensuite des modèles par régression pour chacune des classes (section 2.3). Enfin, des modèles désagrégés de choix discrets ont également été développés. Les choix correspondent à la décision de réaliser 0, 1, 2... déplacements par jour ou par semaine le plus souvent en distinguant selon le motif du déplacement. Le résultat du modèle fournit alors la probabilité d'effectuer 0, 1, 2... déplacements. Nous ne développons toutefois pas ces modèles dont le principe général est décrit au chapitre 9.

2.1. Modèles normatifs

Les premiers modèles de génération développés aux Etats-Unis ou en France ont parfois utilisé des normes de génération des déplacements. Ces normes permettaient de pallier aux manques de données disponibles dans certaines agglomérations. C'était le cas des outils proposés par le SETRA dans le cadre des EPIT dans les années 70 (graphique 4.1, p. 140). Une autre illustration provient du logiciel QRS2 développé par Horowitz aux Etats-Unis, qui propose les normes de génération figurant dans les tableaux 6.2 et 6.3 pour des agglomérations d'une taille comprise entre 50 et 100 000 habitants en 1993. Ces normes sont issues de l'actualisation d'un rapport pour le compte de la Federal Highway Administration qui proposaient une série de normes en fonction de différents paramètres notamment de taille de l'agglomération (Sossau et al., 1978).

Si ces normes sont utiles en l'absence de données, il est évident que leur transférabilité d'une agglomération sur une autre et qui plus est entre pays est problématique. Lorsque des données sont disponibles, il est nécessaire de les utiliser pour déterminer la valeur de ces coefficients par régression linéaire.

Nombre de déplacements attirés par ...	Emploi commercial	Emploi non commercial	Ménage
Domicile-Travail	1.7	1.7	0
Domicile-Autre motif	10	0.5	1
Secondaire	2	2.5	0.5

Tableau 6.2 : Nombre de déplacements attirés selon les caractéristiques de la zone
(Source : Horowitz, 1993)

Revenu en milliers de \$	Nombre de voitures par ménage	Nombre de déplacements par ménage	% domicile-travail	% domicile-autres motifs	% de déplacements secondaires
6.	0.56	4.5	21.	57.	22.
8.	0.81	6.8	21.	57.	22.
10.	0.88	8.4	21.	57.	22.
12.	0.99	10.2	18.	59.	23.
14.	1.07	11.9	18.	59.	23.
16.	1.17	13.2	16.	61.	23.
18.	1.25	14.4	16.	61.	23.
20.	1.31	15.1	16.	61.	23.
25.	1.47	16.4	15.	62.	23.
30.	1.69	17.7	14.	62.	24.
40.	1.85	18.0	13.	62.	25.
50.	2.03	19.0	13.	62.	25.
50. et plus	2.07	19.2	13.	62.	25.

Tableau 6.3 : Nombre de déplacements émis selon le revenu ou la motorisation de la zone
(Source : Horowitz, 1993)

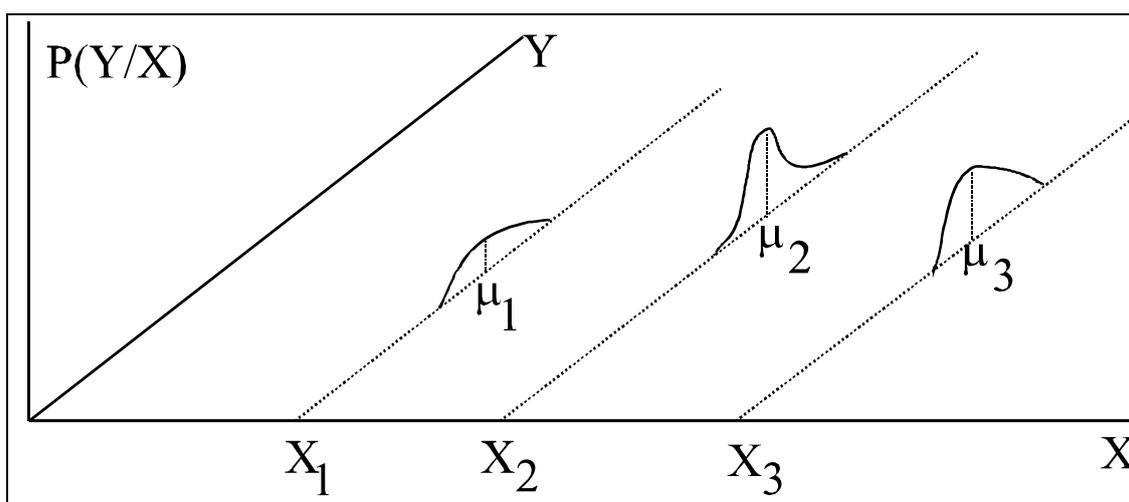
2.2. Modèles par régression linéaire

Le principe de construction du modèle est assez simple. Il nécessite la disposition d'un jeu de données donnant en situation de référence les émissions et attractions par zone et des caractéristiques des zones en termes de population et activités. Il faut alors déterminer les variables expliquant le mieux les émissions et attractions par zone. On utilise pour cela le modèle de régression linéaire multiple.

En général, en milieu urbain, on utilise les données de l'enquête ménages déplacements pour calculer les émissions et attractions par zone à la date de réalisation de la dernière enquête (ou des dernières enquêtes si l'on souhaite prendre en compte la dimension temporelle dans l'estimation des coefficients de régression). On calcule alors le nombre de déplacements émis et reçus par chacune des zones pour chacun des motifs. Les données caractérisant l'activité de chacune des zones doivent être recherchées auprès des organismes les produisant (INSEE, chambre de commerce, agence d'urbanisme, rectorat...). En revanche, les données de caractérisation de la population en situation de référence peuvent être estimées à partir des données de l'enquête ménages déplacements, plutôt qu'à partir des données du recensement de l'INSEE. Ce choix permet de s'assurer d'une bonne cohérence entre les variables à expliquer (émission et attraction) et les variables explicatives (données de population). De plus, on évite le délicat problème de l'actualisation des données, les dates de réalisation du recensement et de l'enquête ménages déplacements coïncidant assez rarement.

2.2.1. Hypothèses du modèle linéaire général

L'application de ce modèle nécessite le respect de certaines conditions que nous rappelons ici (le lecteur souhaitant approfondir le modèle linéaire général peut consulter entre autres Wonnacott, Wonnacott, 1991 auquel nous empruntons largement pour la rédaction de cette section). Considérons une variable Y , par exemple le nombre de déplacements émis pour le motif domicile travail, et une variable explicative X , par exemple le taux d'actifs dans la zone. Supposons que l'on dispose d'un grand nombre d'observations. Pour chaque valeur de X , nous avons plusieurs valeurs de Y (parce que d'autres variables que le taux d'actifs sont également explicatives, parce que l'on a des erreurs de mesure...). Il est possible de représenter (graphique 6.1) la distribution des valeurs de Y , soit un ensemble de distributions conditionnelles Y ($P(Y_1 / X_1), P(Y_2 / X_2), \dots$).

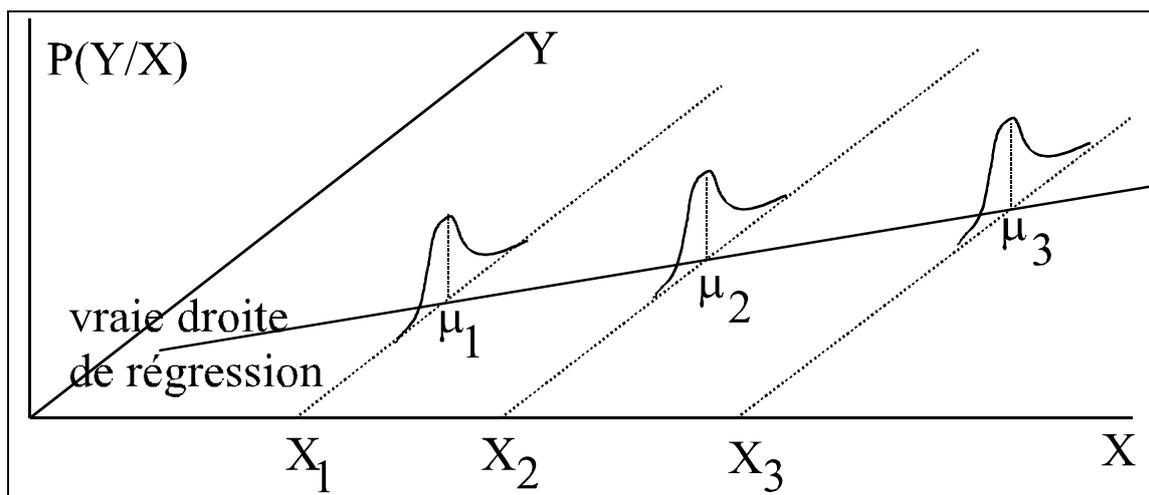


Graphique 6.1 : Distribution générale de Y selon X donné

Les distributions de probabilités conditionnelles de Y sont dans cet exemple différentes pour chaque valeur de X . Le problème ainsi posé n'est pas soluble et nécessite donc de poser des hypothèses que l'on peut représenter graphiquement (graphique 6.2).

Trois hypothèses sont à la base du modèle de régression :

- homogénéité de la variance : les distributions de Y ($P(Y_1 / X_1), P(Y_2 / X_2), \dots$) ont la même variance σ^2 ;
- linéarité : la moyenne de chaque distribution de Y ($\mu_i = E(Y_i / X_i)$) se situe sur une droite qui est la vraie droite de régression, soit $E(Y_i / X_i) = \mu_i = \alpha + \beta X_i$. Cette vraie droite de régression est celle inconnue relative à l'ensemble de la population d'étude ;
- indépendance : les variables aléatoires Y_1, Y_2, \dots sont statistiquement indépendantes. Cela signifie par exemple qu'une forte valeur pour Y_1 ne contribue pas à une forte valeur pour Y_2 . Cette condition peut encore s'exprimer $Y_i = \alpha + \beta X_i + e_i$ où e_1, e_2, \dots sont des aléas indépendants de moyenne nulle et de même variance σ^2 . Ces aléas correspondent d'une part aux erreurs de mesure inhérentes à tout processus de production de données et d'autre part aux erreurs de spécification liées à la non prise en compte de facteurs influant sur la variable étudiée.



Graphique 6.2 : Hypothèses sur la distribution générale de Y selon X dans le modèle linéaire

La vraie droite de régression est évidemment inconnue et le restera. En revanche, les mesures d'échantillon qui elles sont disponibles vont permettre d'estimer l'équation. Pour cela une hypothèse supplémentaire est nécessaire afin de pouvoir estimer $\hat{\beta}$: $\hat{\beta}$ suit une distribution approximativement normale autour de sa vraie valeur β . Cette hypothèse est en principe respectée à condition que la taille de l'échantillon soit suffisante (on considère en général que c'est le cas à partir de 30). L'utilisation de la méthode des Moindres Carrés Ordinaires (MCO) conduit à l'estimation de la droite : $\hat{Y} = \hat{\alpha} + \hat{\beta}X$

Où $\hat{\alpha}$ est un estimateur sans biais de α et $\hat{\beta}$ un estimateur sans biais de β .

Les hypothèses que nous avons faites combinées à l'utilisation de la méthode des MCO, nous permettent d'estimer $\hat{\alpha}$ et $\hat{\beta}$ et leur intervalle de confiance. Elles permettent également de faire des hypothèses sur la significativité des résultats obtenus.

2.2.2. Estimation de β

L'estimation des coefficients de la régression se formalise comme un problème de minimisation de la somme des carrés des distances entre chacune des observations et la droite de régression (distance mesurée en prenant la parallèle à l'ordonnée (c'est-à-dire la distance entre Y_i et \hat{Y}_i sur le graphique 6.3, p. 180). Il s'agit donc de minimiser :

$$\sum_i d^2 = \sum_i (Y_i - \hat{Y}_i)^2 \quad (6.1)$$

Y_i est la valeur observée lors de l'enquête ;

\hat{Y}_i est la valeur estimée de Y_i , c'est-à-dire la valeur sur la droite estimée de régression.

en remplaçant \hat{Y}_i par son estimation $\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i$, l'équation 6.1 devient :

$$\min_{\hat{\alpha}, \hat{\beta}} \left(\sum_i (Y_i - \hat{\alpha} - \hat{\beta}X_i)^2 \right) \quad (6.2)$$

que l'on peut résoudre en annulant la dérivée par rapport à $\hat{\alpha}$ et $\hat{\beta}$ de la quantité entre parenthèse dans l'équation 6.2. On obtient alors la solution suivante pour le coefficient $\hat{\beta}$ (la solution pour $\hat{\alpha}$ est donnée dans la section suivante) :

$$\hat{\beta} = \frac{\sum_i (X_i - \bar{X}) * (Y_i - \bar{Y})}{\sum_i ((X_i - \bar{X})^2)} \quad (6.3)$$

avec \bar{X} moyenne des X_i et \bar{Y} moyenne des Y_i .

Lorsque l'on force la constante à 0, la solution devient :

$$\hat{\beta} = \frac{\sum_i (X_i) * (Y_i)}{\sum_i (X_i)^2} \quad (6.4)$$

L'écart-type de $\hat{\beta}$ est fourni par :

$$SE(\hat{\beta}) = \frac{\sigma}{\sqrt{\sum_i x_i^2}} \quad \text{où } x_i = X_i - \bar{X} \quad (6.5)$$

La variance σ^2 de la population est inconnue. On l'estime alors par :

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_i (Y_i - \hat{Y}_i)^2 \quad (6.6)$$

où n est la taille de l'échantillon, $n-2$ représente le nombre de degrés de liberté qui est n (taille de l'échantillon) $- 2$ (nécessaire à l'estimation de $\hat{\alpha}$ et $\hat{\beta}$) ;

Il y a donc deux manières de réduire la variance de $\hat{\beta}$ et donc la qualité de l'ajustement. La première consiste à accroître la taille de l'échantillon. La seconde consiste à accroître la valeur de la somme des x_i^2 . Il faut pour cela accroître la dispersion des X_i donc choisir des valeurs les plus différentes possibles pour les X_i . Cette dimension est à prendre en compte lors de la définition du plan d'échantillonnage.

Connaissant l'espérance de $\hat{\beta}$ et l'écart-type de $\hat{\beta}$, il est possible de calculer l'intervalle de confiance sur β . L'intervalle de confiance à 95 % sur β est fourni par :

$$\beta = \hat{\beta} \pm t_{0,025} SE(\hat{\beta}) = \hat{\beta} \pm t_{0,025} \frac{\hat{\sigma}}{\sqrt{\sum_i x_i^2}} \quad (6.7)$$

où $t_{0,025}$ est la valeur du t de student à $n-2$ degrés de liberté au seuil de 95 % (on retient la valeur de $t_{0,025}$ et non $t_{0,05}$ lorsque le risque est bilatéral) ;

$$x_i = X_i - \bar{X}.$$

Ceci permet de tester l'hypothèse nulle $H_0 : \beta = 0$. Pour ce test, il faut calculer la valeur du t de student $= \frac{\hat{\beta}}{SE(\hat{\beta})}$ et comparer cette valeur à la valeur critique du t de student à $n-2$ degrés de liberté au seuil choisi (en général 95 %). Le plus souvent, on utilise plutôt la probabilité de se situer au-delà de la valeur de t qui est fournie par les logiciels et qui dans ce cas ne permet plus de rejeter l'hypothèse nulle.

2.2.3. Estimation de α

La résolution de l'équation 6.2 (p. 176) en annulant la dérivée par rapport à $\hat{\alpha}$ fournit l'équation de calcul de $\hat{\alpha}$:

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X} \quad (6.8)$$

L'écart-type de $\hat{\alpha}$ peut être estimé par :

$$SE(\hat{\alpha}) = \sqrt{\frac{\hat{\sigma}^2}{n} + \bar{X}^2 * (SE(\hat{\beta}))^2} \quad (6.9)$$

L'intervalle de confiance à 95 % sur α est fourni par :

$$\alpha = \hat{\alpha} \pm t_{0,025}SE(\hat{\alpha}) = \hat{\alpha} \pm t_{0,025} \sqrt{\frac{\hat{\sigma}^2}{n} + \bar{X}^2 * \frac{\hat{\sigma}^2}{\sum x_i^2}} \quad (6.10)$$

où $t_{0,025}$ est la valeur du t de student à n-2 degrés de liberté au seuil de 95 %.

De même, on peut tester l'hypothèse nulle $H_0 : \alpha = 0$. Il faut alors calculer la valeur du $t = \frac{\hat{\alpha}}{SE(\hat{\alpha})}$ et comparer cette valeur à la valeur critique du t à n-2 degrés de liberté au seuil choisi ou lire la valeur de la probabilité de se situer au-delà de la valeur de t fournie par les logiciels.

2.2.4. Test F sur le modèle complet

Le test de Fisher permet de tester l'hypothèse nulle H_0 : les coefficients β des variables explicatives sont tous nuls simultanément. Le test de F est bidimensionnel à 1 (le nombre de degrés de liberté de la régression, soit le nombre de variables explicatives) et n - 2 (le nombre de degrés de liberté des résidus, soit le nombre d'observations – le nombre de variables explicatives – 1) degrés de liberté. Comme pour le t de student, les logiciels fournissent généralement la probabilité pour le F. Toutefois, ce test n'est pas très performant, comparativement aux tests de nullité des coefficients β . Dans la pratique, l'hypothèse nulle est souvent rejetée.

Ce test est obtenu par la formule suivante :

$$F = \frac{\left(\sum_i (Y_i - \bar{Y}_i)^2 - \sum_i (Y_i - \hat{Y}_i)^2 \right) / (k)}{\left(\sum_i (Y_i - \hat{Y}_i)^2 \right) / (n - k - 1)} \quad (6.11)$$

où n est la taille de l'échantillon et k le nombre de variables explicatives = 1 dans la régression simple. k exprime le nombre de degrés de liberté de la régression et n - k - 1 le nombre de degrés de liberté des résidus.

Il s'interprète donc comme le rapport entre les carrés expliqués (somme des carrés totaux – somme des carrés résiduels) et les carrés non expliqués (somme des carrés résiduels), pondéré par les degrés de liberté de la régression et des résidus. Le F augmente donc avec le pouvoir explicatif de la régression tout en étant pondéré par les degrés de liberté.

2.2.5. Coefficient de détermination R^2

En l'absence de variable explicative (X), la meilleure estimation pour Y serait la moyenne des Y_i , soit \bar{Y} . Pour chaque valeur Y_i , l'écart avec la valeur estimée de Y est alors $Y_i - \bar{Y}$.

Connaissant, la valeur des X_i , il est possible d'avoir une autre estimation pour Y qui est la valeur fournie par la droite de régression :

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i \quad (6.12)$$

L'écart du modèle sans les X_i peut donc se décomposer de la manière suivante :

$$Y_i - \bar{Y} = \hat{Y}_i - \bar{Y} + Y_i - \hat{Y}_i \quad (6.13)$$

écart total = écart expliqué par la régression + écart non expliqué ou résiduel.

Prenons un exemple chiffré pour illustrer graphiquement cette décomposition (graphique 6.3). Soit le jeu de données présenté dans le tableau 6.4.

Observation N°	Variable à expliquer Y	Variable explicative X
1	1,50	0,50
2	2,50	1,10
3	5,00	2,10
4	7,30	2,90
5	5,00	3,30
6	8,00	3,95

Tableau 6.4 : Jeu de données

Si l'on ne dispose d'aucune information complémentaire (pas de variable explicative notamment), la meilleure estimation pour Y sera la moyenne des Y_i , soit $\bar{Y} = 4,88$. Mais on dispose en fait d'une information complémentaire avec une variable X dont on pense qu'elle peut expliquer la valeur prise par Y . Sous réserve du respect des hypothèses nécessaires à la régression, on peut estimer la droite de régression. Son équation est fournie sur le graphique 6.3 : $\hat{Y}_i = 2,04 X_i$, si l'on prend comme référence un modèle sans constante ($\alpha = 0$). On dispose donc d'une estimation pour Y fournie par cette équation. Si l'on choisit l'observation n°4, on obtient $\hat{Y}_4 = 5,92$ (graphique 6.3). Si l'on reprend l'équation 6.13 pour cette observation, on obtient :

$$\begin{aligned} Y_i - \bar{Y} &= \hat{Y}_i - \bar{Y} + Y_i - \hat{Y}_i \\ 7,30 - 4,88 &= 5,92 - 4,88 + 7,30 - 5,92 \\ \text{écart total} &= \text{écart expliqué par la régression} + \text{écart non expliqué ou résiduel.} \end{aligned}$$

Ecart représenté sur le graphique 6.3.

Si l'on somme le carré des écarts sur n , on obtient la variance totale. On peut facilement montrer que cette variance peut être décomposée de la manière suivante (propriété de l'analyse de variance) :

$$\sum \frac{(Y_i - \bar{Y})^2}{n} = \sum \frac{(\hat{Y}_i - \bar{Y})^2}{n} + \sum \frac{(Y_i - \hat{Y}_i)^2}{n} \quad (6.14)$$

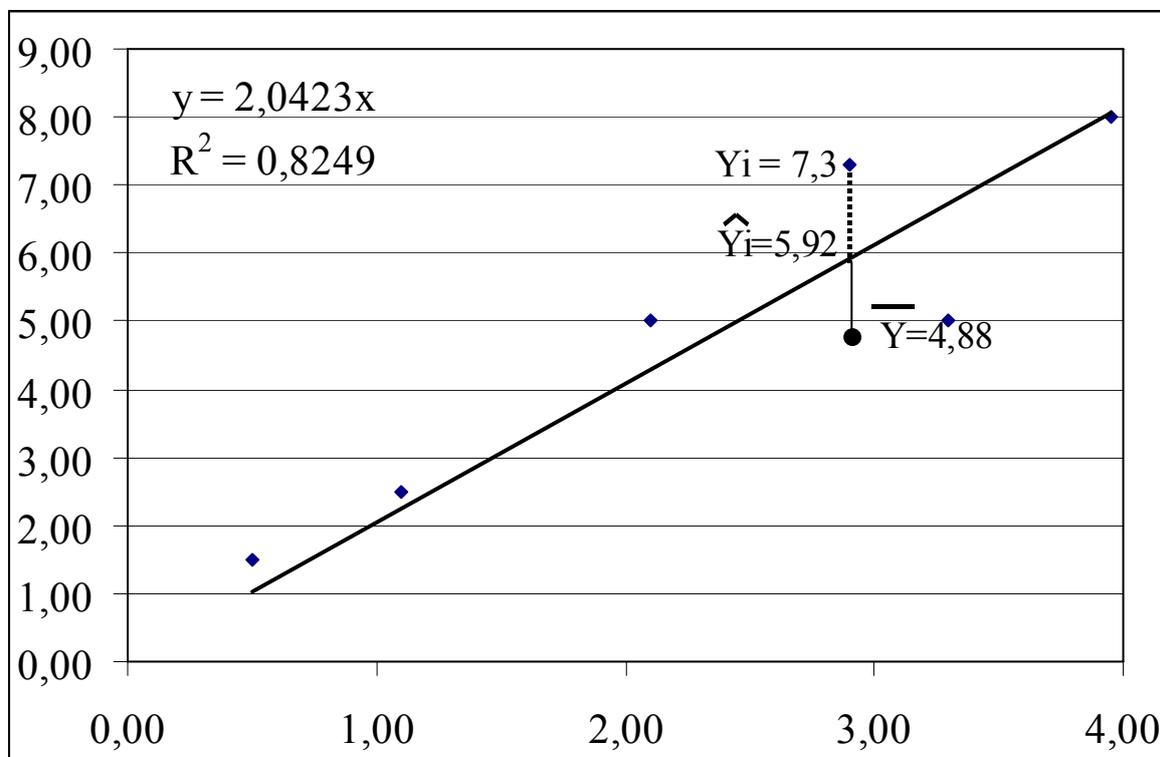
variance totale = variance expliquée par la régression + variance non expliquée ou résiduelle.

Il est alors possible de calculer le pourcentage de la variance expliquée par le modèle explicatif construit par régression. C'est ce que l'on appelle le coefficient de détermination R^2 fournit par :

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \quad (6.15)$$

Si l'on reprend l'illustration du graphique 6.3, on peut donc interpréter le R^2 comme étant le rapport entre la moyenne des carrés des écarts à la droite de régression sur la moyenne des carrés des écarts à la valeur moyenne des Y_i . Il correspond donc à la partie des écarts initiaux qui est expliqué par la droite.

Sa valeur est comprise entre 1 (la totalité de la variation initiale est expliquée par le modèle, c'est-à-dire l'ensemble des points sont situés sur la droite) et 0 (le modèle n'apporte aucune explication par exemple si tous les X ont la même valeur). On notera que le coefficient de détermination correspond au carré du coefficient de corrélation entre X et Y qui exprime le degré de relation entre X et Y .



Graphique 6.3 : Illustration de la décomposition des écarts et de l'interprétation graphique du coefficient R^2

2.2.6. Prévision de Y pour une valeur donnée de X

Comme pour les coefficients α et β , il est possible de calculer un intervalle de confiance pour la vraie valeur de Y relative à une valeur donnée de X . Cette vraie valeur inconnue a été notée (section 2.2.1, p. 175) $E(Y_0 / X_0) = \mu_0 = \alpha + \beta X_0$.

La meilleure estimation que l'on ait de μ_0 est :

$$\hat{Y}_0 = \hat{\alpha} + \hat{\beta} X_0 \quad (6.16)$$

à partir de laquelle, on peut déterminer un intervalle de confiance à 95 % sur μ_0 (cet intervalle est obtenue en prenant en compte la variance sur $\hat{\alpha}$ et $\hat{\beta}$) :

$$\mu_0 = (\hat{\alpha} + \hat{\beta}X_0) \pm t_{0,025} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x_i^2}} \quad (6.17)$$

Pour estimer l'intervalle de confiance pour une valeur donnée Y, il faut ajouter à la variance d'estimation de la vraie droite de régression, la variance de Y. Pour une observation donnée, nous obtenons donc l'intervalle suivant :

$$Y_0 = (\hat{\alpha} + \hat{\beta}X_0) \pm t_{0,025} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum x_i^2} + 1} \quad (6.18)$$

Il est important de noter que plus on s'éloigne de la valeur moyenne des X, plus l'intervalle de confiance est important, donc plus le risque d'erreur est important.

2.2.7. Régression multiple

La présentation peut être aisément généralisée à plusieurs variables explicatives. Les hypothèses sont les mêmes et l'on peut généraliser l'ensemble des résultats présentés ci-dessus au cas de plusieurs variables explicatives (les équations doivent toutefois être modifiées pour tenir compte de l'écriture matricielle qui conduit donc à des écritures un peu plus lourdes).

Lorsque l'on dispose de plusieurs variables explicatives, la tentation est alors grande d'augmenter le nombre de variables prises en compte dans la régression. Mais dans ce cas, le risque est de rencontrer des problèmes de multicollinéarité. Une des variables explicatives peut s'écrire comme une fonction linéaire des autres variables explicatives. Dans ce cas, le problème ne connaît pas une solution unique et ne peut être résolu.

On est rarement en présence d'un problème insoluble. Par contre, lorsque deux ou plusieurs variables sont corrélées, la qualité de la régression diminue et l'on observe pour ces variables un élargissement de l'intervalle de confiance du coefficient $\hat{\beta}_i$ correspondant, qui peut conduire à l'absence de rejet de l'hypothèse nulle pour l'une des variables (la probabilité sur le test de student relatif au coefficient de la variable explicative est supérieur au seuil fixé). Dans ce cas, il est nécessaire d'abandonner la variable correspondante à moins de transformer la variable afin de réduire ou mieux encore supprimer la corrélation entre les variables. Dans tous les cas, il est nécessaire d'analyser le modèle et de s'assurer de la pertinence de la variable choisie pour expliquer le phénomène étudié.

Dans le processus de choix des variables, on pourra utiliser la fonctionnalité de régression pas à pas, offerte par la plupart des logiciels statistiques, qui permet de hiérarchiser les variables explicatives en fonction de leur contribution à l'explication de la variance totale. Il ne s'agit toutefois que d'une aide qui ne doit pas occulter la nécessaire réflexion sur la pertinence économique ou comportementale de la variable choisie.

Lorsque l'on augmente le nombre de variables explicatives, on augmente toujours le coefficient de détermination R^2 . Pour éliminer ce problème, il est possible de calculer un coefficient R^2 ajusté en utilisant la formulation suivante :

$$\bar{R}^2 = \left(R^2 - \frac{k}{n-1} \right) * \left(\frac{n-1}{n-k-1} \right) \quad (6.19)$$

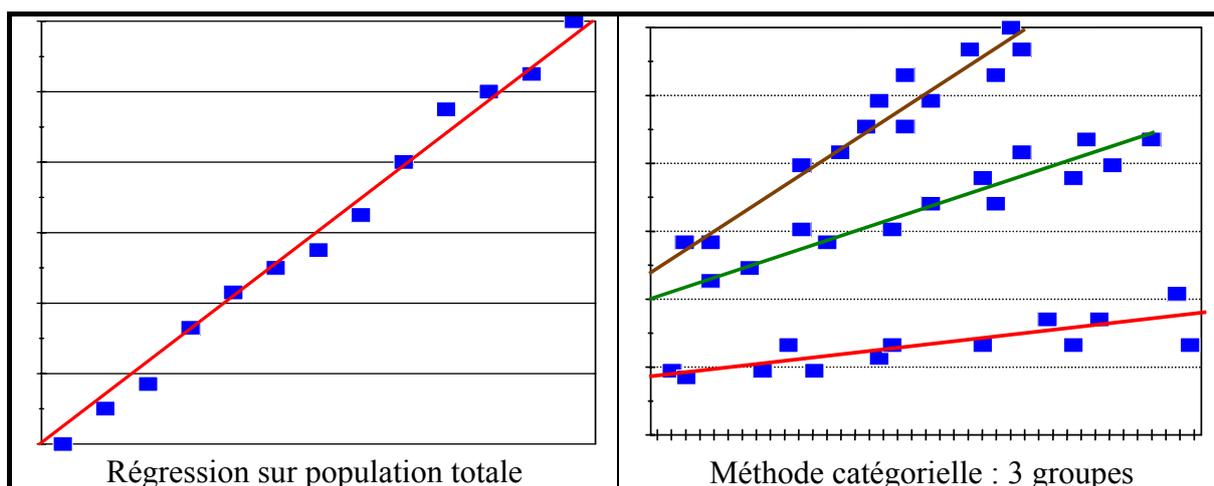
où n est la taille de l'échantillon et k le nombre de variables explicatives.

Lorsque le modèle ne comprend pas de constante, la formule devient :

$$\bar{R}^2 = \left(R^2 - \frac{k}{n-1} \right) * \left(\frac{n-1}{n-k} \right) \quad (6.20)$$

2.3. Modèle catégoriel

Le principe du modèle catégoriel consiste à définir des classes d'individus ou de ménages et à déterminer un modèle explicatif spécifique par régression pour chacune des classes. Ce modèle trouve sa justification dans la variabilité des comportements des individus ou des ménages d'une part et dans la possibilité de réduire cette variabilité par une classification adéquate des individus ou des ménages. Dans ce cas, si l'on dispose de suffisamment de données, il est possible d'estimer avec davantage de précision une droite de régression pour chacune des classes. Un exemple graphique permet d'illustrer le propos (graphique 6.4). Dans le dessin de gauche, la régression sur l'ensemble de la population donne de bons résultats avec des points pas trop éloignés de la droite de régression. En revanche, dans le cas de droite, une régression sur l'ensemble de la population donnerait des résultats de piètres qualités. La diversité des comportements est trop grande. En revanche, une segmentation de la population en trois catégories comme illustrée sur le dessin de droite conduit à des régression de bonnes qualités pour chacune des trois classes. Toutefois, il est bien évident que les données fournissent rarement un diagnostic aussi clair que dans cet exemple.



Graphique 6.4 : Illustration du choix entre un modèle de régression unique ou un modèle catégoriel

La première étape dans la construction d'un modèle catégoriel porte sur la détermination des classes. L'objectif de la segmentation est de réduire la variabilité des comportements par la construction de classes aux comportements de déplacements plus homogènes. La variable à expliquer est une variable quantitative, généralement un nombre de déplacements. Les variables de segmentation permettant de discriminer les comportements sont généralement des variables qualitatives (sexe, âge (sous forme de classes), position dans le cycle de vie, profession, présence d'enfants, motorisation...). L'analyse de la variance permet de déterminer les variables expliquant la variabilité des comportements. Elle permet également de déterminer les classes pertinentes pour chacune de ces variables. Les méthodes de classification permettent enfin de déterminer la segmentation la plus pertinente en combinant les différentes variables explicatives.

A priori, le modèle catégoriel semble plus pertinent car il permet de travailler sur des classes plus homogènes. Il est toutefois soumis à plusieurs contraintes portant sur les données :

- la taille de l'échantillon doit être suffisante pour pouvoir estimer la droite de régression pour chacune des classes. Deux conséquences en découlent. Le nombre de classes doit être le plus faible possible. La taille de chacune des classes doit être la plus proche possible ;
- au-delà du calage, il est nécessaire d'utiliser le modèle en prévision. Le choix des variables doit donc être guidé par la possibilité de prévoir l'évolution de chacune des classes à l'horizon d'étude.

C'est pour cette raison que certains auteurs préconisent plutôt l'utilisation de variables muettes. Supposons que l'on observe que le nombre de déplacements est nettement différent selon que le ménage dispose de 0, 1 ou 2 et plus voitures. Il est possible de segmenter la population en trois classes et d'estimer 3 régressions. Mais il est également possible d'ajouter deux variables muettes dans la régression :

- Z_1 prendra la valeur 1 lorsque le ménage dispose d'une voiture et 0 autrement ;
- Z_2 prendra la valeur 1 lorsque le ménage dispose de deux voitures et plus et 0 autrement.

Il est à noter que le nombre de variables muettes est égal au nombre de classes – 1.

Le choix entre les deux techniques n'est pas toujours facile. Ortuzar et Willumsen (1994) estiment préférable d'utiliser les variables muettes sauf lorsque l'on dispose d'un effectif suffisant dans chacune des classes de la segmentation. L'utilisation des variables muettes permet d'utiliser l'ensemble des données dans l'estimation de la régression alors que la classification nécessite le partage de l'échantillon en classes réduisant la qualité de l'estimation. Si nous rejoignons les auteurs lorsque la taille de l'échantillon est faible, il nous semble que d'autres facteurs doivent également guider le choix. Lorsque l'effectif le permettra, on privilégiera notamment le modèle catégoriel dans les cas suivants :

- si les variables explicatives sont différentes pour chacune des classes ;
- si les coefficients des variables explicatives sont significativement différents (comme dans l'exemple illustré par le graphique 6.4).

Le dernier avantage et non des moindres de l'approche catégorielle est de permettre de combiner le modèle par régression avec des approches démographiques par suivi de cohortes (Madre, Lambert, 1989 ; Galez, 1995 ; Pochet, 1995 ; Madre, Armoogum, 1996).

3. PRINCIPALES LIMITES DE L'ETAPE DE GENERATION

Nous débutons notre analyse des limites en commençant par l'examen des hypothèses formulées pour l'application du modèle linéaire (section 3.1). Nous passons ensuite en revue les trois conditions de l'opérationnalité des modèles :

- cohérence interne à l'étape de génération avec la non égalité des émissions et attractions (section 3.2) et cohérence interne au modèle à quatre étapes avec la non prise en compte de l'offre (bien qu'il s'agisse également d'un problème de pertinence (section 3.3)). Ce dernier point nous renvoie également vers l'interprétation de la fonction de demande (chapitre 5). On pourrait également inclure ici l'examen du respect des hypothèses du modèle linéaire qui constitue un problème de cohérence entre le choix du modèle explicatif et sa formalisation mathématique ;
- pertinence de l'hypothèse de stabilité dans le temps des paramètres de génération (section 3.4) et du choix des variables explicatives contraintes par la nécessité de devoir prévoir l'évolution de ces variables dans le futur (même si c'est aussi très largement un problème de mesurabilité (section 3.5)) ;
- mesurabilité avec l'analyse des liens entre production des données et génération (section 3.6).

3.1. Hypothèses du modèle linéaire

Nous examinons les conséquences des hypothèses du modèle linéaire : linéarité (section 3.1.1) et homoscedasticité (section 3.1.2). Nous ajoutons ici un problème fréquent dans les régressions multiples lorsque les variables explicatives sont corrélées entre elles : la multicollinéarité (section 3.1.3)

3.1.1. Linéarité

Le modèle de régression postule que chacune des variables exerce une influence linéaire sur le phénomène étudié. La connaissance que l'on a de l'effet de chacune des variables explicatives choisies dans le modèle n'est généralement pas suffisante pour s'assurer de cette linéarité. L'hypothèse mérite donc d'être analysée à l'aide des données empiriques disponibles. Cette analyse n'est toutefois pas toujours facile en raison de la relation entre les variables explicatives qui peut laisser croire à une relation linéaire alors que l'on est en présence de non linéarité sur une des variables explicatives. L'analyse de la distribution des résidus permet de détecter certains de ces problèmes (Wonnacott, Wonnacott, 1991).

Pour certaines variables, le constat est plus simple. C'est le cas de la plupart des variables socio-économiques dont l'influence est le plus souvent non linéaire (sexe, âge, profession, équipement automobile...). Face à des variables non linéaires, deux solutions sont disponibles :

- transformer la variable pour la rendre linéaire. Les transformations les plus fréquentes concernent le passage au logarithme ou à des fonctions puissances. Dans certains cas, on utilise des fonctions statistiques (cas du modèle prix-temps avec la transformée de Laplace-Gauss, chapitre 2, p. 54). L'analyse des résidus permet de renseigner sur le choix du type de transformation nécessaire ;
- dans le cas de variables qualitatives discrètes deux solutions se présentent : segmentation de la population ou utilisation de variables muettes. L'exemple de la variable taux d'équipement automobile permet d'explicitier la méthode. Le nombre de déplacements ne varie généralement pas linéairement en fonction du nombre de voitures du ménage. En revanche, la linéarité peut parfois être retrouvée pour chacune des classes de motorisation du ménage (généralement, 0, 1 et 2 et plus voiture(s)). On pourra alors soit utiliser des variables muettes dans la régression ce qui conduit à ajouter un nombre de variables explicatives dans la régression égal au nombre de classes - 1 (section 2.3, p. 182), soit estimer une régression sur chacune des classes d'équipement automobile.

3.1.2. Homoscedasticité

L'hypothèse d'homoscedasticité correspond à l'hypothèse d'égalité de la variance sur les distributions conditionnelles de Y/X . Lorsque cette hypothèse n'est pas vérifiée, on parle d'hétéroscedasticité.

Prenons un exemple simple pour illustrer le problème. L'utilisation du revenu parmi les variables explicatives conduit fréquemment à des problèmes d'hétéroscedasticité. Si l'on cherche à modéliser la consommation des ménages à partir du revenu, l'aléa sera probablement beaucoup plus réduit pour de faibles valeurs de revenu pour lesquels la quasi totalité du revenu sera consommée que pour de fortes valeurs de revenu ou la variabilité de la consommation sera plus forte. L'introduction d'autres variables explicatives, comme l'épargne par exemple, pourra permettre de réduire cette variabilité.

Un autre exemple classique concerne les variables explicatives zonales. Si la taille des zones est très variable et que l'on retient comme variable explicative le nombre de voitures de la zone, il est très

probable que l'écart-type variera avec la taille de la zone. Pour faire face à ce problème deux solutions sont disponibles :

- choisir un zonage comportant des zones de tailles similaires. Cet objectif est rarement compatible avec les autres contraintes provenant notamment des découpages d'enquêtes ou des limites administratives ;
- utiliser des variables exprimant non pas un nombre total pour la zone mais un ratio. Par exemple au lieu du nombre total de voitures de la zone, on privilégiera le nombre de voitures par personne ou par ménage, un taux d'actif au lieu du nombre d'actifs....

3.1.3. Multicolinéarité

La multicolinéarité désigne la corrélation entre les variables explicatives. Dans ce cas, il est difficile de dissocier les effets des deux variables. De ce fait, l'estimation des coefficients de la régression relatifs aux variables corrélées sera peu fiable, car leur écart-type sera élevé (Wonnacott, Wonnacott, 1991).

Ce problème est assez fréquent lorsque l'on utilise des variables socio-économiques de l'individu ou du ménage qui sont fréquemment corrélées entre elles, comme le revenu et la motorisation par exemple. Lorsque la corrélation est trop forte, il est généralement nécessaire de ne retenir qu'une seule des deux variables. Mais dans certains cas, il est possible de décorréler les variables par des transformations adaptées (par exemple lorsque l'on utilise des séries chronologiques faisant intervenir des prix ou des revenus, il est généralement nécessaire de transformer les variables pour les déflater de l'inflation ou d'un indice similaire).

3.2. Divergence entre les émissions et les attractions

Lors du calage, on observe par définition une identité entre la somme des déplacements émis par chacune des zones et la somme des déplacements reçus par chacune des zones. Lorsque l'on rentre dans une zone ..., on finit bien par en sortir. Si cette égalité est effective en situation de référence, ce n'est plus le cas en situation de prévision. L'égalité est pourtant nécessaire pour respecter le bon sens, mais aussi pour l'étape suivante de distribution comme nous l'illustrerons dans le chapitre suivant.

Le calcul des émissions est généralement considéré comme plus fiable, car s'appuyant sur des données socio-économiques des ménages dont l'évolution est mieux maîtrisée que celles des données d'activités zonales. De ce fait, le résultat de l'attraction est pondéré par un facteur identique pour chacune des zones égal au rapport des émissions et attractions :

$$\sum_i E_i \neq \sum_j A_j \quad \Rightarrow \quad A_{j\text{pondéré}} = A_j * \frac{\sum_i E_i}{\sum_j A_j} \quad \text{et donc} \quad \sum_i E_i = \sum_j A_{j\text{pondéré}}$$

3.3. Indépendance vis-à-vis de l'offre de transport et des autres étapes

La formulation de la génération tient compte uniquement des caractéristiques de la population et des activités des zones. De ce fait, l'offre de transport n'influe pas sur le niveau de mobilité. Cette hypothèse est en contradiction avec les phénomènes d'induction observés à la suite de la mise en service de nouvelles infrastructures de transport. Même si le concept d'induction est assez difficile à cerner (Offner, 1985), il y a certains individus, notamment parmi les inactifs non motorisés pour lesquels le nombre de déplacements évolue tout particulièrement pour des motifs non contraints (Bonnel, Thibaud, 1994 ; Bonnel, 1995a)

Plus généralement, il n'y a aucune interaction entre les résultats des phases ultérieures des modèles à quatre étapes et la phase de génération. Même parmi les modèles qui tentent de traiter de manière appropriée les questions d'équilibres (interactions entre les étapes de distribution, répartition modale et affectation), la phase de génération est considérée comme indépendante des suivantes. L'évolution du système de transport que ce soit à travers des améliorations ou au contraire des détériorations suite à des phénomènes de congestion croissante n'exerce donc aucune influence sur les émissions et attractions.

Des tentatives ont été menées pour réduire ce problème à travers la prise en compte de l'accessibilité, mais avec des résultats décevants (Ortuzar, Willumsen, 1994). L'accessibilité au départ d'une zone est une combinaison de facteurs attractifs qui poussent à se déplacer et de facteurs résistants qui réduisent l'intensité des déplacements lointains au profit des déplacements les plus courts. Le facteur attractif correspond à l'ensemble des opportunités présentes dans les autres zones (le plus souvent nombre de personnes, nombre d'emplois, équipement commercial, de loisirs ...) qui permettent la réalisation d'activités hors du domicile. Le facteur résistant est le plus souvent le coût généralisé nécessaire pour se rendre vers ces opportunités. La forme générique de l'accessibilité est la suivante :

$$A_i = \sum_{j \neq i} f(O_j, Cg_{ij}) \quad (6.21)$$

où A_i est l'accessibilité au départ de la zone i ;

O_j représente les opportunités de la zone j ;

Cg_{ij} représente le coût généralisé pour aller de la zone i à la zone j .

La fonction utilisée prend en général la forme d'une fonction gravitaire (chapitre 7, p. 199) :

$$A_i = \sum_{j \neq i} O_j e^{-\alpha Cg_{ij}} \quad (6.22)$$

où α est un coefficient à estimer.

L'absence de prise en compte de l'offre de transport est également en infraction avec les fondements théoriques que nous avons exposés dans le chapitre précédent. Reprenons la fonction de demande dont la forme générique est illustrée au graphique 5.6 (p. 150). Pour un état donné du sous-système de localisation des activités, la fonction de demande dépend du niveau de service. Or ce niveau de service est la traduction de l'état du sous-système de transport.

Dans le modèle de génération, les émissions et attractions traduisent le niveau de demande. Elles sont bien calculées à partir des variables traduisant l'état du sous-système de localisation des activités à un instant donné (nombre de ménages, de personnes, d'emplois...). En revanche, elles ne dépendent pas de l'offre de transport et ne sont donc pas aptes à représenter la fonction de demande telle qu'illustrée au graphique 5.6 (p. 150). C'est pour pallier à cette lacune que l'induction de trafic (qui traduit la relation entre le volume et le niveau de service toutes choses égales par ailleurs) est parfois prise en compte lors de l'introduction d'une nouvelle infrastructure lourde de transport (métro, tramway, rocade...). Le calcul de cette induction se fait toutefois en dehors du modèle de génération. Assez souvent, on utilise une élasticité de la demande à la variation du coût généralisé sur les origines-destinations concernées par l'investissement. Différentes formulations du coût généralisé sont possibles pour tenir compte de l'évolution de l'ensemble des modes et non du seul mode pour lequel il y a investissement. Une formulation du type du logsum est alors assez courante :

$$\text{Log} \left(\sum_k e^{Cg_{ij}^k} \right) \text{ ou } k \text{ représente les modes disponibles.} \quad (6.23)$$

Cette formulation générale est obtenue à partir d'une analogie avec la formulation de l'utilité composite dans le cas du modèle logit emboîté (équation 9.39, p. 261).

3.4. Stabilité des paramètres de génération

Le calage du modèle sur des données en coupe instantanée, telles que les données de l'enquête ménages déplacements, ne permet pas de faire des hypothèses sur une éventuelle évolution des coefficients de calage dans le temps. L'utilisation du modèle en prévision nécessitera donc de faire des hypothèses sur ces coefficients. En l'absence de données complémentaires, il est nécessaire de postuler la stabilité des coefficients de génération dans le temps.

Prenons un exemple pour illustrer le problème. Supposons que l'équipement automobile du ménage est une variable explicative importante du modèle de génération. On sait par ailleurs que la motorisation augmente dans le temps sous l'effet entre autre de la croissance des revenus. La stabilité des coefficients de génération liés à l'équipement automobile signifie donc que le ménage qui aura 2 voitures dans 10 ans fera autant de déplacements que celui qui en a déjà deux aujourd'hui (en faisant l'hypothèse que les autres variables explicatives sont identiques). L'hypothèse est hardie. On peut imaginer que les ménages qui avaient fortement besoin de voitures en raison d'une forte mobilité se sont équipés les premiers et donc que ceux qui s'équiperont dans le futur ne feront pas autant de déplacements. Inversement, on pourrait objecter que la généralisation/banalisation de la voiture dans une civilisation de plus en plus liée à l'automobile conduira à un usage de plus en plus important de l'automobile qui s'accompagnera d'une croissance de la mobilité liée à la facilité d'usage de ce mode de transport. Si l'on ne dispose que d'une enquête ménages déplacements sans données complémentaires, on en sera réduit à diverses conjectures sans possibilité de les confronter aux données existantes. L'hypothèse de stabilité est alors souvent la moins insatisfaisante. Elle peut toutefois, sur le long terme, conduire à des erreurs importantes de sur- ou sous- estimation de la demande conduisant à des investissements sous- ou sur-dimensionnés.

Pour résoudre le problème, il est nécessaire de disposer de données en série chronologiques ou tout au moins de mesures répétées dans le temps. La plupart des grandes agglomérations françaises offre cette opportunité avec les enquêtes ménages déplacements, même si l'intervalle de temps entre les enquêtes est important et le nombre de mesures limité.

3.5. Choix et prévision des variables explicatives

L'utilisation du modèle de génération nécessite de disposer de données sur la valeur des données d'entrée à l'horizon d'étude. Il faudra donc le plus souvent connaître le nombre de personnes, le nombre d'emplois, la motorisation... de chacune des zones. Ces données sont le plus souvent issues d'organismes locaux de planification, de l'INSEE ou de modèles de simulation. Dans tous les cas, il faut garder à l'esprit que ces données sont elles-mêmes issues d'un exercice de prévision ou de simulation. Elles sont donc sujettes aux limites inhérentes à tout exercice de prévision. Les rares études de comparaison des résultats de modélisation avec les trafics observés désignent d'ailleurs les erreurs sur les données d'entrée comme source principale d'erreurs sur les données en sortie (Mackinder, Evans, 1981 ; CEMT, 1982 ; SEMALY, 1996). Ce constat important suggère deux conséquences.

La première concerne le nécessaire exercice de sensibilité du modèle. Il faudra faire varier les paramètres d'entrée (proportionnellement au degré d'incertitude de la prévision de la variable considérée) et analyser les conséquences sur les résultats en sortie. Il faut donc bien se garder de la redoutable précision de certains résultats issus de modélisation laissant l'illusion de la fiabilité de l'exercice. Il faut bien être conscient que quel que soit le phénomène étudié les résultats sont toujours entachés d'incertitude qui s'accroît avec l'horizon de la prévision.

La seconde concerne le choix des variables explicatives. L'analyste devra toujours privilégier, dans la mesure du possible, des variables dont la prévision pourra être menée avec une précision suffisante. Prenons un exemple pour expliciter le problème. Les données d'emplois sont le plus souvent utilisées dans les modèles d'attraction. Elles servent à expliquer le nombre de déplacements domicile-travail, mais aussi les déplacements pour les motifs achats/services. Dans ce cas, il sera souhaitable d'avoir une distinction entre emplois secondaires et tertiaires. Ces derniers sont davantage susceptibles d'attirer des déplacements pour les motifs achats/services que les premiers. Pour affiner l'analyse, on pourra également vouloir isoler les emplois commerciaux au sein des emplois tertiaires. Dans certaines agglomérations ces données peuvent être disponibles en situation de référence. En revanche, la prévision de ces données à un niveau zonal sera d'autant moins fiable que la décomposition sera fine. On retrouve la problématique des modèles avec un choix entre une plus grande pertinence dans la décomposition des emplois, mais une moindre mesurabilité à l'horizon de la prévision.

3.6. Qualité des données

La qualité des données et leur adéquation avec les besoins de modélisation sont évidemment une question cruciale que l'on retrouvera pour chacune des étapes de la modélisation. Les données sont nécessaires à plusieurs niveaux :

- pour choisir les variables explicatives et formaliser le modèle de génération. En général, on utilise pour cela des données d'enquête mobilité comme les enquêtes ménages déplacements. Les exigences principales sur les données sont la représentativité statistique, la taille de l'échantillon (notamment lorsque que l'on souhaite segmenter la demande par motif et/ou par segment de population) et la présence des variables pertinentes. Le choix d'une interrogation à partir des déplacements ou des activités réalisés peut être également important selon le type de modèle utilisé. Il faut également disposer de données de caractérisation des activités par zone. Dans ce recueil, on doit généralement faire face à trois problèmes majeurs :
 - la disponibilité de données récentes ou tout au moins pas trop éloignées de la date de réalisation de l'enquête mobilité ;
 - la compatibilité des découpages zonaux utilisés pour les différentes sources de données ;
 - la fiabilité de l'information disponible ;
- pour caler et valider le modèle de génération. On utilise pour cela les mêmes données que ci-dessus. Aux problèmes précédents, il faut ajouter les conditions nécessaires à l'application du modèle linéaire et la multicollinéarité (dépendance entre les variables explicatives). La validation nécessite l'analyse des résultats produits par le modèle en prévision. La comparaison avec d'autres sources de données et l'expertise disponible sur le phénomène étudié permettent d'analyser la plausibilité des résultats à travers notamment le calcul de l'élasticité des résultats à certaines variables ;
- pour utiliser le modèle en prévision. Les deux problèmes principaux concernent d'une part la prévision des données d'entrée du modèle dont la fiabilité est parfois sujette à caution (notamment pour la prévision des variables d'activités zonales) et d'autre part les hypothèses de stabilité sur les coefficients en l'absence de données d'enquêtes répétées ou de données en série chronologique.

Chapitre 7 – Distribution des déplacements

La distribution constitue généralement la seconde étape des modèles à quatre étapes. Elle prend place après la génération dont elle utilise les sorties : émissions et attractions. La distribution correspond au choix de la destination du déplacement dans la séquence des choix de l'individu. Mathématiquement, elle permet de calculer la matrice origines-destinations des déplacements à partir de ses marges.

Nous utilisons un plan similaire à celui retenu pour l'étape de génération. Nous commençons par préciser certaines définitions (section 1), pour présenter ensuite les différentes familles de modèles disponibles : modèles à facteurs de croissance qui sont les premiers à avoir été développés en milieu urbain, modèles gravitaires, modèles basés sur la maximisation de l'entropie et enfin modèles d'opportunité (section 2). La dernière section est consacrée à l'analyse des limites de l'étape de distribution qui est probablement une des étapes la plus sensible de l'ensemble de la chaîne à quatre étapes (section 3).

1. DEFINITIONS ET CONCEPTS DE BASE

1.1. Notation

Le modèle de génération permet de calculer les émissions et attractions, c'est-à-dire le nombre de déplacements sortant ou entrant de chacune des zones de l'aire d'étude. A l'issue de la génération, l'origine et la destination des déplacements ne sont pas reliées. C'est l'objectif de l'étape de distribution de les relier pour produire la matrice origines-destinations des déplacements sur l'aire d'étude. La contribution des deux étapes est représentée sur le schéma 7.1.

T_{ij} correspond au nombre de déplacements entre la zone i et la zone j . Le schéma 7.1 peut donc être également repris mathématiquement (tableau 7.1).

		destination			Emissions : $\sum_j T_{ij}$
	Zone	1	2	3	Total
origine	1	T_{11}	T_{12}	T_{13}	$T_{1.} \approx E_1$
	2	T_{21}	T_{22}	T_{23}	$T_{2.} \approx E_2$
	3	T_{31}	T_{32}	T_{33}	$T_{3.} \approx E_3$
Attractions : $\sum_i T_{ij}$	Total	$T_{.1} \approx A_1$	$T_{.2} \approx A_2$	$T_{.3} \approx A_3$	T ou $T_{..} = \sum_{i,j} T_{ij}$

Tableau 7.1 : Distribution des déplacements : notations

En reprenant les notations habituelles, nous avons :

- $T_{i.} = \sum_j T_{ij}$ qui correspond aux émissions de la zone i : E_i ;

- $T_{.j} = \sum_i T_{ij}$ qui correspond aux attractions de la zone j : A_j ;
- $T_{.i}$ ou $T_{i.} = \sum_{j,i} T_{ij}$ qui correspond à la somme des émissions ou à la somme des attractions, soit l'ensemble des déplacements, réalisés par la population cible, internes à la zone d'étude.

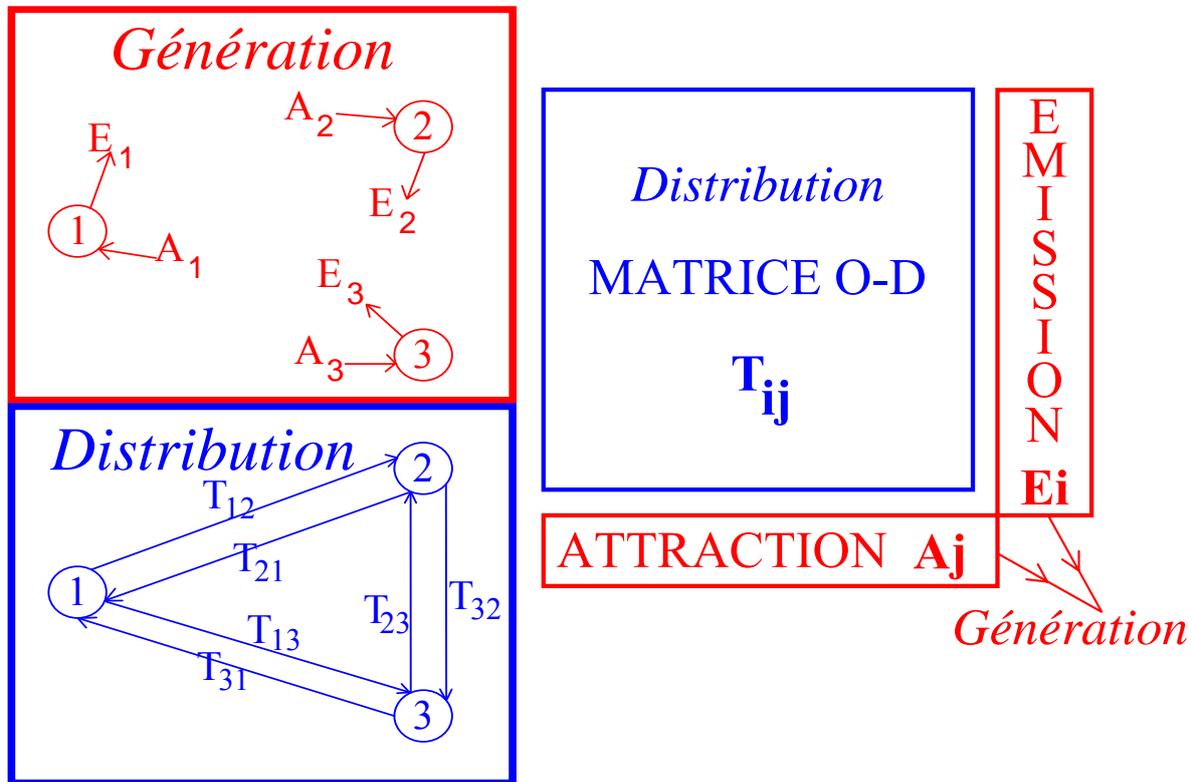


Schéma 7.1 : Distribution des déplacements

Nous reprenons également les notations conventionnelles en représentant :

- par des lettres minuscules les données relatives à la situation de référence, c'est-à-dire celle issue des enquêtes, qui correspond à la situation initiale avant la réalisation du projet ou du plan de déplacements que l'on cherche à simuler : t_{ij} ;
- par des lettres majuscules les données à l'horizon de la prévision : T_{ij} .

1.2. Contraintes aux marges

Le calcul matriciel impose le respect des calculs de sommes sur les lignes et les colonnes. Cette condition est appelée la contrainte aux marges. Elle se formule de la manière suivante :

$$- \sum_j T_{ij} \approx E_i \tag{7.1a}$$

La somme des déplacements sortant de la zone i vers chacune des zones j doit être égale à l'émission de la zone i calculée à l'étape de génération précédente ;

$$- \sum_i T_{ij} \approx A_j \tag{7.1b}$$

La somme des déplacements entrant dans la zone j à partir de chacune des zones i doit être égale à l'attraction de la zone j calculée à l'étape de génération précédente.

Cette contrainte est complétée par l'égalité des sommes des émissions et des attractions que nous avons abordée au chapitre précédent (p. 185).

Si les données sur les attractions ou bien sur les émissions ne sont pas suffisamment fiables compte tenu des données disponibles, la contrainte ne portera que sur l'une des deux dimensions. Toutefois, la qualité du calage ou plutôt de la représentation de la réalité a de fortes chances d'être de moindre qualité. La prévision a donc également toutes les chances de souffrir de cette qualité amoindrie, d'autant plus que les erreurs sur le calage de la distribution ont tendance à s'amplifier en prévision.

1.3. Coût généralisé, temps généralisé

Le coût du déplacement est un élément qui est généralement pris en compte par les modèles de distribution. Il peut prendre la forme d'une distance, mais est le plus souvent mesuré à travers le coût généralisé ou le temps généralisé. Le passage de l'un à l'autre s'effectue en divisant par la valeur du temps. En reprenant les notations du chapitre 5 (équation 5.8, p. 152), on obtient :

$$C_g = P + \left(\sum_i \alpha_i T_i \right) * V_t \text{ ou } T_g = \frac{P}{V_t} + \left(\sum_i \alpha_i T_i \right) \quad (7.2)$$

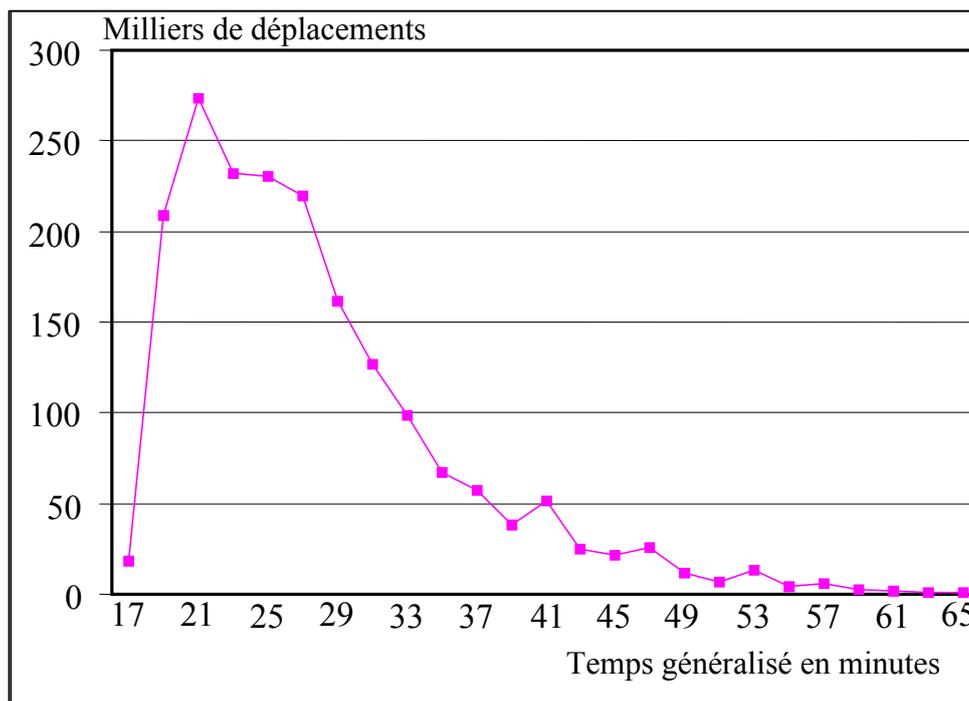
Dans les deux cas, le coût ou le temps ainsi calculé correspond à la perception par l'individu du coût ou du temps du déplacement qui influe sur son processus de choix (chapitre 5, p. 152).

1.4. Répartition des distances de déplacements

Le modèle de distribution vise à construire la matrice origines-destinations des déplacements. Si l'on retenait une méthodologie identique à celle de la génération, le calage du modèle de distribution devrait conduire à reproduire la matrice origines-destinations relative à la situation de référence.

La méthode ne peut toutefois être reprise, car la quantité de données disponibles ne permet pas d'obtenir des données représentatives à ce niveau. Prenons l'exemple de Lyon pour illustrer le problème. 6 000 ménages ont été enquêtés lors de la dernière enquête ménages déplacements en 1995 qui est celle présentant le plus fort effectif parmi les trois enquêtes ménages déplacements réalisées à Lyon au cours de ces 20 dernières années. Ces ménages comprennent 13 997 individus enquêtés qui ont réalisés 53 213 déplacements. Selon les études, les découpages utilisés comprennent généralement entre 200 et 500 zones, soit entre 40 000 et 250 000 cellules pour la matrice. Il y a donc entre 1,3 et 0,2 déplacement en moyenne dans chaque cellule. Sachant que le nombre est forcément entier et que la distribution n'est pas équiprobable, il y a un très grand nombre de cellules qui sont forcément vides. On peut même affirmer que plus de 99 % des cellules ne sont pas statistiquement différentes de 0. A ce niveau de découpage, il n'est donc pas possible d'utiliser la matrice de référence pour le calage.

Nous verrons plus loin quelles sont les procédures de calage utilisées pour faire face à ce problème. Parmi celles-ci, on utilise notamment la courbe de répartition des distances de déplacements (graphique 7.1). La distance utilisée est en principe un temps ou un coût généralisé. Cette distribution est calculée à partir des données d'enquête en situation de référence.



Graphique 7.1 : Forme de la distribution des distances (temps généralisés) des déplacements (source : modèle stratégique appliqué à Lyon LET-SEMALY (Lichère, Raux, 1997a, b))

1.5. Echanges et transit

Les déplacements d'échanges ont une extrémité au sein de la zone d'étude, mais l'autre extrémité en dehors de celle-ci. Les déplacements de transit ont l'origine et la destination externes à la zone d'étude, mais traversent la zone et donc utilisent les infrastructures de transport de la zone.

La connaissance de ces déplacements est généralement nécessaire pour l'affectation de la demande sur les réseaux afin de connaître la charge des infrastructures. La modélisation de ces déplacements n'est toutefois pas prise en compte dans les modèles de distribution. Le plus souvent ces déplacements sont moins bien connus. Pour leur estimation en situation de référence, on utilise le plus souvent les données des enquêtes cordon (chapitre 4, p. 121). Ils font l'objet d'une modélisation spécifique simplifiée (souvent un taux de croissance annuel pouvant dépendre de certaines variables comme le PIB ou les revenus).

2. PRINCIPALES FAMILLES DE MODELES DE DISTRIBUTION

Historiquement en milieu urbain, les modèles à facteurs de croissance (section 2.1) sont les premiers modèles à avoir été développés. Les limites de ces modèles évoquées dans le premier chapitre ont conduit à développer les modèles gravitaires (section 2.2). Leur nom est issu de l'analogie avec la loi de la gravitation universelle de Newton. Ces modèles ont d'abord été développés pour modéliser la demande de déplacements entre agglomérations dès la fin du XIXe siècle. Nous présentons ensuite les modèles découlant de la théorie de la maximisation de l'entropie qui permet en fait de retrouver en apportant une autre justification théorique les modèles à facteurs de croissance ou gravitaires (section 2.3). Enfin, nous présentons des modèles peu connus car assez rarement appliqués : les modèles d'opportunité (section 2.4).

2.1. Modèles à facteurs de croissance

Le principe du modèle à facteurs de croissance est très simple. Il permet à partir de la matrice de référence par déformations successives de parvenir à la matrice horizon en respectant la contrainte aux marges.

Le point de départ est donc constitué par la matrice de référence. Les marges de cette matrice correspondent aux émissions et attractions en situation de référence. Le modèle de génération permet de calculer les nouvelles marges de la matrice horizon, soit les émissions et attractions à l'horizon d'étude. La matrice de référence ne respecte évidemment pas ces nouvelles contraintes. Le principe du modèle est donc de déformer progressivement cette matrice afin de satisfaire à ces nouvelles contraintes.

Les méthodes utilisées par les modèles à facteurs de croissance se distinguent principalement par l'algorithme de déformation de la matrice origine.

2.1.1. Facteur uniforme

La méthode la plus simple utilise un facteur de croissance uniforme. Cette méthode n'est évidemment pas très réaliste et reste confinée à des contextes où l'agglomération évolue peu et de manière uniforme sur le court terme ou à des contextes où l'on ne dispose pas de données permettant de caler un modèle de génération. Le facteur de croissance est évidemment :

$$F = T / t, \text{ de telle sorte que pour chaque case de la matrice } T_{ij} = F * t_{ij}. \quad (7.3)$$

Les tableaux 7.2. et 7.3. fournissent un exemple.

Origine \ destination	1	2	3	e_i
1	5	100	50	155
2	150	10	120	280
3	200	350	15	565
a_i	355	460	185	1000

Tableau 7.2 : Matrice de référence

Origine \ destination	1	2	3	E_i	$F = T / t$
1	6	120	60	186	1,2
2	180	12	144	336	1,2
3	240	420	18	678	1,2
A_i	426	552	222	1200	
$F = T / t$	1,2	1,2	1,2		

Tableau 7.3 : Matrice à l'horizon d'étude avec un facteur de croissance uniforme (= 1,2)

2.1.2. Définition du facteur de croissance

Le plus souvent, on dispose de données permettant de construire un modèle de génération et donc d'utiliser l'information disponible sur les marges de la matrice.

Le facteur de croissance en ligne et en colonne est alors défini à partir des marges de la matrice en situation de référence et à l'horizon :

$$\text{Sur les lignes } F_i = \frac{E_i (\text{horizon})}{e_i (\text{référence})} \quad (7.4a)$$

$$\text{et sur les colonnes } F_j = \frac{A_j (\text{horizon})}{a_j (\text{référence})} \quad (7.4b)$$

Si l'on reprend l'exemple du tableau 7.2 et que l'on dispose des résultats du modèle de génération, le problème peut être illustré à travers le tableau 7.4.

La ligne des A_j fournit le résultat du modèle de génération sans égaliser la somme des émissions et attractions. La procédure d'égalisation conduit à la ligne des A'_j . Vient ensuite le calcul des facteurs de croissance en ligne et en colonne.

Ces facteurs de croissance sont ensuite utilisés afin de calculer les cases de la matrice. La méthode la plus simple (mais pas la plus rapide en termes de convergence) consiste à prendre la moyenne des facteurs de croissance en ligne et en colonne pour calculer le trafic entre i et j . L'équation 7.3 devient alors :

$$T_{ij} = \frac{1}{2} (F_i + F_j) * t_{ij} \quad (7.5)$$

O/D	1	2	3		E_i	$F_i = E_i / e_i$
1	?	?	?		250	1,61
2	?	?	?		280	1,00
3	?	?	?		600	1,06
A_j	500	400	300	$(\sum A_j)=1200$	$(\sum E_i)=1130$	
A'_j	471	377	283	$(\sum A_j)=1130$		
$F_j = A'_j / a_j$	1,33	0,82	1,53			

Tableau 7.4 : Facteurs de croissance en ligne et en colonne

Les trafics T_{ij} ne respectent pas encore les marges après ce premier calcul. L'opération doit donc être renouvelée plusieurs fois avant d'obtenir une différence acceptable entre les sommes des T_{ij} en ligne et en colonne et les émissions et attractions à l'horizon d'étude.

D'autres méthodes sont disponibles pour obtenir une convergence plus rapide. Nous présentons ici les deux méthodes les plus fréquemment utilisées par les logiciels de modélisation de la demande de transports pour résoudre le problème : algorithme de Furness (qui est le plus répandu) et algorithme de Fratar. Un panorama plus complet des méthodes disponibles pour résoudre ce problème a été établi par Madre (Madre, 1980). L'auteur analyse les propriétés et les avantages et limites des différentes méthodes.

2.1.3. Méthode de Furness

Le principe de la méthode est très simple. Il consiste à traiter successivement les lignes puis les colonnes afin qu'elles respectent la contrainte aux marges, puis de renouveler l'opération jusqu'à l'obtention d'une précision suffisante. Compte tenu de la forte incertitude sur les résultats de l'étape de

distribution, 3 à 5 itérations sur les lignes et les colonnes sont en général suffisantes pour obtenir un résultat satisfaisant.

Mathématiquement la méthode est la suivante :

itération 1 : calcul des $F_i^{(1)} = \frac{E_i}{e_i}$, puis calcul des $T_{ij}^{(1)} = t_{ij} * F_i^{(1)}$ (7.6)

par le mode de calcul, la contrainte aux marges est respectée sur les lignes, mais pas les colonnes ;

itération 2 : calcul des $F_j^{(1)} = \frac{A_j}{\sum_i T_{ij}^{(1)}}$, puis calcul des $T_{ij}^{(2)} = T_{ij}^{(1)} * F_j^{(1)}$ (7.7)

par le mode de calcul, la contrainte aux marges est maintenant respectée sur les colonnes, mais plus sur les lignes ;

itération 3 : calcul des $F_i^{(2)} = \frac{E_i}{\sum_j T_{ij}^{(2)}}$, puis calcul des $T_{ij}^{(3)} = T_{ij}^{(2)} * F_i^{(2)}$ (7.8)

par le mode de calcul, la contrainte aux marges est de nouveau respectée sur les lignes, mais plus les colonnes ;
et ainsi de suite....

Si l'on reprend notre exemple des tableaux 7.2 et 7.4, le résultat à l'issue de la 4^e itération sur les lignes (soit 7 itérations au total sur les lignes ou les colonnes) figure dans le tableau 7.5. Les facteurs de croissance sur les lignes utilisés pour cette itération sont déjà compris entre 0,99 et 1,01. A l'issue de la 5^e itération (soit 10 itérations au total), les facteurs de croissance sur les colonnes sont compris entre 0,999 et 1,001.

Itération 4 (lignes)						
O/D	1	2	3	$\sum_j T_{ij}$	E_i	F_i
1	11,63	111,27	127,10	250,00	250	1,00
2	146,88	4,69	128,43	280,00	280	0,99
3	312,62	261,75	25,63	600,00	600	1,01
$\sum_i T_{ij}$	471,14	377,71	281,15	1130,00		
A_j	470,83	376,67	282,50			

Tableau 7.5 : Résultat de l'algorithme de Furness (4^e itération)

Cette formulation itérative est souvent présentée mathématiquement à travers l'équation 7.9, qui est en fait une reprise des équations 7.3 (p. 195) ou 7.5 (p. 196) en faisant apparaître explicitement les émissions et les attractions :

$$T_{ij} = \alpha_i \beta_j E_i A_j * t_{ij} \tag{7.9}$$

Les facteurs α_i et β_j sont appelés les facteurs de balancement dans la mesure où ils sont recalculés alternativement à chaque nouvelle itération. A la première itération, α_i se déduit de l'équation 7.6 :

$$\alpha_i = \frac{1}{e_i} ; \text{ tandis que } \beta_j = \frac{1}{a_j} , \tag{7.10}$$

pour les itérations suivantes, on utilise les équations 7.7, 7.8 et ainsi de suite.

2.1.4. Méthode de Fratar

Le principe est similaire à celui de Furness, si ce n'est que les facteurs de croissance sur les lignes et les colonnes sont traités simultanément en ajoutant un terme λ_{ij} que l'on peut assimiler à un facteur de résistance qui est fonction de la croissance des autres zones :

$$\lambda_{ij} = \left[\frac{L_i + L_j}{2} \right], \text{ avec } L_i = \frac{\sum_j t_{ij}}{\sum_j t_{ij} * F_j} \text{ et } L_j = \frac{\sum_i t_{ij}}{\sum_i t_{ij} * F_i} \quad (7.11)$$

On obtient ensuite T_{ij} par la formule suivante :

$$T_{ij} = t_{ij} * F_i * F_j * \lambda_{ij} \quad (7.12)$$

où F_i et F_j sont les facteurs de croissance.

Comme pour l'algorithme de Furness, l'équation 7.12 peut être réécrite sous la forme de l'équation 7.9, faisant apparaître explicitement les émissions et les attractions à l'horizon d'étude.

Itération 3						
O/D	1	2	3	$\sum_j T_{ij}$	E_i	F_i
1	11,57	112,03	126,24	249,84	250	1,01
2	148,96	4,76	129,02	282,74	280	0,97
3	309,56	262,50	25,36	597,42	600	1,01
$\sum_i T_{ij}$	470,09	379,29	280,62	1130,00		
A_j	470,83	376,67	282,50			
F_j	1,01	0,97	1,02			
λ_{ij}	1	2	3	L_i		
1	1,00	1,00	1,00	1,00		
2	0,99	0,99	1,00	0,99		
3	1,00	1,00	1,01	1,01		
L_j	1,00	0,99	1,01			

Tableau 7.6 : Résultat de l'algorithme de Fratar (3^e itération)

A l'itération suivante, on recalcule tous les termes en remplaçant t_{ij} par le T_{ij} obtenu à l'itération précédente.... La rapidité de la convergence de cet algorithme est similaire à celle de l'algorithme de Furness, même si elle peut être plus rapide dans certaines configurations de données. Toujours sur le même jeu de données, les résultats à l'issue de la 3^e itération figurent dans le tableau 7.6 avec la valeur des facteurs ayant servi à ce calcul. On vérifiera évidemment que les deux algorithmes convergent vers la même solution (au degré de précision attendu près).

2.1.5. Avantages et limites des facteurs de croissance

Le principal avantage des méthodes à facteurs de croissance est leur simplicité de calcul. L'algorithme peut être très facilement programmé sous tableur. La méthode est également facile à comprendre et semble intuitivement plutôt pertinente. Enfin, elle permet de conserver la structure de la matrice de référence qui n'est sujette qu'aux déformations liées à l'évolution des zones origines et destinations.

Cet avantage est aussi sa principale limite. La pertinence de la matrice horizon dépend fortement de la qualité de la matrice de référence. L'utilisation d'un découpage assez fin ne permet pas de garantir cette qualité. Même avec un découpage fortement agrégé comme dans le cas du modèle stratégique développé sur l'agglomération lyonnaise (Lichère, Raux, 1997a) qui ne comprend que 25 zones, la précision d'un grand nombre d'origines-destinations reste très faible. De ce fait, dans la plupart des cas, de nombreuses cellules de la matrice sont nulles. La faible précision ne permet pas de s'assurer de cette nullité avec une précision suffisante. Inversement de nombreuses cellules, bien qu'elles ne soient pas nulles ne sont pas statistiquement différentes de 0, même à un seuil de confiance assez faible. Avec les deux algorithmes présentés ci-dessus, une case nulle en référence restera nulle à l'horizon d'étude quelle que soit la valeur des facteurs de croissance en ligne et en colonne (il existe toutefois d'autres méthodes permettant de modifier une case nulle, mais elle présente d'autres problèmes qui ont été analysés par Madre (Madre, 1980)).

La seconde limite principale tient dans la non prise en compte de l'offre de transport. Une modification du système, même importante, n'aura aucun effet sur la matrice horizon. Cette méthode ne peut donc être utilisée pour tester des scénarios contrastés de politiques de transport.

La pertinence de cette méthode reste donc assez limitée. Son domaine d'utilisation reste confiné à un horizon temporel assez limité à la condition que la structure spatiale des activités n'évolue pas trop fortement. De plus, le découpage utilisé ne pourra pas être trop fin et devra toujours être compatible avec la quantité de données disponible en termes de nombre de déplacements enquêtés. Enfin, la méthode ne pourra pas être utilisée pour tester des mesures contrastées de politiques de transport ou incluant une modification significative du système de transport.

Pour ces raisons, les modèles à facteurs de croissance ont été assez rapidement remplacés par les modèles gravitaires.

2.2. Modèles gravitaires

2.2.1. Principe

Le principe du modèle gravitaire est très simple. Il s'appuie sur une analogie avec la loi de la gravitation universelle de Newton :

$$F_{ij} = \frac{G * (M_i * M_j)}{d_{ij}^2} \quad (7.13)$$

où F_{ij} représente l'attraction entre deux corps de masse M_i et M_j ;

d_{ij} est la distance séparant ces deux corps ;

et G est une constante.

L'analogie est assez intuitive. L'attraction entre deux zones, c'est-à-dire le trafic, sera d'autant plus forte que le poids de chacune des zones sera important, mais cette attraction est pondérée par l'importance de la distance séparant les deux zones.

Les premières applications, apparues dès la fin du XIXe siècle, concernent le trafic interurbain. Il s'agit d'estimer le trafic entre deux agglomérations en fonction de leur taille. L'expression de la taille ou du poids de l'agglomération s'exprime alors le plus souvent à travers la population et les revenus de chacune des agglomérations. La distance devient assez vite une distance généralisée (ou un temps généralisé) et non plus seulement une distance géométrique pour tenir compte des performances du système de transport entre les deux agglomérations :

$$T_{ij} = \frac{K * (P_i + P_j)^\alpha * (R_i + R_j)^\beta}{Cg_{ij}^\gamma} \quad (7.14)$$

où P_i et P_j sont la population de la ville i et de la ville j ;

R_i et R_j les revenus des villes i et j ;

Cg_{ij} le coût généralisé du déplacement entre i et j ;

K est une constante de proportionnalité ;

α , β , γ des élasticités à la population (proche de 1), au revenu et au coût généralisé.

2.2.2. Transposition à l'urbain

La transposition au contexte urbain s'est faite en remplaçant la population et les revenus (ou l'ensemble des facteurs pouvant exprimer la propension à se déplacer) par les émissions et les attractions. Ces dernières expriment le poids de chacune des zones parmi l'ensemble des déplacements réalisés au sein d'une agglomération. Le coût généralisé entre les deux zones exprime toujours le facteur résistant, limitant la quantité de déplacements entre deux zones lorsque le coût augmente. Toutefois, d'autres fonctions du coût généralisé sont apparues afin d'améliorer le calage des fonctions gravitaires sur les données disponibles. La forme générique devient ainsi :

$$T_{ij} = a E_i A_j f(Cg_{ij}) \quad (7.15)$$

où a est un coefficient de proportionnalité, E_i l'émission de la zone i , A_j l'attraction de la zone j .

L'utilisation du modèle en situation de prévision ne garantit plus le respect de la contrainte aux marges. Il est donc nécessaire d'introduire des coefficients pour traduire ces contraintes :

$$T_{ij} = \alpha_i \beta_j E_i A_j f(Cg_{ij}) \quad (7.16)$$

où α_i et β_j sont les coefficients traduisant le respect de la contrainte aux marges, ce qui nous permet de les calculer :

$$\alpha_i = \frac{1}{\sum_l \beta_l A_l f(Cg_{il})} \quad \text{et} \quad \beta_j = \frac{1}{\sum_k \alpha_k E_k f(Cg_{kj})} \quad (7.17)$$

Les coefficients α_i et β_j étant interdépendants, le calcul des T_{ij} nécessite donc une procédure itérative. Les algorithmes de Furness (le plus fréquent) ou de Fratar permettent de conduire cette procédure itérative. Les coefficients α_i et β_j deviennent ainsi les facteurs de balancement de l'algorithme de Furness ou de Fratar. Ils sont recalculés à l'issue de chaque itération permettant une convergence en générale assez rapide (le plus souvent 4 à 5 itérations sont suffisantes comme pour le modèle à facteurs de croissance).

L'équation 7.16 est parfois présentée en intégrant directement la formulation des coefficients α_i et β_j dans l'équation, conduisant à la formulation suivante :

$$T_{ij} = \frac{\beta_j E_i A_j f(Cg_{ij})}{\sum_l \beta_l A_l f(Cg_{il})} \quad (7.18)$$

où les β_l sont définis de telle sorte que la contrainte aux marges sur les colonnes soit respectée (équation 7.1b : $\sum_i T_{il} = A_l$, p. 192)

L'équation 7.18 peut également être présentée de la manière suivante en inversant le rôle des α_i et β_j :

$$T_{ij} = \frac{\alpha_i A_j E_i f(Cg_{ij})}{\sum_k \alpha_k E_k f(Cg_{kj})} \quad (7.19)$$

où les α_k sont définis de telle sorte que la contrainte aux marges sur les lignes soit respectée (équation 7.1a : $\sum_j T_{kj} = E_k$, p. 192)

Cette formulation du modèle de distribution (équations 7.18 ou 7.19) est également appelée modèle gravitaire doublement contraint. La mention doublement contraint fait référence à la double contrainte sur les marges d'une part sur les émissions et d'autre part sur les attractions.

Dans certains cas, les données disponibles ne permettent pas de forcer au respect de la contrainte aux marges à la fois sur les émissions et sur les attractions. Supposons que l'on force au respect de la contrainte aux marges uniquement sur les émissions, les équations 7.18 et 7.19 se résument alors dans l'équation 7.20 :

$$T_{ij} = \frac{E_i A_j f(Cg_{ij})}{\sum_l A_l f(Cg_{il})} \quad (7.20)$$

qui correspond au modèle gravitaire simplement contraint.

L'équation 7.15 correspond quant à elle à un modèle gravitaire sans contrainte aux marges. Le choix entre ces trois formes de modèles (sans contrainte, contrainte simple ou double) dépend des données disponibles et de leur fiabilité. En interurbain, c'est le modèle sans contrainte qui est privilégié. De ce fait, le trafic entre deux zones (généralement deux agglomérations) ne dépend que des caractéristiques de ces zones et du coût généralisé du déplacement entre ces deux zones. En revanche, en urbain, on utilise le modèle simplement ou doublement contraint qui fait dépendre le trafic entre deux zones, non seulement des caractéristiques de ces deux zones et du coût généralisé entre ces zones, mais aussi des caractéristiques des autres zones et des autres coûts généralisés. Le modèle doublement contraint doit être privilégié lorsque la fiabilité des données tant en émission qu'en attraction le permet.

2.2.3. Fonction de résistance ...

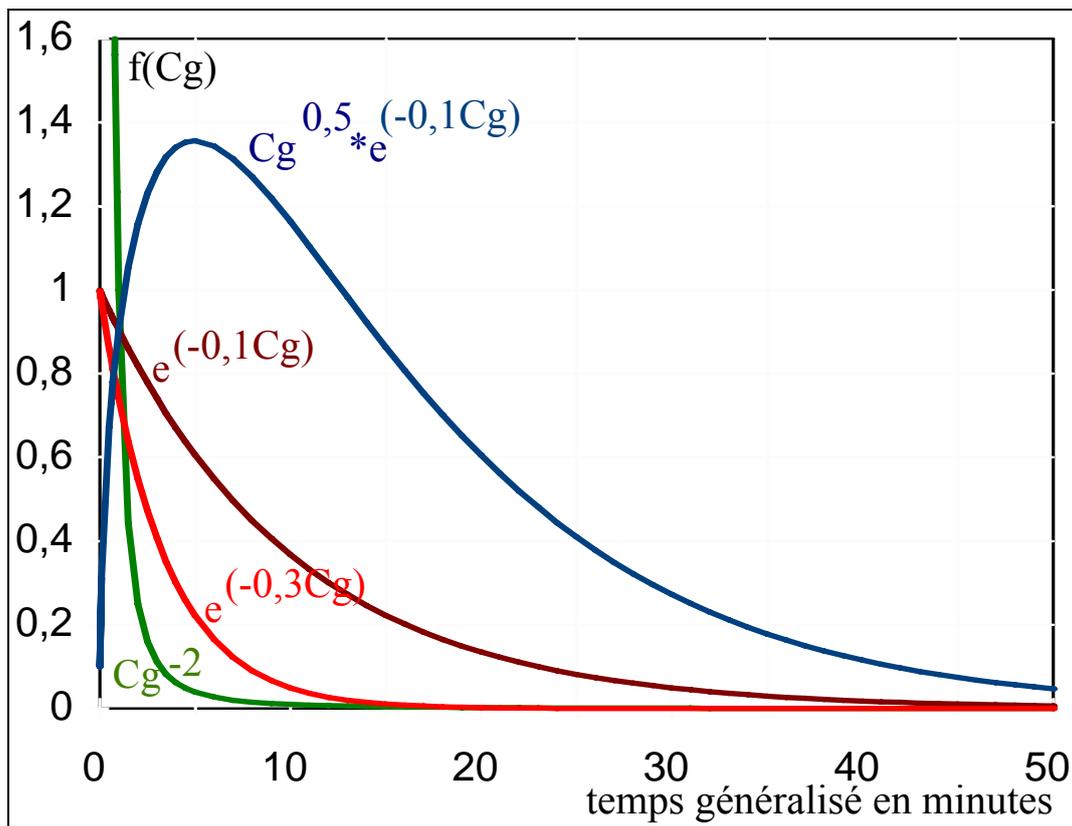
La fonction f connaît de nombreuses appellations : fonction de résistance, de dissuasion, d'impédance, de friction ou parfois encore de déterrence (probablement par « traduction » de l'anglais *deterrence function* qui signifie fonction de dissuasion). Elle prend le plus souvent l'une des trois formes suivantes :

$$f(Cg_{ij}) = Cg_{ij}^{-b} \text{ fonction puissance ;} \quad (7.21a)$$

$f(Cg_{ij}) = e^{-bCg_{ij}}$ fonction exponentielle, le coefficient b est parfois appelé coefficient de conductance ; (7.21b)

$f(Cg_{ij}) = Cg_{ij}^{-c} * e^{-dCg_{ij}}$ fonction combinée exponentielle et puissance. (7.21c)

Les deux dernières formes sont en général privilégiées car elles permettent un meilleur calage sur les données d'enquête. La forme générale des courbes est présentée au graphique 7.2.



Graphique 7.2 : Formes de la courbe de résistance

Le calage du modèle de distribution consiste à déterminer la valeur des inconnues de la fonction de résistance (coefficient b ou c et d selon la fonction choisie, section 2.2.6, p. 205).

On utilise parfois une autre forme de fonction de résistance. L'objectif dans ce cas est de mieux coller aux données empiriques :

$$f(Cg_{ij}) = \sum_k F^k \delta_{ij}^k \quad (7.22)$$

Cette forme est intéressante lorsque les fonctions précédentes ne permettent pas un bon calage de la distribution des longueurs de déplacements (en termes de coût ou de temps généralisé). Le principe est de segmenter la distribution des coûts généralisés en un certain nombre de classes (de 10 à 15 selon Ortuzar et Willumsen (1994)). Chaque classe k a alors une valeur F^k pour la fonction de coût généralisé qui est généralement la valeur de la fonction de résistance observée pour la valeur moyenne des coûts généralisés des déplacements appartenant à la classe k . Le paramètre δ_{ij}^k prend la

valeur 1 si le coût généralisé sur l'origine-destination est inclus dans la classe k et la valeur 0 autrement.

Cette méthode permet évidemment de mieux reproduire la distribution des coûts généralisés puisque l'on dispose d'un plus grand nombre de facteurs de calage comparativement aux fonctions puissance et exponentielle qui n'en ont qu'un et la combinaison des deux fonctions qui n'en a que deux.

2.2.4. ... ou fonction de demande

Jusqu'ici le modèle gravitaire est justifié par analogie avec la théorie physique de la loi de gravitation de Newton. On peut en fait en donner une autre interprétation se rapprochant de la théorie développée au chapitre 5.

La fonction de résistance est alors interprétée comme illustrant la fonction de demande. Reprenons les concepts utilisés au chapitre 5. Le graphique 5.6 (p. 150) illustre, pour un état donné du sous-système de localisation des activités, la fonction de demande en fonction du niveau de service (équation 5.6, p. 149), c'est-à-dire une certaine mesure de l'état du sous-système des transports. De plus, le coût généralisé est l'expression généralement utilisée pour formaliser le niveau de service (équation 5.8, p. 152).

Ces éléments peuvent être repris pour commenter les courbes du graphique 7.2. A mesure que le coût généralisé s'accroît (c'est-à-dire que le niveau de service décroît), la fonction du coût généralisé décroît et donc le trafic sur l'origine-destination décroît (c'est-à-dire la fonction de demande décroît). On retrouve donc une courbe similaire à celle du graphique 5.6 (p. 150, croissance en fonction du niveau de service ou décroissance en fonction du coût généralisé).

L'interprétation peut être prolongée. Les émissions et les attractions représentent en fait le sous-système de localisation des activités. Le précédent chapitre nous a en effet montré que les émissions et attractions sont calculées à partir des caractéristiques d'activités des zones. L'utilisation du modèle gravitaire s'effectuant à un moment donné, correspondant à un certain état du sous-système de localisation des activités, les émissions et attractions sont logiquement constantes.

Cette interprétation permet de proposer une autre justification théorique du modèle gravitaire. Il est possible en fait d'en donner une justification économique plus rigoureuse. Nous reprenons pour cela les travaux de Cochrane (1975). La section 2.3 (p. 207) nous permettra d'en donner une nouvelle justification s'appuyant sur la maximisation de l'entropie.

2.2.5. Justification économique du modèle gravitaire

Cochrane (1975) fournit une justification du modèle gravitaire cohérente avec la théorie de la maximisation de l'utilité individuelle. On pose pour cela que chaque destination potentielle j_k au sein de la zone j a une utilité U_{jkn} inconnue pour l'individu n , mais qui peut s'exprimer de la manière suivante :

$$U_{jkn} = V + \varepsilon_{jkn} \quad (7.23)$$

où V est une quantité positive, mais inconnue ;

ε_{jkn} est une variable aléatoire.

Pour une origine i et une destination potentielle j_k située dans la zone j , l'utilité deviendra :

$$U_{jkn} = V - Cg_{ij} + \varepsilon_{jkn} \quad (7.24)$$

avec Cg_{ij} le coût généralisé du déplacement de i à j .

Pour résoudre le problème, il est nécessaire de postuler une distribution pour les ε_{jkn} . On retient une distribution de Gumbel (chapitre 9, p. 255). Il est de plus nécessaire de postuler que les ε_{jkn} ont des distributions indépendantes et identiques de paramètres $(0, \mu)$. En vertu des propriétés de la loi de Gumbel (chapitre 9, propriété 6 de la loi de Gumbel, p. 256), le maximum de U_{jkn} (noté U_{jn}) au sein de la zone j suit lui aussi une loi de Gumbel de paramètres :

$$\left(\frac{1}{\mu} \ln \left(\sum_k e^{\mu(V - Cg_{ij})} \right), \mu \right) \quad (7.25)$$

le terme $\mu(V - Cg_{ij})$ étant constant pour une zone donnée, si l'on pose N_j le nombre de destinations potentielles de la zone j , l'équation 7.25 devient :

$$\left(\frac{1}{\mu} (\ln(N_j) + \mu(V - Cg_{ij})), \mu \right) \quad (7.26)$$

Le nombre de destinations potentielles dans la zone j est fonction des opportunités de la zone j . Il est aussi égal à l'attraction de la zone j si l'on a utilisé un modèle de génération. L'équation 7.26 devient :

$$\left(\frac{1}{\mu} (\ln(A_j) + \mu(V - Cg_{ij})), \mu \right) \quad (7.27)$$

La propriété 4 de la loi de Gumbel (chapitre 9, p. 255 et équation 9.29, p. 258), permet d'écrire :

$$U_{jn} = V_j + \varepsilon_{jn} \quad (7.28)$$

avec $V_j = \frac{1}{\mu} (\ln(A_j) + \mu(V - Cg_{ij}))$;

et ε_{jn} une variable aléatoire qui suit une loi de Gumbel de paramètres $(0, \mu)$.

Dans ce cas, l'application du logit multinomial (équation 9.19, p. 254) nous permet d'écrire :

$$P(j) = \frac{e^{\mu V_j}}{\sum_m e^{\mu V_m}} = \frac{e^{\mu \left(\frac{1}{\mu} (\ln(A_j) + \mu(V - Cg_{ij})) \right)}}{\sum_m e^{\mu \left(\frac{1}{\mu} (\ln(A_m) + \mu(V - Cg_{im})) \right)}}$$

$$P(j) = \frac{A_j e^{\mu(V - Cg_{ij})}}{\sum_m A_m e^{\mu(V - Cg_{im})}} = \frac{A_j e^{-\mu(Cg_{ij})}}{\sum_m A_m e^{-\mu(Cg_{im})}}$$

Si l'on reprend le nombre de déplacement émis par la zone i , on obtient :

$$T_{ij} = \frac{E_i A_j e^{-\mu(Cg_{ij})}}{\sum_m A_m e^{-\mu(Cg_{im})}} \quad (7.29)$$

On retrouve l'équation 7.20 (p. 201) du modèle gravitaire simplement contraint. Dans ce modèle, seule la contrainte aux marges sur les émissions est respectée. En revanche, la contrainte sur

les attractions n'est pas respectée. Cochrane (1975) l'explique en indiquant que dans le modèle simplement contraint plusieurs déplacements peuvent avoir la même destination potentielle, ce qui n'est pas possible (prenons l'exemple des déplacements domicile-travail pour expliciter le propos. Un emploi ne peut être occupé que par individu et non par plusieurs individus à la fois. L'analogie peut être transposée aux autres motifs en retenant comme nombre de destinations possibles dans une zone le nombre d'attractions de cette zone). Il est donc nécessaire de corriger l'utilité d'avoir une destination dans la zone j par un facteur additif λ_j . Dans ce cas, l'équation 7.28 devient :

$$V_j = \frac{1}{\mu} \left(\ln(A_j) + \mu(V - C_{gij} - \lambda_j) \right); \quad (7.30)$$

et l'équation 7.29 devient :

$$T_{ij} = \frac{E_i A_j e^{-\mu(C_{gij} + \lambda_j)}}{\sum_m A_m e^{-\mu(C_{gim} + \lambda_m)}} = \frac{E_i A_j e^{-\mu(\lambda_j)} e^{-\mu(C_{gij})}}{\sum_m A_m e^{-\mu(\lambda_m)} e^{-\mu(C_{gim})}} \quad (7.31)$$

En posant $e^{-\mu(\lambda_j)} = \beta_j$, on retrouve l'équation 7.18 (p. 201) et le respect de la contrainte aux marges sur les attractions et donc le modèle doublement contraint.

Enfin, Cochrane (1975) dérive le modèle gravitaire sans contrainte, en posant comme hypothèse que l'individu n effectue un déplacement entre i et j seulement lorsque le coût généralisé C_{gij} est inférieur à la partie positive V de l'utilité. Dans cette dérivation du modèle gravitaire, il n'est pas nécessaire d'introduire de contrainte de budget temps de déplacement ou de contrainte de budget monétaire, ce qui constitue assurément une limite pour ce modèle sans contrainte sur les marges, donc sans contrainte sur le nombre total de déplacements.

2.2.6. Calage du modèle gravitaire

Compte tenu de la taille des échantillons d'enquête, la matrice origines-destinations de référence présente en général une fiabilité très faible au niveau de chacune des cases de la matrice avec un grand nombre de 0. De ce fait, il ne serait pas raisonnable de chercher à reproduire cette matrice lors du calage du modèle de distribution, car on obtiendrait des coefficients de calage présentant également une très faible précision. Il est donc nécessaire de déterminer d'autres indicateurs plus agrégés permettant d'offrir un niveau de précision acceptable. Plusieurs éléments sont pris en compte :

- la distribution des distances de déplacements ;
- la comparaison entre la matrice de référence et la matrice calculée par le modèle s'il y a des cellules de cette matrice qui présentent une fiabilité suffisante (et uniquement pour ces cellules). Et plus généralement la comparaison avec la matrice de référence à un niveau suffisamment agrégé ;
- le trafic traversant une ligne écran.

Nous reprenons ces trois dimensions ci-dessous.

La courbe de distribution des distances de déplacements est suffisamment agrégée pour offrir une précision acceptable pour le calage du modèle de distribution (graphique 7.1, p. 194). La distance utilisée est généralement le coût ou le temps généralisé. C'est l'examen de cette courbe qui permet de privilégier l'une ou l'autre des fonctions de résistance. Il est toutefois évident qu'avec un seul paramètre de calage (ou deux pour la fonction combinée), il peut être assez difficile de caler correctement cette courbe. Par exemple si la concavité de la courbe n'est pas respectée pour la partie des plus fortes valeurs de coût, la courbe ne pourra pas être calée correctement. Ce pourra être le cas si deux zones ou des couples de zones ont une attraction particulière (pour des raisons historiques ou parce qu'un employeur important y loge une partie de son personnel...) qui ne s'explique pas

uniquement par la fonction de résistance et les valeurs des émissions et attractions. Il est toutefois assez logique qu'en ne disposant que d'un (ou deux dans le cas de la fonction combinée) degré de liberté, il ne soit pas possible de caler très finement la courbe. Il arrive ainsi que lorsque l'on se trouve dans la configuration évoquée ci-dessus d'attraction particulière entre deux zones ou couples de zones, on introduise un coefficient correcteur spécifique entre ces deux zones ou couples de zones. La formulation devient alors :

$$t_{ij} = k_{ij} \alpha_i \beta_j e_i a_j f(Cg_{ij}) \quad (7.32)$$

L'utilisation de tels coefficients doit toutefois rester très limitée. En généralisant, l'utilisation de ces coefficients à l'ensemble des couples origines-destinations i-j, il serait toujours possible d'ajuster des coefficients k_{ij} quelle que soit la valeur du(des) coefficient(s) de calage de la fonction de résistance. Le modèle perdrait toutefois toutes ses qualités de prévision. Sauf cas très spécifique, il est donc préférable de ne pas utiliser de tels coefficients.

Lorsque certaines cases de la matrice disposent d'un effectif enquêté suffisant pour garantir une fiabilité acceptable, il peut être intéressant de vérifier la qualité du modèle gravitaire pour les cases correspondantes. Cette possibilité reste toutefois très limitée compte tenu de la taille des échantillons d'enquêtes et du nombre de zones généralement utilisé. Il est donc généralement nécessaire de procéder à des agrégations de zones afin de disposer d'un échantillon enquêté suffisant pour au moins une partie importante des cases de la matrice (même avec un découpage assez grossier, il reste des cases de la matrice présentant une fiabilité assez faible, car le nombre de déplacements sur ces relations est assez faible. C'est souvent le cas de certaines relations entre zones périphériques). Il est alors possible de calculer des indicateurs de distances entre la matrice de référence et la matrice obtenue par le modèle. Ces indicateurs agrégés peuvent être complétés par l'analyse des résultats pour les cellules les moins bien représentées par le modèle ou pour les relations origines-destinations importantes au regard des mesures qui doivent être testées à l'horizon de la prévision. Le choix du découpage agrégé doit être étudié attentivement. Il doit répondre principalement à deux objectifs qui peuvent être partiellement contradictoires :

- fournir un effectif suffisant dans chacune des cases de la matrice au niveau de confiance souhaité ;
- être pertinent au regard des mesures devant être simulées par le modèle à l'horizon d'étude.

La dernière source de données fréquemment utilisée lors du calage du modèle de distribution concerne le trafic traversant une ligne écran. Il faut disposer d'enquêtes ou de comptages sur les flux traversant cette ligne écran. L'analyse n'est pas toujours aisée pour plusieurs raisons :

- la détermination des origines-destinations concernées par la ligne écran. Les origines-destinations peuvent être déterminées a priori. Mais il est également possible de ne retenir que les déplacements qui traversent effectivement cette ligne à l'issue du modèle d'affectation ;
- les données disponibles ne concernent pas toujours l'ensemble des modes de déplacements. Bien souvent seuls les comptages routiers sont disponibles faute de moyen permettant d'organiser une enquête sur la ligne écran auprès des individus la traversant ;
- les déplacements traversant la ligne écran ne sont pas tous inclus dans la population cible de l'enquête déplacements utilisée pour le calage. Si l'on prend l'exemple des enquêtes ménages déplacements françaises, les déplacements d'échanges ou de transit, les déplacements professionnels, les déplacements de marchandises et les déplacements réalisés par des non-résidents permanents de la zone d'enquête sont exclus de la population cible. Selon la position de la ligne écran, ces déplacements peuvent représenter une part non négligeable des déplacements la traversant, part qui est généralement inconnue.

Le choix des lignes écrans dépend bien souvent des données disponibles. Mais lorsque le choix est possible, celui-ci doit être guidé par les objectifs d'utilisation du modèle en prévision ou par des

considérations topographiques qui facilitent l'analyse. Ainsi, on utilise souvent les ponts d'un fleuve traversant l'agglomération ce qui permet d'isoler plus aisément les origines-destinations concernées.

2.3. Maximisation de l'entropie

Nous présentons tout d'abord le principe général de la maximisation de l'entropie (section 2.3.1) avant de montrer que cette approche permet de générer le modèle gravitaire (section 2.3.2).

2.3.1. Principe général

Pour cette présentation, nous empruntons largement à Ortuzar et Willumsen (1994) dont le mérite est la clarté comparativement à des présentations beaucoup plus mathématiques qui pourraient rebuter le lecteur peu familier avec ce langage. On pourra toutefois trouver une présentation plus académique dans Wilson (1974).

Considérons un système fait d'un très grand nombre d'éléments distincts. Une description complète d'un tel système nécessite une spécification de tous les micro-états, c'est-à-dire une énumération de tous les individus avec la description des caractéristiques de chacun des déplacements réalisés. Une telle énumération est hors de portée et de plus non nécessaire dans la mesure où bien souvent seule la description agrégée ou méso-état suffit. Un tel méso-état peut correspondre à la matrice origines-destinations somme des micro-états.

Ce méso-état résulte d'un très grand nombre de combinaisons différentes possibles de micro-états différents. Par exemple deux individus résidant dans une même zone qui échangent leur destination correspondent à deux micro-états différents, mais conduisent à la même matrice origines-destinations donc au même méso-état.

Il y a un état encore plus agrégé du système qui se situe au niveau macro, qui correspond au macro-état du système. Dans notre cas, il peut s'agir du nombre total de déplacements, ou du nombre de déplacements émis ou reçus par zone ou de toutes autres données agrégées. Compte tenu des moyens (et donc des données) disponibles, seul ce macro-état peut être connu à un niveau de précision suffisant. De plus, dans une perspective prévisionnelle, l'incertitude sur les données estimées en prévision est moindre à ce macro-état qu'elle ne l'est au niveau des méso-états. Il est plus facile de prévoir la population par zone à un horizon donné, que sa décomposition selon certaines classes pour chacune des zones.

L'hypothèse de base des approches de maximisation de l'entropie est que tous les micro-états du système sont équiprobables à moins que l'on dispose d'informations avérant le contraire. Cette hypothèse semble acceptable dans la mesure où l'on ne connaît pas le système au niveau de ses micro-états. Le principe de la méthode est alors de trouver le méso-état le plus probable qui satisfasse les connaissances que l'on a sur le macro-état du système. C'est-à-dire le méso-état le plus probable compte tenu des contraintes posées sur le macro-état. Les contraintes traduisant mathématiquement les connaissances que l'on a sur le macro-état du système.

On peut montrer que le nombre de micro-états $W\{T_{ij}\}$ correspondant au méso-état T_{ij} est obtenu par :

$$W\{T_{ij}\} = \frac{T!}{\prod_{i,j} T_{ij}!} \quad (7.33)$$

Comme l'on a posé par hypothèse que tous les micro-états sont équiprobables, le méso-état le plus probable est celui qui pourra être généré par le plus grand nombre de micro-états différents. On est donc en présence d'un problème de maximisation. Pour simplifier le problème, on peut rechercher le maximum du log de $W\{T_{ij}\}$ dans la mesure où les deux termes ont le même maximum :

$$\log W\{T_{ij}\} = \log \frac{T!}{\prod_{i,j} T_{ij}!} = \log T! - \sum_{i,j} \log T_{ij}! \quad (7.34)$$

En utilisant l'approximation de Stirling pour $\log X! = X \log X - X$, on obtient :

$$\log W\{T_{ij}\} = \log T! - \sum_{i,j} (T_{ij} \log T_{ij} - T_{ij}) \quad (7.35)$$

T est une constante, $\log T!$ aussi. Le problème d'optimisation devient :

$$\log W' = - \sum_{i,j} (T_{ij} \log T_{ij} - T_{ij}) \quad (7.36)$$

L'introduction des contraintes relatives à notre connaissance du macro-état du système permet de résoudre le problème d'optimisation. Il est également possible d'ajouter des informations complémentaires sur la connaissance du méso-état du système ou d'un méso-état antérieur du système. Par exemple la connaissance d'une matrice origines-destinations relative à une date antérieure peut être ajoutée, pour conduire au problème d'optimisation suivant :

$$\log W'' = - \sum_{i,j} (T_{ij} \log \frac{T_{ij}}{t_{ij}} - T_{ij} + t_{ij}) \quad (7.37)$$

Cette fonction est intéressante dans la mesure où chacun des éléments est nul si $T_{ij} = t_{ij}$ ou prend une valeur positive dans le cas contraire. Ainsi $-\log W''$ est une mesure intéressante de la distance entre les deux matrices T et t . De plus, il est possible de montrer que l'on peut approximer $-\log W''$ par la formule suivante :

$$-\log W'' \approx 0,5 \sum_{i,j} \frac{(T_{ij} - t_{ij})^2}{t_{ij}} \quad (7.38)$$

qui est connue pour être une bonne mesure de la distance entre deux matrices. Il est ainsi possible de déterminer la matrice T_{ij} la plus proche (au sens de la distance de l'équation 7.37 ou de son estimation approchée équation 7.38) de la matrice t_{ij} antérieure connue et qui respecte les contraintes sur le macro-état du système. Cette solution est particulièrement intéressante lorsque l'on connaît avec précision une matrice relative à un état antérieur du système (ce qui est malheureusement assez rarement le cas).

2.3.2. Génération du modèle gravitaire

Pour résoudre le problème de maximisation, il est nécessaire de traduire mathématiquement les connaissances que l'on a sur le macro-état du système. Cette connaissance découle des résultats de l'étape de génération à savoir le calcul des émissions et des attractions. Cette connaissance est traduite par l'intermédiaire de la contrainte aux marges :

$$\sum_j T_{ij} = E_i \text{ ou encore } E_i - \sum_j T_{ij} = 0 \quad (7.39a)$$

$$\sum_i T_{ij} = A_j \text{ ou encore } A_j - \sum_i T_{ij} = 0 \quad (7.39b)$$

Enfin, tous les flux sont forcément positifs ou nuls :

$$\forall i, j \quad T_{ij} \geq 0 \quad (7.40)$$

Ce problème de maximisation sous contraintes est résolu en formant le Lagrangien :

$$L = \log W' + \sum_i \alpha_i' \left\{ E_i - \sum_j T_{ij} \right\} + \sum_j \alpha_j'' \left\{ A_j - \sum_i T_{ij} \right\} \quad (7.41)$$

qui est résolu en égalisant à 0 la dérivée selon T_{ij} :

$$\frac{\partial L}{\partial T_{ij}} = -\log T_{ij} - \alpha_i' - \alpha_j'' = 0 \quad (7.42)$$

$$\text{soit } T_{ij} = e^{(-\alpha_i' - \alpha_j'')} = e^{-\alpha_i'} * e^{-\alpha_j''} \quad (7.43)$$

La valeur des multiplicateurs de Lagrange est obtenue par simple transformation de variables :

$$\alpha_i' E_i = e^{-\alpha_i'} \quad \text{et} \quad \beta_j A_j = e^{-\alpha_j''} \quad (7.44)$$

$$\text{On obtient ainsi } T_{ij} = \alpha_i \beta_j E_i A_j \quad (7.45)$$

Si l'on connaît une matrice antérieure, la résolution du problème de maximisation avec le même jeu de contraintes (avec l'équation 7.37), on obtient :

$$T_{ij} = \alpha_i \beta_j E_i A_j t_{ij} \quad (7.46)$$

qui correspond au modèle à facteur de croissance (équation 7.9, p. 197) où α_i et β_j sont les facteurs de balancement et E_i et A_j les émissions et attractions issues du modèle de génération.

Pour retrouver le modèle gravitaire, il manque la fonction de résistance. Pour l'obtenir, il faut ajouter une contrainte supplémentaire :

$$\sum_{i,j} T_{ij} * C_{gij} = C_g \quad (7.47)$$

Cette contrainte exprime le fait que la somme des temps ou coûts généralisés de l'ensemble des déplacements présents dans le macro-état du système est une constante. Cette contrainte supplémentaire s'exprime par :

$$C_g - \sum_{i,j} T_{ij} * C_{gij} = 0 \quad (7.48)$$

Le problème de maximisation sous contraintes 7.39, 7.40 et 7.48 s'écrit de manière similaire :

$$L = \log W' + \sum_i \alpha_i' \left\{ E_i - \sum_j T_{ij} \right\} + \sum_j \alpha_j'' \left\{ A_j - \sum_i T_{ij} \right\} + b \left\{ C_g - \sum_{i,j} T_{ij} * C_{gij} \right\} \quad (7.49)$$

qui se résout de même par dérivation :

$$\frac{\partial L}{\partial T_{ij}} = -\log T_{ij} - \alpha_i' - \alpha_j'' - b C_{gij} = 0 \quad (7.50)$$

$$\text{soit } T_{ij} = e^{(-\alpha'_i - \alpha''_j - bCg_{ij})} = e^{-\alpha'_i} * e^{-\alpha''_j} * e^{-bCg_{ij}} \quad (7.51)$$

En reprenant la même transformation pour les coefficients α'_i et α''_j (équation 7.44), on obtient :

$$T_{ij} = \alpha_i \beta_j E_i A_j e^{-bCg_{ij}} \quad (7.52)$$

qui correspond au modèle gravitaire avec une fonction de résistance exponentielle.

En modifiant la contrainte précédente (7.47), par la formulation suivante :

$$\sum_{i,j} T_{ij} * \log(Cg_{ij}) = Cg, \quad (7.53)$$

on retrouve la fonction de résistance puissance :

$$T_{ij} = \alpha_i \beta_j E_i A_j Cg_{ij}^{-b} \quad (7.54)$$

Et en en prenant les deux contraintes (7.47) et (7.53), on retrouve la fonction de résistance combinant les fonctions exponentielles et puissances.

La justification de l'une des trois formes de contraintes (7.47, 7.53 ou la combinaison des deux) d'un point de vue théorique dans l'approche par la maximisation de l'entropie reste évidemment problématique. Elle l'est tout autant pour le modèle gravitaire dérivé de l'analogie avec des lois de la physique. Toutefois, les données empiriques disponibles conduisent le plus souvent à privilégier la contrainte exprimée par l'équation 7.47 ou sa combinaison avec la contrainte exprimée par l'équation 7.53.

L'équation 7.22 (p. 202), exprimant une fonction de résistance en escalier, peut également être retrouvée. Il suffit d'introduire une contrainte spécifique pour chacune des k classes de coûts généralisés. Elle s'obtient en reprenant l'équation 7.47 (on pourrait également reprendre l'équation 7.53) et en l'adaptant à chacune des classes k :

$$Cg^k = \sum_{i,j} T_{ij} Cg_{ij}^{m_k} \delta_{ij}^k \quad (7.55)$$

δ_{ij}^k a la même signification que dans l'équation 7.22 (p. 202) ;

$Cg_{ij}^{m_k}$ est le coût généralisé moyen de la classe k.

Après transformation, on retrouve l'équation 7.22.

Il est intéressant de voir que le même modèle peut être construit à partir d'analogie avec des théories de la physique ou à partir de l'approche par la maximisation de l'entropie qui est utilisée dans plusieurs domaines. Toutefois, nous privilégions la justification s'appuyant sur l'approche par la maximisation de l'utilité individuelle qui est plus cohérente avec l'approche théorique développée dans cet ouvrage.

2.4. Modèles d'opportunité

Le modèle gravitaire est le modèle le plus couramment utilisé. Nous présentons toutefois également le modèle d'opportunité même s'il a donné lieu à peu d'applications (SETRA, CETUR ; Salenius, 1972). D'après Ortuzar et Willumsen (1994), les fondements de la méthodologie sont dus à Stouffer (1940), mais c'est Schneider (1959) qui a développé théoriquement la méthode dans le sens où elle est exposée aujourd'hui.

Le principe du modèle est simple. Il considère que les individus choisissent la destination la plus proche de leur lieu d'origine, leur offrant l'opportunité de satisfaire leur besoin (de déplacement ou d'activité). La distance peut être géométrique, mais elle est plus souvent un coût ou un temps généralisé. La distance n'est donc plus prise en compte de manière absolue ou continue comme dans le modèle gravitaire, mais de manière relative ou ordinale par rapport aux autres destinations possibles.

L'application de la méthode consiste donc à énumérer les destinations possibles (en fait l'ensemble des zones) et à les classer par ordre croissant de distance. La probabilité qu'un individu choisisse une zone est donc égale à la probabilité qu'il ne choisisse pas les zones plus proches multipliée par la probabilité de trouver des opportunités susceptibles de le satisfaire dans la zone.

Soit p la probabilité qu'une opportunité donnée satisfasse l'individu. Si la zone j comporte N_j opportunités susceptibles de satisfaire l'individu, la probabilité qu'au moins une opportunité satisfasse l'individu dans la zone j sera $p * N_j$.

La probabilité qu'un individu choisisse la zone j dépend donc de son ordre dans le classement des zones par ordre croissant de distance à partir de la zone i origine. Si l'on note par k l'indice traduisant cet ordre, la probabilité de choisir une destination dans cette k -ième zone devient :

$$P_k = Q_{k-1} * pN_k \quad (7.56)$$

Où P_k est la probabilité qu'une opportunité de la k -ième zone soit choisie par l'individu ;

Q_{k-1} est la probabilité qu'aucune des opportunités présentes dans les $k-1$ zones plus proches ne satisfasse l'individu.

N_k est le nombre d'opportunités de la k -ième zone dans l'ordre croissant de distance.

$$\text{Donc } Q_k = Q_{k-1} - P_k = Q_{k-1} (1 - pN_k) \quad (7.57)$$

$$\text{D'où, on obtient : } pN_k = \frac{Q_{k-1} - Q_k}{Q_{k-1}} \quad (7.58)$$

On exprime par O_k le cumul des opportunités présentes dans les k zones les plus proches, soit

$$O_k = \sum_k N_k \quad (7.59)$$

Si l'on remplace N_k dans l'équation 7.58 par la valeur déduite de l'équation 7.59, on obtient :

$$-p(O_k - O_{k-1}) = \frac{Q_k - Q_{k-1}}{Q_{k-1}} \quad (7.60)$$

Si l'on considère l'expression limite pour une variation infinitésimale de O , l'équation 7.60 devient :

$$\frac{dQ(O)}{Q(O)} = -pdO \quad (7.61)$$

En intégrant cette équation, on obtient :

$$\log(Q(O)) = -pO + \text{constante} \quad (7.62)$$

$$\text{ou } Q(O) = \alpha e^{-pO} \quad (7.63)$$

La probabilité que l'individu ne trouve pas de destination qui le satisfasse dans les k zones les plus proches de la zone origine i s'exprime donc par :

$$Q_{k,i} = \alpha e^{-pO_{k,i}} \quad (7.64)$$

Si l'on prend en compte le nombre d'individus dans la zone i ou plus exactement le nombre de déplacements ayant une origine dans la zone i, c'est-à-dire E_i l'émission de la zone i, le nombre de déplacements T_{ij} entre i et j (j étant la kième zone la plus proche de i) devient :

$$T_{ij} = E_i (Q_{k-1,i} - Q_{k,i}) = E_i \alpha (e^{-pO_{k-1,i}} - e^{-pO_{k,i}}) \quad (7.65)$$

Le respect de la contrainte aux marges permet de calculer α :

$$E_i = \sum_j T_{ij} = \sum_k E_i \alpha (e^{-pO_{k-1,i}} - e^{-pO_{k,i}}) = \alpha E_i (1 - e^{-pO_N}) \quad (7.66)$$

où N est le nombre de zones et donc O_N le nombre total d'opportunités de l'aire d'étude, d'où

$$T_{ij} = E_i \frac{(e^{-pO_{k-1,i}} - e^{-pO_{k,i}})}{(1 - e^{-pO_N})} \quad (7.67)$$

avec $O_{k,i}$ le cumul des opportunités présentes dans les k zones les plus proches de la zone i,

soit : $O_{k,i} = \sum_k N_k$, N_k étant le nombre d'opportunités susceptibles de satisfaire l'individu dans la zone k.

D'un point de vue pratique, dans une application pour le motif domicile-travail, le nombre d'opportunités peut être estimé à travers le nombre d'emplois. Pour le motif domicile-achats, on pourra utiliser le nombre d'emplois de commerce ou la surface commerciale ou une autre variable disponible pour exprimer la propension à attirer des déplacements pour ce motif. D'une manière générale, c'est donc le même type de variables qui est utilisé pour estimer le nombre d'opportunités que pour estimer l'attraction dans le modèle de génération.

Le modèle utilise donc les mêmes variables explicatives que le modèle gravitaire, mais les hypothèses de choix des individus sont différentes et l'influence de la distance est également différente : continue pour le modèle gravitaire, ordinaire pour le modèle d'opportunité. Cela a évidemment des conséquences sur l'incidence de la distance dans les deux modèles. Une modification du système de transport qui ne remettrait pas en cause l'ordre des distances n'aurait aucune incidence dans le modèle d'opportunité (cette éventualité est toutefois peu probable lorsque le nombre de zones est important) à la différence du modèle gravitaire. Inversement le modèle d'opportunité sera fortement sensible à la précision du calcul des distances si les intervalles de distance sont faibles entre les zones successives dans le classement par ordre croissant de distance. Dans ce cas, une modification marginale du mode de calcul de la fonction de coût ou temps généralisé aura des conséquences beaucoup plus grandes avec le modèle d'opportunité qu'avec le modèle gravitaire.

Ces éléments constituent la principale limite du modèle d'opportunité. Il présente par contre l'avantage de permettre le calcul des déplacements intrazonaux en utilisant la même méthodologie que pour le calcul des déplacements interzonaux.

Les indicateurs de qualité de reproduction de la situation de référence utilisés pour le calage du modèle gravitaire peuvent également être utilisés pour le calage du modèle d'opportunité. Les logiciels du marché ne proposent toutefois pas cette option ce qui explique en partie la confidentialité de cette méthode (à moins que ce ne soit cette confidentialité qui explique son absence des logiciel du marché ... On retrouve un classique problème de causalité !).

3. PRINCIPALES LIMITES DE L'ETAPE DE DISTRIBUTION

Nous reprenons les trois conditions de l'opérationnalité des modèles (chapitre 2, p. 45) pour exposer les principaux éléments de la problématique des modèles de distribution :

- pertinence : prise en compte des déplacements internes (section 3.1), pertinence du coût généralisé comme variable explicative du choix de la destination (section 3.2), stabilité des coefficients de calage (section 3.3), segmentation selon les individus et/ou les motifs (section 3.4), déplacements d'échange réalisés par les résidents du périmètre d'étude (section 3.5) ;
- cohérence interne à l'étape de distribution : cellules nulles (section 3.6) et cohérence interne au modèle à quatre étapes : quelle valeur pour le coût généralisé (section 3.7), le choix du mode succède ou précède le choix de la destination ou les problèmes d'incohérences avec le choix du mode (section 3.8);
- mesurabilité : qualité des données (section 3.9).

Les questions de cohérence externe ne sont pas examinées ici car elles dépendent étroitement des objectifs spécifiques de chaque étude de modélisation, qui sont par nature variables. Les choix de classification proposés ci-dessus sont évidemment critiquables. Ils le sont par la définition même de la problématique des modèles puisque les trois conditions de l'opérationnalité des modèles sont contradictoires (chapitre 2, p. 48). La prise en compte d'un aspect selon l'une des trois dimensions a donc inévitablement des conséquences sur les deux autres. L'entrée que nous avons privilégiée nous semble mieux mettre en valeur les limites des modèles de distribution ou les difficultés rencontrées dans la pratique.

3.1. Déplacements intrazonales

Le choix d'un zonage impose certaines réductions dans la représentation de la demande et des réseaux. L'ensemble de la demande d'une zone est ainsi concentré dans son centroïde. Un déplacement interne est donc un déplacement du centroïde vers lui-même. Sa distance est donc nulle car le déplacement ne nécessite pas un usage du réseau. Il serait toujours possible de lui affecter une distance fonction des caractéristiques de la zone (surface, offre disponible...), mais on doit faire face à d'autres difficultés. Une part importante des déplacements internes est effectuée à pied. Si l'on retient une distance temps généralisé, et que l'on cherche à calculer une sorte de moyenne des temps généralisés des déplacements internes à la zone, le temps intrazone pourrait être supérieur au temps nécessaire pour aller dans les zones adjacentes voire plus lointaines.

Pour solutionner le problème, la distance étant en quelque sorte fictive compte tenu de la représentation choisie, on pourrait retenir une valeur spécifique à chaque zone de telle sorte que le calage permette une bonne reconstitution des déplacements intrazonaux. Il n'est toutefois pas certain que cette valeur conserve sa pertinence en prévision dans la mesure où elle n'est pas explicative des comportements de mobilité de proximité et surtout d'usage de la marche.

Autrefois, la marche n'était en fait pas réellement considérée dans les exercices de modélisation. Le problème était donc de faible ampleur. L'évolution des problématiques de planification des déplacements urbains pose aujourd'hui les enjeux de manière un peu différente.

Des solutions existent. Mais elles passent en général par un traitement spécifique des déplacements intrazonaux. Ce traitement est effectué en dehors du modèle de distribution choisi et les déplacements sont ensuite réintégrés dans la matrice origines-destinations.

3.2. Pertinence de la distance comme critère de choix de la destination

Dans l'introduction de ce chapitre, nous présentons la distribution comme l'étape du choix de la zone de destination. Il est donc légitime de se poser la question de la pertinence des variables prises en compte pour expliquer ce choix.

Dans le modèle à facteur de croissance ce choix n'est pas explicité. On considère seulement que la matrice de référence traduit ce choix. Donc en conservant la structure de la matrice sous contrainte du respect des marges, on conserve les mêmes choix.

Dans le modèle gravitaire, le choix est exprimé à travers la distance (généralement une distance en termes de coût ou de temps généralisé). L'individu préfère choisir des destinations proches et sa satisfaction décroît selon une courbe qui est en fait la courbe exprimée par la fonction de résistance.

Dans le modèle d'opportunité, la distance est toujours prise en compte, mais plus de manière absolue ou continue, mais de manière relative ou ordinale. Une autre dimension est explicitement prise en compte. C'est le nombre d'opportunités disponibles dans une zone pour réaliser une activité. En fait, cette dimension est également présente implicitement dans les deux autres modèles car elle correspond aux attractions.

Les mêmes variables sont donc utilisées par les trois familles de modèles sauf pour la distance qui n'est pas considérée par le modèle à facteurs de croissance. Toutefois, indépendamment de la distance, le choix entre plusieurs localisations possibles n'est pas complètement aléatoire. Dans le modèle d'opportunité, on considère explicitement que toutes les opportunités sont équiprobables de probabilité p (si l'on ne tient pas compte de la distance).

Dans les deux autres modèles, il n'y a pas d'hypothèse explicite. Dans le modèle à facteurs de croissance, on retient les choix de la situation de référence si tant est que les effectifs enquêtés puissent donner une image suffisamment fidèle de ces choix.

Dans le modèle gravitaire, une première présentation s'approche du choix du modèle à facteurs de croissance, puisque le calage conduit à essayer de reproduire la situation de référence, mais seul un ou deux coefficient(s) de la fonction de distance peut traduire ce choix (parallèlement à la prise en compte de la distance). Mais une autre présentation pourrait privilégier l'approche du modèle d'opportunité, puisque l'attraction d'une zone ne tient compte que de la quantité d'opportunités possibles comme dans le modèle d'opportunité. Si dans une agglomération les émissions et attractions sont constantes, malgré une évolution de la nature des emplois offerts par exemple, l'évolution de la matrice origines-destinations ne dépendra que de l'évolution des temps généralisés. Enfin et surtout, la justification du modèle gravitaire à partir de la théorie de la maximisation de l'utilité explicite l'hypothèse. Dans notre démonstration, nous avons défini l'utilité par une composante déterministe constante quel que soit l'individu ou la destination retenue (éventuellement après segmentation selon le motif ou tout autre segmentation qui serait utilisée lors de l'étape de distribution) et une composante aléatoire. Cela signifie que nous avons postulé qu'en moyenne toutes les destinations potentielles ont une utilité identique.

C'est probablement une des raisons pour lesquelles le modèle de distribution donne fréquemment des résultats peu crédibles face à des modifications importantes du système de transport, comme par exemple des scénarios de péage urbain. Il est en effet probable que les choix ne dépendent pas que de la distance et de la quantité d'opportunités, mais aussi et peut être surtout de la

qualification de ces opportunités, c'est-à-dire que pour un individu donné, les opportunités ne sont pas équiprobables ou n'ont pas la même utilité pour la composante déterministe (indépendamment de la distance).

3.3. Stabilité des coefficients

La stabilité dans le temps des coefficients de calage est un problème récurrent qui est présent pour chacune des étapes des modèles à quatre étapes. Le calage étant souvent effectué sur des données correspondant à une seule date, l'influence du temps ne peut être prise en compte.

En supposant que le coefficient de la fonction puissance ou exponentielle (ou la combinaison des deux) est stable dans le temps, on suppose en fait que la distribution des déplacements selon les coûts généralisés est stable dans le temps. Cette hypothèse est loin d'être établie. L'allongement des distances de déplacements automobiles est un fait connu de l'évolution des caractéristiques de mobilité de ces dernières décennies. Dans le même temps, on observe une augmentation des vitesses moyennes automobiles. Il n'est toutefois pas certain que ces deux évolutions contraires se compensent pour donner une distribution stable des coûts généralisés. L'utilisation d'un modèle calé pour les k classes de la distribution des coûts généralisés (équation 7.22, p. 202) conduit également à la même hypothèse.

Pour résoudre le problème, il est nécessaire de disposer de données en série chronologiques ou tout au moins de mesures répétées dans le temps. Ainsi sur le cas de l'agglomération lyonnaise, le calage d'un modèle gravitaire sur les données des trois dernières enquêtes ménages déplacements a mis en évidence une évolution des coefficients de calage de la fonction de résistance exponentielle (Lichère, Raux, 1997b). Cette évolution a pu être reliée très fidèlement à la motorisation qui évolue elle aussi dans le temps. Cette relation est « *interprétée en considérant que la disposition d'une voiture à la fois ouvre le champ des destinations possibles et donne accès à un mode qui est souvent le plus rapide, ce qui pousse à choisir des destinations « plus lointaines » en termes de temps tous modes* ». Nous n'avons toutefois que déplacé le problème car nous devons faire une nouvelle hypothèse de stabilité non plus du coefficient mais de la relation établissant son évolution dans le temps. Cette hypothèse est probablement de portée plus réduite sur les résultats du modèle, mais on est toujours face à une hypothèse d'invariants (chapitre 2, p. 42).

3.4. Segmentation

L'expérience montre que la courbe de la fonction de résistance du modèle gravitaire est très différente selon le motif du déplacement. Prenons un exemple trivial pour s'en convaincre. Les déplacements pour le motif domicile-enseignement primaire sont pour la plupart internes ou bien limités aux zones les plus proches de la zone du domicile. Inversement, les contraintes des marchés du logement et du travail et les aspirations individuelles ne permettent pas cette concordance des lieux d'emplois et de domicile. Il est donc généralement souhaitable de segmenter la matrice origines-destinations observée selon différents motifs de déplacements.

Si un modèle catégoriel a été choisi pour le modèle de génération, il est nécessaire de se poser la question du maintien de la segmentation pour le modèle de distribution. Cette question sera d'autant plus importante selon que les variables de segmentation sont pertinentes ou non, également pour le choix du mode de déplacement.

Toutefois, cette segmentation, surtout si elle nécessite un croisement des variables socio-économiques et des motifs de déplacements, aiguë le problème de la fiabilité des données en référence. Le choix de la segmentation dépend donc de la possibilité de segmenter les données en référence et donc de la qualité des données disponibles en référence. Le problème est encore accru par le fait que pour une partie des données de calage, les variables de segmentation ne sont pas toujours

disponibles. C'est fréquemment le cas des enquêtes embarquées dans les transports collectifs ou des enquêtes au bord de la route. Le temps disponible pour ce type d'enquête étant très limité, le contenu du questionnaire est limité au maximum. Enfin, pour les données de comptage, aucune segmentation n'est possible.

3.5. Zones externes

Le problème des déplacements d'échanges ou de transit non pris en compte par les enquêtes ménages déplacements a déjà été évoqué (chapitre 4, p. 119). Mais il y a une autre catégorie de déplacements d'échange généralement pris en compte par les enquêtes ménages déplacements, qui pose problème dans le calage du modèle. Il s'agit des déplacements réalisés par les individus enquêtés qui sortent du périmètre d'étude. Ces déplacements sont pris en compte par les individus dans leur processus de choix de destination. Ils correspondent à un choix pour une destination lointaine (tout au moins externe à la zone d'étude). Il devrait donc être pris en compte dans le calage de la fonction de résistance du modèle gravitaire ou dans le modèle d'opportunité. Ceci d'autant plus que la proportion de ces déplacements a tendance à s'accroître dans le temps.

La connaissance de ces déplacements est toutefois très limitée. Bien souvent, la destination n'est connue que de manière très approximative ne permettant pas de calculer un temps généralisé avec une précision suffisante. De plus, la définition des réseaux est bien souvent limitée au périmètre d'étude. La connaissance fine de la destination ne permet donc pas toujours un calcul fin du temps généralisé. Enfin, l'attraction des zones externes n'est généralement pas calculée faute de données sur ces zones externes.

Ces déplacements ne sont donc généralement pas pris en compte dans le calage du modèle de distribution. Ils sont estimés par des procédures externes en les englobant avec les autres déplacements d'échange émanant de personnes non résidentes de la zone d'étude.

3.6. Cellules nulles

Ce problème a déjà été largement évoqué dans les sections précédentes, nous ne faisons donc que rappeler ici qu'il s'agit d'une des limites des modèles de distribution.

3.7. Quelle valeur pour le coût généralisé ?

La valeur du coût généralisé est calculée à partir de la fonction de coût généralisé et de la description des réseaux. Le problème du choix de la valeur pour le modèle de distribution est triple.

La distribution est effectuée avant la répartition modale. Quelle valeur de coût généralisé faut-il retenir : celle de la voiture, celle des transports collectifs, de la marche, des deux-roues ... voire d'une subdivision des modes collectifs. Ces valeurs sont en général différentes. Le problème n'ayant pas une bonne solution, deux solutions sont privilégiées :

- valeur du mode présentant le plus faible coût généralisé. Seuls la voiture et les transports collectifs sont pris en compte ;
- moyenne pondérée des coûts généralisés des modes voiture et transports collectifs. La pondération est en général le poids de chacun des modes sur l'origine-destination considérée. Cette solution semble apparemment plus pertinente que la première, car elle tient compte du choix effectif du mode réalisé par les individus. Elle pose toutefois deux problèmes principaux. En prévision tout d'abord, la répartition n'est connue qu'à l'issue de l'étape de répartition modale qui prend place après la distribution (le problème est similaire pour la solution précédente)... Deuxièmement, cette solution est inconsistante avec l'hypothèse de rationalité des choix des individus. Considérons une situation dans laquelle nous avons x modes avec une certaine répartition entre ces modes. Considérons l'ajout d'un mode

supplémentaire dont le coût généralisé est supérieur à tous les autres sur toutes les origines-destinations. L'expérience montre qu'il existe toujours une proportion, même faible, d'utilisateurs utilisant ce mode. Le nouveau coût généralisé moyen obtenu avec ce mode supplémentaire est donc supérieur à celui calculé dans la situation antérieure. Le résultat est donc inconsistant puisque l'introduction d'une option supplémentaire quelle qu'elle soit sans modification des options antérieures ne peut conduire qu'à une diminution du coût généralisé moyen selon le principe de rationalité du choix du consommateur. La contradiction n'est toutefois qu'apparente. Elle découle de l'approche agrégée, car si le coût généralisé de cette nouvelle option est en moyenne supérieur à celui des autres modes, cela n'est plus vrai à niveau désagrégé. Il y a des individus qui perçoivent ce nouveau mode comme étant moins coûteux (en coût généralisé) que les autres, ce qui explique justement leur choix pour ce mode selon l'hypothèse de rationalité du choix du consommateur.

Nous verrons dans le chapitre suivant qu'il existe en fait une troisième solution applicable dans certains cas (équation 8.29, p. 237).

La contradiction peut donc être levée à un niveau désagrégé, mais pas au niveau agrégé. C'est une des raisons qui amène à privilégier le choix du mode présentant le coût généralisé le plus faible. En général, ce choix permet de plus un meilleur calage et une plus grande stabilité du modèle en prévision.

Lorsque les transports collectifs sont subdivisés en plusieurs modes, on retient en général la valeur du mode le plus rapide (en termes de coûts généralisés).

Le second problème résulte de la congestion. Celle-ci découle de la charge sur les infrastructures. Mais la charge sur les infrastructures n'est connue qu'à l'issue de l'étape d'affectation qui prend place après la répartition modale ...

Enfin, pour disposer d'un jeu de données suffisant, le calage de la distribution est généralement effectué sur des données journalières et non pas d'heure de pointe. Il est pourtant évident que le coût généralisé n'est pas identique pendant les heures creuses et pendant les heures de pointe. Autrefois, les modèles servaient principalement au dimensionnement des infrastructures, ce qui conduisait à retenir les coûts généralisés d'heure de pointe. Aujourd'hui la problématique des déplacements urbains ne se limite plus à cette question, rendant le choix encore plus délicat.

Le problème n'est apparemment pas soluble. Dans la pratique, on retient en général une première valeur résultant de l'affectation des matrices de référence voiture particulière et transports collectifs et de la valeur correspondante de répartition modale si nécessaire pour le calcul du modèle de distribution. A l'issue du modèle de distribution, on fait tourner les modèles de choix du mode et d'affectation qui fournissent une nouvelle valeur pour les coûts généralisés. Il est alors possible de recommencer à faire tourner les modèles de distribution, répartition modale et affectation pour s'assurer de la cohérence des valeurs de coûts généralisés et au besoin renouveler l'itération.

3.8. Ordre des étapes et cohérence avec le choix du mode

L'étape de choix du mode prend place après celle du choix de la destination. Il n'est toutefois pas certain que l'individu dans son processus de choix suive toujours cette séquence.

De plus, la séquence distribution – répartition modale conduit dans certains cas à des résultats aberrants, notamment lorsque les caractéristiques du système de transport évoluent fortement. Considérons un scénario incluant un péage de cordon pour les automobilistes. Ce péage conduit à une croissance du coût généralisé sur les origines-destinations concernées. Cette croissance induit donc une diminution du trafic sur les origines-destinations concernées. Vient ensuite la répartition modale

qui sera plus favorable aux transports collectifs compte tenu de la croissance du coût généralisé voiture. Mais il arrive dans certains cas, que la baisse du trafic sur certaines origines-destinations est telle qu'elle conduit à une baisse du volume de trafic sur les transports collectifs, même si leur part de marché s'est accrue. L'introduction du péage pour les automobiles conduit alors à une baisse de l'usage des transports collectifs ... ce qui est pour le moins surprenant !

Certains analystes préconisent ainsi dans certains cas d'inverser les deux étapes. Toutefois, le problème de cohérence évoquée ci-dessus n'est pas résolu. Le principal déterminant du choix du mode est le coût généralisé. Comment celui-ci peut-il être calculé si la répartition spatiale n'a pas été faite ?

Il est en fait plus probable que l'individu fasse un choix combiné destination et mode de transport. Ceci explique le développement des approches combinant le choix de la destination et le choix du mode (CERTU, 1998b ; chapitre 8, p. 236). Morellet propose une autre solution à partir d'une segmentation extrêmement fine de la demande prenant en compte les caractéristiques des individus, mais aussi des déplacements. Le modèle simule l'évolution du poids de chacun des segments ce qui permet de traiter les problèmes de cohérence entre distribution et choix du mode (modèle MATISSE : Morellet, Julien, 1990 ; Morellet, Marchal, 1995).

3.9. Qualité des données

Le problème a déjà été évoqué dans ce chapitre, nous nous contentons donc de le rappeler brièvement. La taille des échantillons est en général trop faible pour obtenir des matrices origines-destinations suffisamment fiables y compris pour des découpages assez grossier. A cette limite s'ajoute le problème spécifique des cellules nulles. Il s'agit d'un problème récurrent commun à toutes les étapes de la modélisation.

Chapitre 8 – Répartition modale – approche agrégée

La répartition modale constitue généralement la troisième étape des modèles à quatre étapes. Elle prend place après la distribution dont elle utilise les sorties : les matrices origines-destinations. La répartition modale correspond au choix du mode de déplacement dans la séquence des choix de l'individu. Mathématiquement, elle permet d'éclater la(les) matrice(s) origines-destinations en autant de matrices origines-destinations par mode qu'il y a de modes pris en compte.

Nous utilisons un plan similaire à celui retenu pour les étapes précédentes. Nous commençons par préciser certains éléments relatifs à la répartition modale (section 1), pour présenter ensuite les différentes familles de modèles disponibles : modèles normatifs, modèles logit ou régression logistique et enfin les modèles combinant certaines des étapes du modèle à 4 étapes (section 2). La dernière section est consacrée à l'analyse des limites de l'étape de répartition modale (section 3).

Notre chapitre est limité comme les précédents aux approches agrégées. Le choix du mode dans une approche désagrégée sera abordé dans le chapitre suivant.

1. DEFINITION ET VARIABLES EXPLICATIVES

L'étape du choix du mode a longtemps été négligée en France à la différence des pays anglo-saxons (section 1.1). Ce constat est d'autant plus surprenant que les travaux ont été nombreux pour expliquer les facteurs explicatifs du choix du mode (section 1.2), à moins que ce ne soit justement ce foisonnement qui ait conduit à remettre également en cause les développements apparus outre-Atlantique.

Nous examinons ensuite comme pour les étapes précédentes, la question du calcul des coûts généralisés (section 1.3), puis celle de la segmentation de la demande (section 1.4).

1.1. Choix du mode : une étape négligée en France

Les pratiques de modélisation sont très différentes en France et dans la plupart des pays anglo-saxons, notamment pour des raisons historiques propres à la France que nous avons eu l'occasion d'aborder au premier chapitre. Parmi les 4 étapes, la répartition modale est celle pour laquelle ces différences sont le plus évidentes. Jusqu'à une période récente, la plupart des études de modélisation ont été unimodales (à l'exception de la région parisienne qui comme dans beaucoup de domaines se distingue) s'appuyant notamment sur les deux logiciels qui étaient les plus utilisés en France le logiciel DAVIS pour les études routières et le logiciel TERESE pour les études de transports collectifs (GESMAD, 2000). Ces deux outils ont assez bien répondu aux problématiques qui étaient posées d'estimation de clientèle. Cette estimation permettait d'une part de dimensionner l'infrastructure et d'autre part de procéder à l'évaluation socio-économique en utilisant une approche monomodale.

L'évolution des problématiques marquées par les nouveaux Plans de Déplacements Urbains et une volonté de plus grande maîtrise de la voiture au profit des autres modes a imposé une évolution des besoins de modélisation. La réponse à ces questions ne peut plus faire l'économie d'une réflexion sur le partage modal. Elle ne peut plus se faire uniquement à l'aide des modèles monomodaux. C'est une des raisons du regain d'intérêt pour l'étape de répartition modale en France.

L'évolution a été très différente dans les pays anglo-saxons et tout particulièrement aux Etats-Unis. Les critiques sur la faible pertinence des modèles agrégés à représenter les processus de choix

des individus (ces critiques ont été tout à fait similaires des deux côtés de l'Atlantique, mais elles n'ont pas conduit aux mêmes évolutions) et donc à permettre de simuler efficacement l'impact de scénarios contrastés de politiques de transport ont conduit à des travaux de recherche permettant de mieux fonder sur le plan théorique la modélisation de la demande de transport. Ils ont débouché sur la théorie des choix discrets que nous développons dans le prochain chapitre.

Le domaine d'application privilégiée des modèles issus de cette théorie a justement été la répartition modale. C'est donc cette phase qui la première (et aujourd'hui encore) a bénéficié de ces efforts de recherche aux Etats Unis. C'est notamment pour sa contribution théorique à l'analyse des modèles de choix discrets que McFadden a reçu le prix Nobel d'économie en 2000 (ses travaux ne sont toutefois pas limités au secteur des transports).

Le contraste est donc saisissant entre un foisonnement de travaux sur la phase de répartition modale dans les pays anglo-saxons et une absence quasi-totale de travaux de modélisation sur cette phase en France. Gageons que l'évolution des problématiques de transport en France permettra un renouvellement des pratiques.

1.2. Variables explicatives du choix du mode

Il est classique d'organiser les facteurs explicatifs du choix du mode en trois groupes relatifs à :

- l'offre de transport ;
- les caractéristiques du déplacement ;
- les caractéristiques de l'individu ou du ménage.

1.2.1. Déterminants liés à l'offre de transport

Les caractéristiques de l'offre de transport influent sur le mode de transport choisi. Ces caractéristiques sont assez nombreuses. Les deux principaux paramètres pris en compte sont le prix et le temps. Le stationnement apparaît aussi comme un facteur déterminant du choix. La faiblesse des données disponibles conduit souvent à le négliger dans les études de modélisation. Enfin d'autres éléments plus qualitatifs, souvent assez délicats à mesurer, influent sur le choix.

Le prix considéré est un prix ressenti par l'individu. Il diffère généralement du coût réel. C'est notamment le cas de la voiture pour laquelle les usagers ne considèrent généralement que les coûts d'usage directs (essence et péage éventuel) dans leur processus de choix.

Le temps de déplacement est également un temps perçu afin de tenir compte de la perception plus ou moins négative associée aux différents moments du déplacement. Par exemple, pour les transports collectifs, les temps de marche ou d'attente sont perçus comme étant au moins deux fois plus pénibles que les temps en véhicule. Selon les objectifs de l'analyse, il sera ainsi possible de décomposer assez finement le temps de déplacement en temps élémentaires perçus par les individus comme homogènes et affectés chacun de facteurs pondérateurs afin de déterminer un temps « pénibilisé » (Lichère, Raux, 1997a, b). C'est ce temps qui est pris en compte dans le processus de choix. La décomposition des temps élémentaires permet d'isoler le temps d'attente qui découle de la fréquence d'une part et de la régularité d'autre part. La SEMALY définit ainsi un coefficient de régularité dans son logiciel TERESE (SEMALY, 2000) qui permet de calculer l'attente en multipliant ce coefficient par le demi intervalle entre deux véhicules. Une amélioration de la régularité (couloirs réservés, priorité aux feux...) peut ainsi être prise en compte dans la modélisation.

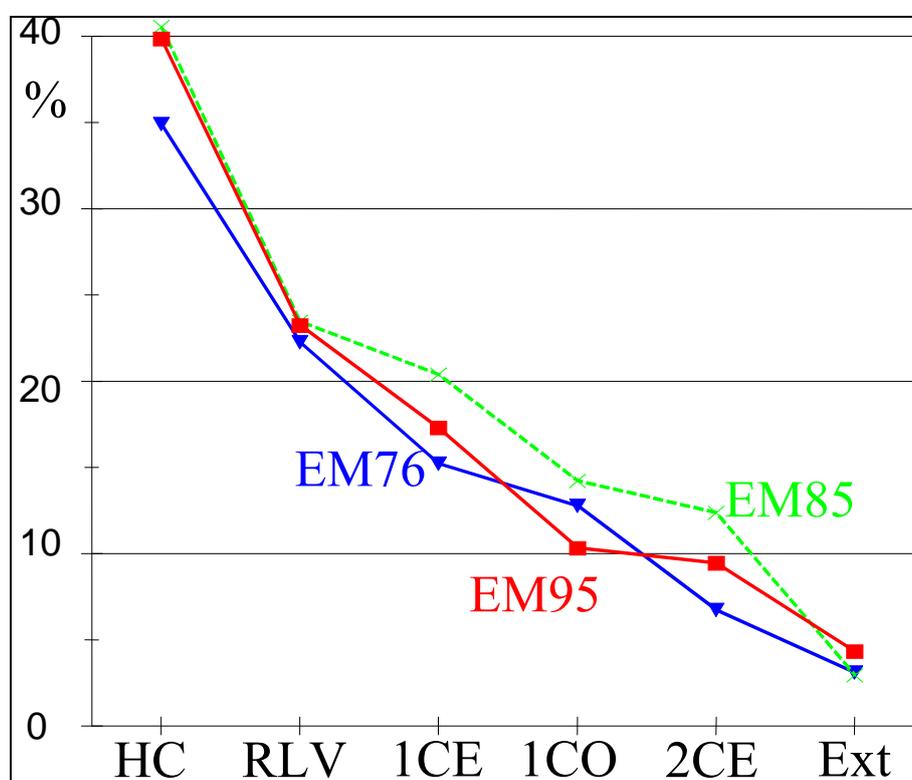
L'offre de stationnement est également un déterminant important du choix du mode (Massot, Orfeuill, 1991 ; Kaufmann, Guidez, 1996)). Massot et Orfeuill ont ainsi montré que les individus disposant d'une place de stationnement sur leur lieu de travail utilisaient deux fois moins les transports

collectifs pour se rendre au travail que ceux qui n'en disposaient pas. La différence est même plus forte lorsque les individus résident dans le centre d'une grande agglomération disposant d'une offre de transports collectifs de qualité. Ce constat explique probablement une partie du succès rencontré par des villes comme Berne et Zurich où l'usage des transports collectifs est nettement plus fort qu'en France (Bonnel, 1993).

D'autres facteurs plus qualitatifs influent également. C'est le cas du confort, de la sécurité, de la fiabilité... autant de facteurs souvent assez difficiles à mesurer. Le plus souvent, ils sont pris en compte de manière implicite dans la formule de coût généralisé. L'estimation de ces facteurs est toutefois possible à l'aide d'enquêtes développées en marketing ou de méthodes de préférences déclarées permettant d'estimer l'utilité associée (Ampt, Swanson, 1996).

1.2.2. Déterminants liés aux caractéristiques du déplacement

Chacune des caractéristiques du déplacement influe fortement sur l'usage des modes de transport. Nous nous limitons aux principales : origine-destination, motif, heure de la journée, chaînage.



Graphique 8.1 : Part de marché des TC sur (TC+VP) pour les déplacements ayant une origine ou une destination dans la zone..., périmètre 76

(Sources : Bonnel, 2000b à partir des enquêtes ménages déplacements Lyon 1976, 1985, 1995 ; définition du zonage : p. 76)

L'origine et la destination des déplacements sont évidemment des facteurs importants du choix, car le niveau d'offre des deux modes motorisés varie fortement et souvent de manière opposée selon la localisation des extrémités du déplacement. La part de marché des transports collectifs décroît ainsi à mesure que l'une des extrémités du déplacement se périphérise (graphique 8.1). Si l'on prend en compte les deux extrémités, il apparaît que les flux les plus centraux ont la plus forte part de marché des transports collectifs. Celle-ci reste assez forte pour les déplacements radiaux. Elle devient par

contre très faible pour les flux périphériques. En évolution, compte tenu de l'amélioration de l'offre de transports collectifs à Lyon entre 1976 et 1995, la part de marché des transports collectifs augmente pour chacune des zones (sauf 1CO : 1^{ère} couronne Ouest, graphique 8.1). En revanche, en moyenne, la part de marché diminue très faiblement (de 21,5 % en 1976 à 21,4 % en 1995). On observe ainsi un paradoxe avec une part de marché qui augmente pour l'ensemble des zones, mais qui diminue en moyenne. Le paradoxe n'est qu'apparent, car dans le même temps, la répartition des flux dans l'agglomération a évolué sous l'effet de l'étalement urbain au profit des flux périphériques (Bonnel, 2000b). Cet étalement a de fortes chances de se poursuivre sous l'effet des aspirations fortes de la population à la maison individuelle et des stratégies de localisation des activités (Orfeuill, 2000), mais aussi du fait de l'allongement des distances parcourues permises par l'accroissement de la vitesse sur les voiries routières (Wiel, 1999).

Nous insistons sur cette dimension car elle justifie d'une certaine manière la séquence des étapes plaçant le choix du mode après la distribution. Il est ainsi possible de prendre en compte l'origine et la destination des déplacements à travers la matrice origines-destinations dans le processus de choix du mode. L'ordre inverse ne permettrait pas cette prise en compte.

Les autres caractéristiques du déplacement influent également sur le choix du mode comme nous l'illustrons dans les sections 1.3 à 1.5.

1.2.3. Déterminants liés aux caractéristiques individuelles

Analyser les facteurs déterminants le choix du mode présuppose que tous les individus disposent effectivement du choix entre les modes de transport. Ce n'est évidemment pas le cas. Malgré la progression de la motorisation, tous les individus n'ont pas accès à la voiture en tant que conducteur (Massot, Orfeuill, 1989). La première condition est évidemment d'avoir le permis de conduire. La seconde est d'avoir un véhicule à sa disposition. La motorisation est ainsi généralement prise en compte dans un modèle de choix du mode de transport. Elle peut intervenir à travers une motorisation moyenne, à partir d'une segmentation de la population selon le nombre de véhicules du ménage ou encore à partir d'une estimation de la disponibilité réelle d'un véhicule pour un déplacement donné (il faut alors définir des règles d'affectation du ou des véhicules du ménage entre ses membres).

On peut caractériser l'incidence de la motorisation à partir de données agrégées sur Lyon en 1995. Les individus qui ne disposent pas de voiture dans leur ménage effectuent 4 fois plus de déplacements en transports collectifs qu'en voiture, alors que ceux qui disposent de deux voitures et plus en font 10 fois moins (source : enquête ménages déplacements 1995). Il ne faudrait toutefois pas interpréter la relation de manière trop mécanique, car de nombreux autres facteurs sont corrélés à la motorisation et peuvent expliquer ces résultats. De plus, la relation peut être assez différente lorsque l'offre de transport est également différente. Les écarts sont ainsi beaucoup moins marqués à Berne ou Zurich malgré une motorisation similaire à celle de Lyon (Bonnel, 1993).

Le statut de l'individu est une variable issue des recherches sur l'analyse de la mobilité (Averous et al., 1979-80 ; Jones et al., 1980 ; Andan et al., 1988...). Elle permet de combiner plusieurs facteurs socio-économiques influant sur les pratiques de mobilité : le sexe, l'âge, la profession, la position dans le ménage, la présence d'enfant... Sa définition et notamment son degré de finesse, varie selon les objectifs des analyses. L'intérêt de cette variable est d'être fortement reliée aux rôles des individus dans le ménage, notamment dans la répartition des tâches au sein du ménage, et dans la nature des activités contraintes pratiquées par chaque classe. L'usage des modes est également fortement contrasté selon le statut. De plus, la motorisation est assez différente selon les classes de statut avec des taux de croissance pouvant être modélisés à l'aide d'approches démographiques (Madre, Lambert, 1989, Pochet, 1995). Cette variable est également intéressante, car le poids de chacune des classes de statut est assez facilement prévisible dans le temps.

Le poids des habitudes n'est généralement pas présenté comme un déterminant socio-économique de l'individu. Il s'agit pourtant bien d'une caractéristique de l'individu fortement liée à sa position dans le cycle de vie. Même si la théorie économique sur laquelle s'appuie la plupart des modèles de prévision stipule que l'individu est rationnel et donc qu'il remet en cause en permanence ces choix en fonction de l'évolution de la performance des modes de transport, force est de constater que les moments de choix sont plutôt rares. Ou plus exactement que l'individu remet rarement en cause ces choix. Bourgin (1978) a ainsi mis en évidence l'importance des moments de transition dans le cycle de vie dans la remise en cause des habitudes modales. Goodwin (1984) a également montré que les effets d'une évolution de l'offre peuvent se faire sentir très longtemps après la mise en service d'une nouvelle infrastructure.

Cette brève revue de la littérature sur les facteurs explicatifs du choix du mode ne saurait être complète, car de nombreuses autres variables interviennent dans le processus de choix du mode de transport. Les dimensions sociologiques ou psychologiques ont ainsi été assez peu prises en compte dans cette analyse (Averous et al, 1979-80 ; CEMT, 1977 ; Fichelet, 1976 ; Mercadal, 1970...) qui a privilégié les variables les plus fréquemment prises en compte dans la modélisation.

Dans la pratique, seule une partie de ces facteurs est fréquemment prise en compte dans la modélisation du choix du mode. Les facteurs liés à l'offre sont le plus souvent intégrés dans le calcul du coût généralisé. Ceux liés aux caractéristiques du déplacement interviennent par le biais de la matrice origines-destinations pour ce qui est de la localisation des extrémités du déplacement. Les motifs interviennent fréquemment à travers une segmentation de la matrice (section 1.4, p. 224). Enfin, les caractéristiques individuelles sont prises en compte à travers une segmentation de la demande ou à travers la motorisation. L'impact de cette variable en prévision nécessite une bonne estimation de la motorisation à l'horizon d'étude. Plusieurs modèles sont disponibles dans ce but. Toutefois, les approches démographiques sont généralement celles qui produisent les résultats les plus fiables en prévision (Madre, Lambert, 1989 ; Madre, Armoogum, 1996 ; Pochet, 1995).

1.3. Coût généralisé de la période de pointe ou de la journée ?

Le coût généralisé apparaît donc comme central dans l'estimation du partage modal, tout comme pour l'étape de distribution. Nous verrons qu'il l'est encore pour celle d'affectation. Son estimation est donc cruciale pour la modélisation. Nous ne revenons pas sur le mode de calcul (chapitre 5, p. 152), ni sur les problèmes de cohérence entre étapes (chapitre 7, p. 216). En revanche, nous identifions une nouvelle difficulté d'estimation qui n'est pas sans conséquence pour le partage modal.

En période de pointe de semaine, les performances des transports collectifs sont généralement meilleures que le reste de la journée du fait d'une meilleure fréquence. En revanche, pour l'automobile la période de pointe est synonyme de congestion, généralement absente le reste de la journée. Les valeurs respectives de coût généralisé sont donc très différentes : plus élevées en pointe que le reste de la journée pour la voiture et plus faibles pour les transports collectifs.

Nous avons évoqué l'évolution des objectifs de modélisation (chapitre 6, p. 168) qui nécessite de ne plus se limiter à la pointe qui servait à dimensionner les infrastructures, pour produire des données à la journée. Mais dans ce cas, quelle valeur de coût généralisé doit-on choisir pour chacun des modes ? La solution la plus pertinente serait de découper la demande en période de pointe et période creuse et de modéliser le choix modal pour chacun des segments à partir des valeurs de coûts généralisés de chacune des deux périodes. Cette solution présente un inconvénient majeur. Elle impose de définir et de caler deux réseaux : un pour la pointe, un pour le reste de la journée. Elle est donc généralement rédhitoire en termes de budget (temps ou monétaire).

La solution généralement retenue consiste à segmenter la demande pour isoler la matrice heure de pointe et à ne modéliser le choix du mode que pour cette période. Si des données doivent être produites à la journée ou sur des périodes temporelles plus longues, elles sont obtenues à partir des résultats d'heures de pointe auxquels sont affectés des ratios de passage de la pointe à la période temporelle visée.

1.4. Quelle segmentation ?

L'usage des modes est fortement différencié selon le motif du déplacement. Si l'on se limite aux modes motorisés, la part de marché des transports collectifs est beaucoup plus importante pour le motif école (à Lyon, 55 % en 1995). Contrairement, à ce qui est souvent dit le motif suivant ayant la plus forte part de marché n'est pas le travail (14 % à Lyon en 1995), mais les loisirs (21 %) et les achats/services (16 %). Enfin, pour les accompagnements, la part des transports collectifs est infime (4 %). En général, ces différences justifient une segmentation selon le motif.

L'analyse des variables socio-économiques montre aussi des différenciations d'usage des modes pouvant être assez marquées. Nous nous attardons quelque peu sur la motorisation en raison du rôle généralement tenu par cette variable dans la modélisation. Les données du tableau 8.1 sur Lyon confirment une différence très marquée entre les ménages non motorisés et les autres. Les différences subsistent au sein des ménages motorisés entre les mono-motorisés et les multi-motorisés, mais elles sont faibles comparativement à la précédente segmentation.

Nombre de voitures à libre disposition au sein du ménage	0 voiture	1 voiture	2 voitures et plus
Part voiture particulière	20 %	79 %	90 %
Part transports collectifs	80 %	21 %	10 %

Tableau 8.1 : Répartition modale selon le nombre de voitures à disposition, Lyon 1995
(Source : enquête ménages déplacements de Lyon 1995)

Ces résultats inciteraient à recommander une segmentation selon la motorisation. La segmentation n'est toutefois envisageable que si les matrices origines-destinations issues de l'étape de distribution sont elles-mêmes segmentées selon cette variable. Sinon, il faudrait disposer de règles d'éclatement de la matrice qui seraient probablement assez délicates à définir. Il faut donc aussi une segmentation au niveau de l'étape de génération pour les mêmes raisons.

Les données du tableau 8.2 semblent justifier cette segmentation pour l'étape de génération, la mobilité (au sens du nombre de déplacements quotidiens) étant plus faible pour les ménages non motorisés que pour les autres. De nouveau, une faible différence subsistent entre les mono-motorisés et les multi-motorisés. Ces résultats militent pour une segmentation de la population au moins en deux classes comprenant d'un côté les personnes appartenant aux ménages non motorisés et de l'autre les autres individus (la validation de la segmentation nécessite toutefois au préalable une analyse de variance afin de s'assurer que les différences observées sont statistiquement significatives). L'examen des effectifs (tableau 8.2) conduit à nuancer le propos. La classe des individus appartenant aux ménages non motorisés a un faible effectif. Si cette segmentation doit être croisée avec le motif, il est évident que les effectifs seront trop faibles au moins pour certains motifs pour permettre un calage avec une précision statistique suffisante.

Nombre de voitures à libre disposition au sein du ménage	Nombre moyen de déplacements par jour	Nombre de personnes enquêtées dans la classe	Nombre de déplacements enquêtés dans la classe
0	3,12	1 955	5 864
1	3,94	5 808	22 208
2 et +	4,13	6 234	25 141

Tableau 8.2 : Nombre de déplacements quotidiens et population enquêtée selon le nombre de voitures à disposition, Lyon 1995
(Source : enquête ménages déplacements de Lyon 1995)

Ces résultats nous conduisent à nuancer la pertinence de la segmentation. En effet le gain de pertinence lié à une meilleure discrimination des comportements conduit à une cohérence réduite et à des problèmes de mesurabilité (on retrouve toujours la problématique des modèles !). Le choix résulte donc chaque fois d'une analyse des données disponibles conduisant à un compromis entre le gain de pertinence et la précision statistique du calage. Toutefois, si la motorisation n'est pas retenue comme variable de segmentation, elle reste toutefois une variable clef du modèle de répartition modale. Elle intervient alors le plus souvent à travers un taux de motorisation ou d'équipement du ménage.

1.5. Choix du mode : déplacements ou sorties

Jusqu'ici nous avons raisonné comme si l'individu faisait son choix de mode pour chacun des déplacements pris isolément. Les déplacements sont pourtant inclus dans une boucle ou sortie (ensemble des déplacements et activités inclus entre un départ et un retour successifs au domicile). C'est au niveau de cette boucle ou sortie que le choix du mode est effectué. Si l'on reprend les données lyonnaises, en 1995, 96 % des déplacements appartenant à une boucle dont le premier déplacement est effectué en voiture, sont eux-mêmes effectués en voiture (Bonnel, 2000a). Le pourcentage est un peu plus faible pour les transports collectifs avec 82 %. Les boucles comportent donc une forte homogénéité modale au moins pour les modes motorisés. Les modèles devraient donc conserver ce processus de choix. C'est globalement vrai pour les déplacements inclus dans des boucles constituées d'un seul aller retour, le mode présentant le plus faible coût généralisé est souvent le même pour le trajet aller et pour le retour. C'est, en revanche, beaucoup moins vrai pour les boucles plus complexes (boucles comprenant au moins trois déplacements) dont le poids s'accroît dans le temps.

Ce constat rejoint les résultats de travaux de recherche déjà anciens. De nombreux travaux menés à la fin des années 70 et tout au long des années 80 insistaient sur la nécessité d'analyser les comportements de mobilité à travers les schémas d'activités (cheminement spatio-temporel de l'individu permettant la réalisation d'activités) plutôt qu'à travers les déplacements (Hägerstrand, 1974 ; Lenntorp, 1978 ; Averous et al., 1979-80 ; Jones et al., 1980 ; Andan et al., 1988...). Ce renouveau (déjà ancien !) méthodologique des analyses de mobilité tarde toutefois à produire des outils opérationnels de modélisation.

Il est assez fréquent que des modèles à quatre étapes tirent partie de ces avancées au niveau de l'étape de génération (Lichère, Raux, 1997a, b ; Zhang et al., 2001). En revanche, nous ne connaissons qu'un seul modèle véritablement opérationnel utilisant cette notion dans la chaîne à quatre étapes (Fellendorf et al, 1997). Il s'agit du logiciel VISEM inclus dans le logiciel VISION de la société PTV. Le calage du modèle semble toutefois particulièrement gourmand en données pour obtenir des résultats garantissant une fiabilité suffisante.

D'autres modèles s'appuyant sur ce type d'approche sont en cours de développement, mais ils restent souvent confinés dans les milieux de recherche (Ettema, Timmermans, 1997 ; Timmermans et al., 1997 ; Bowman, Ben-Akiva, 2001).

2. PRINCIPALES FAMILLES DE MODELES DE REPARTITION MODALE

Les premiers modèles utilisés en France, comme pour l'étape de génération peuvent être classés au sein des modèles normatifs (section 2.1). Leur piètre qualité prévisionnelle a conduit à leur abandon rapide dans la plupart des pays au profit des modèles logit obtenus par régression logistique (section 2.2). Les problèmes de cohérence entre les étapes de distribution et de répartition modale ont conduit certains analystes à combiner ces deux étapes dans un seul modèle (section 2.3). Certains modèles proposent même de traiter les trois premières étapes dans une seule équation (section 2.4). Ces modèles sont appelés des modèles directs par opposition aux modèles séquentiels.

2.1. Modèles normatifs

Le principe est tout à fait similaire à celui que nous avons évoqué pour la génération (chapitre 6, p. 173). Il consiste à définir une norme de répartition modale. Cette norme peut être assez simple, voire plus complexe avec l'utilisation d'abaques donnant différentes courbes en fonction d'un certain nombre de paramètres. Nous donnons quelques illustrations de ces méthodes.

Type de destination	Mode principal	Bus sur site propre	Bus sur voie réservée	Bus sur voie normale
destinations centrales à fortes contraintes de stationnement	trajet direct	80 % ± 10 %	60 % ± 10 %	40 % ± 10 %
	trajet avec correspondance(s)	60 % ± 10 %	40 % ± 10 %	20 % ± 10 %
destinations périphériques à faibles contraintes de stationnement	trajet direct	50 % ± 10 %	30 % ± 10 %	10 % ± 5 %
	trajet avec correspondance(s)	30 % ± 10 %	10 % ± 5 %	5 % ± 5 %

Tableau 8.3 : Exemple de grille de part modale pour les transports en commun pour le motif domicile-travail (CETUR, 1990a)

Le tableau 8.3 fournit un exemple de grille tenant compte de la centralité de la destination du déplacement, des correspondances et de la protection du réseau. Ce tableau peut bien entendu être adapté en fonction des données locales lorsqu'elles sont disponibles. Dans le même esprit, le logiciel CartOPERA (anciennement OPERA) développé par le CETE Méditerranée (CETE Méditerranée, CERTU 2000a,b) propose une grille de part de marché de la voiture à partir d'une typologie de zones (schéma 8.1). Le pouvoir prédictif de telles normes est évidemment assez limité.

D'autres normes sont un peu plus élaborées utilisant des courbes (ou abaques) donnant la répartition modale en fonction du rapport de coût généralisé entre les transports collectifs et la voiture. Nous en donnons une illustration (graphique 8.2), mais d'autres exemples plus complexes peuvent être trouvés (notamment une série de 12 courbes en fonction de l'indice économique élevé ou faible de la zone origine, et de la fréquence des transports collectifs. Chacune de ces courbes fournit, dans le contexte de Toronto, la part modale des transports collectifs en fonction du rapport des temps réels transports collectifs et voiture particulière sur une liaison donnée (IAURP, 1963)).

Ces normes ne sont pratiquement plus utilisées actuellement avec le développement des modèles logit qui peuvent être facilement calés sous tableur dans leurs versions les plus simples.

Grilles de répartition modale

Etude: Variante:

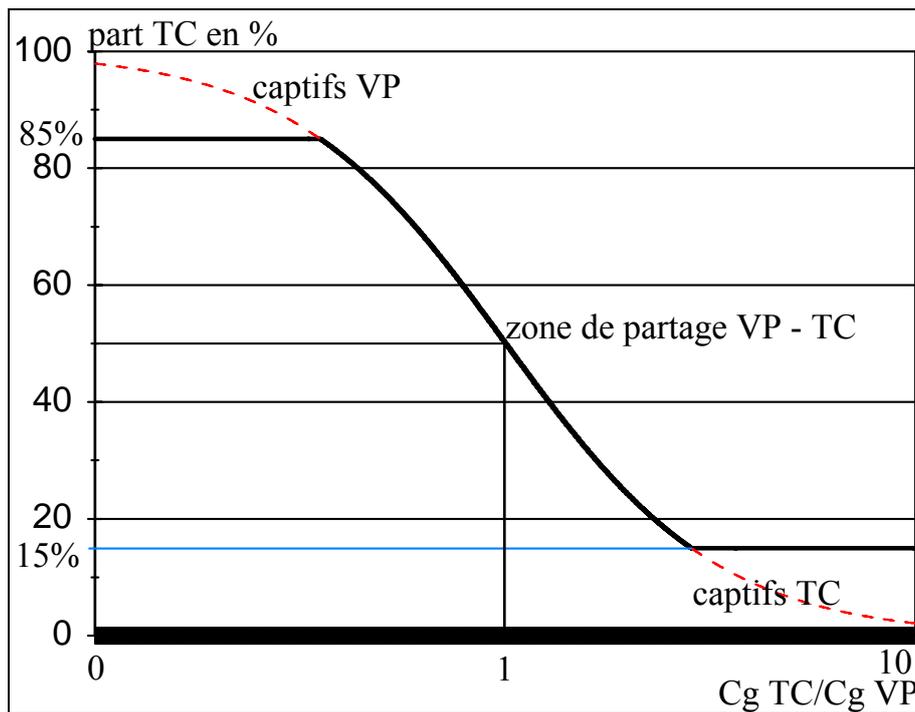
Motif du déplacement: Mode de transport:

VP TC 2R
 Domicile-Travail
 Domicile-Autres
 Secondaires

Type de relation	Taux d'utilisation du mode de transport selon la typologie des zones									
	1	2	3	4	5	6	7	8	9	10
1	93	80	82	100						
2	80	88	90	100						
3	82	90	100	100						
4	100	100	100	100						
5										
6										
7										
8										
9										
10										

Totaux 100% Imp fenêtre OK Annuler Aide

Schéma 8.1 : Exemple de grille de part modale pour la voiture pour le motif domicile-travail (extrait de CETE Méditerranée, CERTU 2000b)



Graphique 8.2 : Exemple de courbe de partage modal voiture-transports collectifs pour le motif domicile-travail (extrait de CETUR, 1990a)

2.2. Modèles logit ou régression logistique

Les fondements théoriques de ces modèles découlent de la théorie micro-économique (section 2.2.1). L'utilisation de la régression logistique par rapport à une formulation linéaire découle de l'analyse des données empiriques conduisant généralement à une courbe en « S » qui est très proche

de la courbe de répartition logit. Cette présentation trouve également sa justification théorique dans une approche probabiliste de l'utilité définie dans sa seule composante déterministe (section 2.2.2). La présentation avec deux modes peut en fait être très facilement généralisée à plusieurs modes avec le logit multinomial (section 2.2.3). Cette structure de modèle nécessite toutefois le respect de l'hypothèse d'indépendance entre les alternatives disponibles. Lorsque cette hypothèse n'est pas respectée, il est recommandé d'utiliser la structure de choix emboîté ou hiérarchique (section 2.2.4). Enfin, nous abordons le calage du logit (section 2.2.5).

2.2.1. Fondements théoriques

Nous reprenons rapidement les développements du chapitre 5. Nous avons indiqué que l'individu est supposé rationnel et que cette rationalité s'exprime à travers le choix de l'alternative présentant la plus grande utilité. Dans l'approche déterministe (et agrégée) qui est la notre dans ce chapitre, l'utilité s'exprime par (équation 5.11, p. 155) :

$$V = \sum_k \beta_k X_k \quad (8.1)$$

où V est l'utilité du bien. Par commodité de notation, nous avons abandonné l'indice n relatif à l'individu et l'indice i relatif à l'alternative ;

X_k correspond aux différentes variables aléatoires permettant d'estimer l'utilité ;

β_k est le coefficient de la variable X_k .

Les variables généralement utilisées appartiennent aux trois groupes de variables que nous avons identifiés dans la première section :

- coût généralisé. Il est exprimé dans sa forme généralisée ou à travers une description des temps élémentaires et des éléments de coûts monétaires dont les paramètres sont calés par régression logistique ;
- motorisation. Elle est définie à travers le taux de motorisation (nombre de voitures à disposition des ménages sur le nombre de ménages de la zone) ou le taux d'équipement (pourcentage de ménage équipé, c'est-à-dire disposant d'au moins une voiture ou multi-équipé, c'est-à-dire disposant de 2 voitures ou plus). Parfois, la population est segmentée selon le nombre de voitures du ménage ;
- stationnement. Les données étant rarement disponibles sur l'offre de stationnement, on trouve principalement deux types de solution. Dans la première, on utilise la densité (nombre de personnes + nombre d'emplois sur la surface de la zone) qui exprime d'une certaine manière la pression sur le stationnement (Raux, Lichère 1997a, b). Cet indicateur permet en général de bien exprimer la pression relative du stationnement entre les différentes zones de la ville centre de l'agglomération. Elle surestime parfois un peu la contrainte de stationnement dans les zones périphériques où celle-ci est généralement absente. La seconde méthode consiste à introduire une variable muette pour chacune des zones centrales où la pression sur le stationnement existe (Bonnel, Cabanne, 2000). Le calage permet alors de déterminer la pression sur le stationnement. En général, ces deux méthodes conduisent à des résultats proches. La recherche citée pour la seconde méthode a ainsi conduit à un rapport des coefficients de calage pour les deux zones de stationnement choisies à Lyon, identique au rapport des densités des deux zones ainsi définies. Enfin, on trouve également des modèles qui intègrent le stationnement dans le coût généralisé sous forme de temps de recherche de stationnement et de coût moyen du stationnement par zone.

Les autres variables sont plus rarement prises en compte. Ces variables sont le plus souvent intégrées dans la fonction d'utilité sous une forme additive (équation 8.1). D'autres formulations sont également choisies dans la mesure où elles permettent un meilleur calage, mais aussi et surtout correspondent aux comportements de choix des individus. Nous en donnons deux illustrations (dans les deux cas, nous exprimons non pas l'utilité d'un mode mais la différence d'utilité entre deux modes,

à savoir la voiture et les transports collectifs, qui constituent l'ensemble des choix disponibles. Ce choix est expliqué plus loin (équations 8.14 et 8.15, p. 232) :

$$V_{TCij} - V_{VPij} = \left(k_m + \frac{T_{g_{tc}ij}}{\tau_{tc_m} / \text{mot}_i} - \frac{T_{g_{vp}ij}}{\tau_{vp_m} * \text{mot}_i} - \frac{d_j}{\delta_m} \right) \quad (8.2)$$

où V_{TCij} , V_{VPij} sont respectivement l'utilité du mode transports collectifs, voiture particulière ;

$T_{g_{tc}ij}$ et $T_{g_{vp}ij}$ sont les temps généralisés sur l'origine-destination i/j ;

d_j est la densité de la zone de destination j . Cette variable représente la contrainte de stationnement ;

mot_i est la motorisation de la zone i . Dans cet exemple, la motorisation n'intervient pas de manière additive, mais exprime d'une certaine manière la pénibilité relative des temps généralisés des deux modes selon le niveau de motorisation. Une forte motorisation accroît la perception de pénibilité du temps transports collectifs et inversement réduit celle de la voiture ;

k_m , τ_{tc_m} , τ_{vp_m} , δ_m sont des paramètres de calage relatifs au motif m de déplacement (Lichère, Raux, 1997a, b).

$$(V_{VP} - V_{TC}) = a \frac{vp}{\text{mot}^m} + b_{tc} \text{mot}^m + c + b_1 B_1 + b_2 B_2 \quad (8.3)$$

où, mot : taux de motorisation ;

vp , tc : coût généralisé voiture et transports collectifs ;

B_1 et B_2 sont deux variables destinées à schématiser la contrainte de stationnement :

- B_1 telle que : $B_1 = 1$ si 1 (l'hypercentre) est la zone d'origine ou de destination, $B_1 = 0$ sinon ;

- B_2 telle que : $B_2 = 1$ si 2 (le reste de Lyon et Villeurbanne) est la zone d'origine ou de destination, $B_2 = 0$ sinon ;

A , b , b_1 , b_2 , c , m sont des coefficients de calage (Bonnel, Cabanne, 2000, équation 3.11, p. 85).

Considérons le cas le plus simple d'un choix entre deux modes de transport : voiture particulière et transports collectifs. L'individu choisit le mode dont l'utilité est la plus grande. Il choisira donc la voiture si :

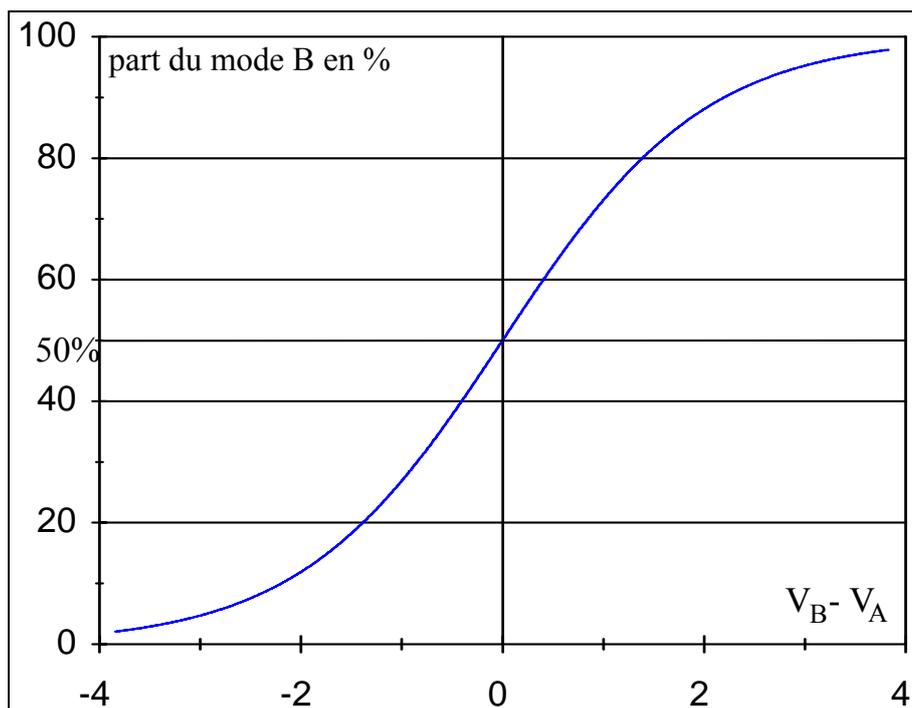
$$V_{VP} \geq V_{TC}, \text{ donc si } V_{VP} - V_{TC} \geq 0 \quad (8.4)$$

V_{VP} correspond à l'utilité du mode voiture particulière et V_{TC} à l'utilité du mode transports collectifs.

Une utilisation déterministe du modèle à partir de l'utilité moyenne de la voiture et des transports collectifs conduirait à un choix de la forme tout-ou-rien (0 ou 100 %) pour chacune des origines-destinations. Ce résultat ne correspond bien évidemment pas aux données empiriques disponibles.

2.2.2. Distribution logitistique

La confrontation avec les données empiriques montre au contraire que le choix entre les deux modes suit une courbe en S (graphique 8.3) en fonction de la différence d'utilité des deux modes de transport.



Graphique 8.3 : Illustration de la forme générale des courbes de partage modal en S

Cette forme de courbe en S qui correspond à l'observation empirique des données paraît assez intuitive. La part des individus choisissant le mode B lorsque l'utilité du mode A est nettement plus grande que celle du mode B est logiquement très faible. La croissance n'est que faible à mesure que la différence s'amenuise pour s'accélérer lorsque la différence tend vers 0 et conduire à une répartition égale entre les deux modes. L'évolution est ensuite inverse car le rôle de A et B s'inverse.

Cette répartition est similaire à celle d'une fonction de probabilité logit, dont l'équation de la fonction de répartition (graphique 8.3) est donnée par la formule suivante :

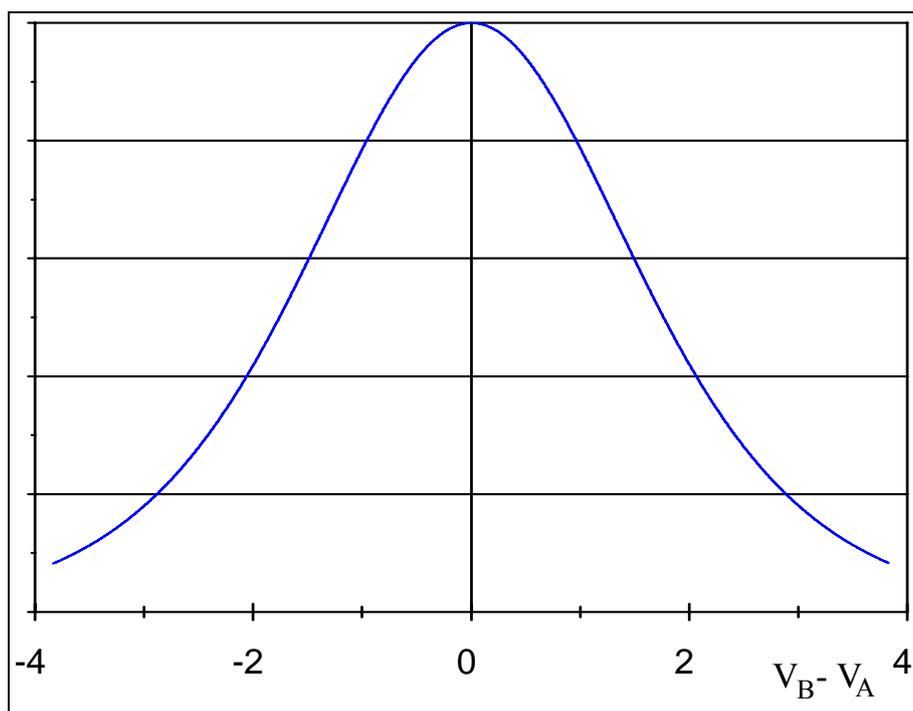
$$F(V_B - V_A) = \frac{1}{1 + e^{-\mu(V_B - V_A)}} \quad (8.5)$$

où μ est un paramètre positif,
et l'équation de la fonction de densité (graphique 8.4) par :

$$f(V_B - V_A) = \frac{\mu e^{-\mu(V_B - V_A)}}{\left[1 + e^{-\mu(V_B - V_A)}\right]^2} \quad (8.6)$$

Nous revenons dans le chapitre suivant sur cette distribution et sa justification théorique (chapitre 9, p. 254 et 255).

Ce choix apparaît en contradiction avec notre approche théorique dans un cadre agrégé et déterministe. En fait, implicitement nous avons développé une approche probabiliste. L'approche agrégée nous conduit à considérer un individu moyen se déplaçant entre une zone i et une zone j qui est confronté à un choix de mode de transport. Nous calculons ensuite pour cet individu moyen une utilité moyenne pour chacun des modes compte tenu des caractéristiques socio-économiques moyennes de cet individu moyen. Nous savons très bien que cet individu moyen est une fiction et qu'en fait les individus se distribuent autour de cet individu moyen. Nous pouvons alors utiliser l'approche probabiliste de l'utilité constante développée par Luce et Suppes (1965).



Graphique 8.4 : Illustration de la forme générale des fonctions de densité logistique

Nous reprenons ici la présentation qu'en font Ben-Akiva et Lerman (1985), soit :

$P_n(i)$ ou $P\langle i | C_n \rangle$, la probabilité que l'individu n choisisse l'alternative i appartenant à l'ensemble C_n des alternatives disponibles pour l'individu n . Nous utilisons la seconde notation lorsque nous faisons explicitement référence à l'univers de choix disponible pour l'individu n .

Cette probabilité doit évidemment respecter certaines propriétés :

$$0 \leq P\langle i | C_n \rangle \leq 1, \forall i \in C_n \quad (8.7)$$

l'égalité correspondant au cas de l'approche déterministe.

$$\sum_{i \in C_n} P\langle i | C_n \rangle = 1 \quad (8.8)$$

Pour tout sous-échantillon $\tilde{C}_n \subseteq C_n$, la probabilité de \tilde{C}_n dans C_n , est obtenue par :

$$P\langle \tilde{C}_n | C_n \rangle = \sum_{i \in \tilde{C}_n} P\langle i | C_n \rangle \quad (8.9)$$

D'où, on peut déduire la probabilité conditionnelle de i dans \tilde{C}_n :

$$P\langle i | \tilde{C}_n \subseteq C_n \rangle = \frac{P\langle i | C_n \rangle}{P\langle \tilde{C}_n | C_n \rangle} \quad (8.10)$$

à condition que l'une au moins des probabilités conditionnelles d'une alternative appartenant à \tilde{C}_n soit non nulle de telle sorte que $P\langle \tilde{C}_n | C_n \rangle$ soit non nul.

Le modèle le plus simple s'appuyant sur une approche probabiliste de l'utilité constante a été construit selon les auteurs précités par Luce (1959). Il découle de l'hypothèse suivante appelée « *choice axiom* » :

« Le « *choice axiom* » stipule que pour toute alternative du sous-ensemble \tilde{C}_n de l'ensemble C_n telle que $i \in \tilde{C}_n \subseteq C_n$, on a

$$P\langle i | \tilde{C}_n \subseteq C_n \rangle = P\langle i | \tilde{C}_n \rangle \quad (8.11)$$

En d'autres termes, si certaines alternatives sont enlevées de l'ensemble des alternatives disponibles, les probabilités relatives entre les alternatives au sein du sous-ensemble ainsi construit restent inchangées » (Ben-Akiva, Lerman, 1985)

En reprenant les équations 8.10 et 8.11, on en déduit :

$$P\langle i | C_n \rangle = P\langle i | \tilde{C}_n \rangle * P\langle \tilde{C}_n | C_n \rangle \quad \text{pour tout } i \in \tilde{C}_n \subseteq C_n \quad (8.12)$$

De cet axiome, on peut déduire la propriété d'indépendance vis-à-vis des autres alternatives (*Independence from Irrelevant Alternatives property*) encore appelée IIA à partir des initiales en anglais :

$$\frac{P\langle i | \tilde{C}_n \rangle}{P\langle j | \tilde{C}_n \rangle} = \frac{P\langle i | C_n \rangle}{P\langle j | C_n \rangle}, \quad \text{pour tout } i, j \in \tilde{C}_n \subseteq C_n \quad (8.13)$$

Luce (1959) interprète cette propriété IIA comme pouvant être une version probabiliste du concept de transitivité. Nous revenons sur cette propriété dans la section suivante, car si elle permet une formulation très simple de la probabilité de choix, elle est aussi la principale limite de cette approche.

Luce (1959) a montré que dans le cas où cet axiome est respecté, la probabilité de choix de l'alternative i s'écrit très simplement. Dans un premier temps, nous nous limitons à un ensemble C_n ne comportant que deux alternatives i et j :

$$P\langle i | C_n \rangle = \frac{1}{1 + e^{-\mu(V_{in} - V_{jn})}} = \frac{e^{\mu V_{in}}}{e^{\mu V_{in}} + e^{\mu V_{jn}}} \quad (8.14)$$

μ étant une constante, elle est généralement intégrée dans la formulation de l'utilité, l'équation devient alors :

$$P\langle i | C_n \rangle = \frac{1}{1 + e^{-(V_{in} - V_{jn})}} = \frac{e^{V_{in}}}{e^{V_{in}} + e^{V_{jn}}} \quad (8.15)$$

On retrouve donc l'expression de l'équation 8.5. La distribution logistique que nous avons justifiée à partir de l'observation empirique des données d'enquête et d'une commodité analytique trouve donc ici une justification théorique. Nous verrons section 2.3 (équations 8.21 à 8.26, p. 236) que cette formulation peut être également dérivée de la théorie de la maximisation de l'entropie. Cette présentation avec deux alternatives conduisant au logit binomial peut être très facilement généralisée à plus de deux alternatives.

2.2.3. Logit multinomial et propriété IIA

Le logit est multinomial lorsqu'il y a plus de 2 choix possibles. Considérons les alternatives j appartenant à l'ensemble C_n des alternatives disponibles pour l'individu n . L'expression devient alors :

$$P(i|C_n) = \frac{e^{V_{in}}}{\sum_{j \in C_n} e^{V_{jn}}} \quad (8.16)$$

Ce modèle présente une propriété intéressante appelée IIA d'indépendance vis-à-vis des autres alternatives (Ben-Akiva, Lerman, 1985). Cette propriété s'exprime très simplement (équation 8.13). Elle signifie que le rapport de probabilité entre deux alternatives ne dépend pas des autres alternatives disponibles. Cette propriété fait la force du modèle logit, car elle permet d'introduire facilement un mode supplémentaire sans devoir redéfinir le modèle. Elle constitue également sa faiblesse illustrée par le paradoxe du bus bleu et du bus rouge (Ben-Akiva, Lerman, 1985 ; Ortuzar et Willumsen (1994) attribuent ce paradoxe à Mayberry (1973)).

Supposons que l'on soit en présence de deux modes voiture et bus et que chacun des modes ait une probabilité identique pour simplifier la présentation :

$$P_n(\text{VP}) = P_n(\text{bus}) = 0,5.$$

Supposons maintenant que par un éclair de génie, l'opérateur décide de peindre en bleu la moitié des bus, sachant que l'ensemble des bus étaient auparavant peint en rouge. On est donc maintenant en présence de 3 modes disponibles. Si la fréquence n'est pas un attribut de l'utilité du mode bus, l'utilité du mode bus bleu est identique à l'utilité du mode bus rouge (qui reste inchangée par rapport à la situation précédente). Compte tenu de la propriété IIA, on a toujours :

$$\frac{P_n(\text{VP})}{P_n(\text{bus rouge})} \text{ qui reste constant donc } = \frac{0,5}{0,5},$$

et $P_n(\text{bus rouge}) = P_n(\text{bus bleu})$, donc :

$$P_n(\text{VP}) = 1/3,$$

$$P_n(\text{bus bleu}) = 1/3,$$

$$P_n(\text{bus rouge}) = 1/3.$$

L'opérateur a donc réussi à accroître sa part de marché au détriment de la voiture.

Ce résultat n'est évidemment pas très réaliste, car, on s'attendrait en fait à conserver la même part de marché pour la voiture, les bus se partageant leur part de marché, soit :

$$P_n(\text{VP}) = 1/2,$$

$$P_n(\text{bus bleu}) = 1/4,$$

$$P_n(\text{bus rouge}) = 1/4.$$

Ce paradoxe découle en fait du « *choice axiom* ». Le respect de cet axiome impose d'avoir des alternatives dont les caractéristiques ne soient pas corrélées. Ici nous avons un cas extrême où les deux modes bus possèdent en fait des caractéristiques rigoureusement identiques à l'exception de la couleur qui a priori ne fait pas partie des variables déterminantes du choix du mode de transport et qui n'est pas un attribut pris en compte dans la fonction d'utilité.

Dans la pratique, la corrélation n'est pas toujours aussi explicite en particulier face à une nouvelle alternative (Hivert et al., 1988). Lorsque cette corrélation existe, il est recommandé de définir une structure emboîtée.

2.2.4. Logit emboîté (nested logit)

Il est assez fréquent que les modes collectifs violent la propriété IIA. Les caractéristiques de ces alternatives sont en effet souvent assez fortement corrélées. L'individu fait ainsi souvent un choix entre des modes collectifs ou individuels. S'il choisit un mode collectif, il fera ensuite un choix entre les différents modes collectifs concurrents disponibles. On peut illustrer cette structure dans le schéma suivant (schéma 8.2).

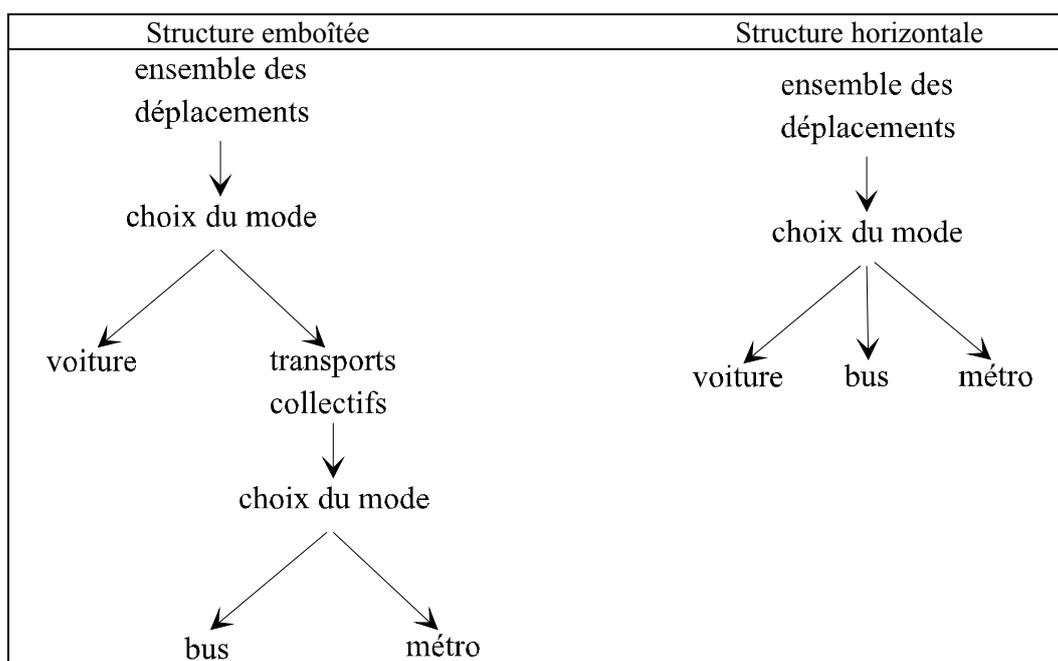


Schéma 8.2 : Structures de modèles logit multinomiaux

La structure emboîtée ou hiérarchique ne résout pas complètement le problème. Il n'est généralement pas très difficile de déterminer les variables pertinentes et d'estimer leur valeur pour caler le modèle logit pour le mode bus ou métro pour reprendre notre exemple. En revanche, il reste à déterminer la valeur de l'utilité pour le mode composite transports collectifs. Sur une origine-destination donnée, quelle valeur de temps de parcours doit-on retenir (moyenne, moyenne pondérée, plus faible valeur...) ? Nous verrons dans le prochain chapitre la méthode généralement utilisée pour résoudre le problème.

Le problème est encore accru en prévision lorsque l'on introduit un nouveau mode de transport. Si l'on reprend l'exemple précédent, et que l'on introduit des lignes de tramways. L'absence de données en référence pour ce mode pose évidemment le problème de la définition de l'utilité relative à ce mode et de sa position dans la structure hiérarchique (l'utilisation des méthodes d'enquêtes de préférences déclarées est souvent utile dans ce cas).

En revanche, on vérifiera aisément que la structure hiérarchique permet de résoudre le paradoxe précédent. Les utilités des bus bleus et rouges étant identiques, leur part de marché respective est également identique au sein du mode transports collectifs :

$$P_n(\text{bus rouge}) = P_n(\text{bus bleu}) = 0,5 * P_n(\text{transports collectifs}).$$

L'utilité du mode transports collectifs restant identique suite à la coloration de la moitié de la flotte de bus, on a toujours l'égalité des utilités transports collectifs et voiture :

$$P_n(\text{transports collectifs}) = P_n(\text{voiture}) = 0,5.$$

Donc :

$$P_n(\text{voiture}) = 0,5,$$

$$P_n(\text{bus rouge}) = 0,5 * P_n(\text{transports collectifs}) = 0,5 * 0,5 = 0,25,$$

$$P_n(\text{bus bleu}) = 0,5 * P_n(\text{transports collectifs}) = 0,5 * 0,5 = 0,25.$$

Dans la pratique, le calage du modèle emboîté est mené dans l'ordre inverse du processus de choix de l'individu. On calera donc un premier logit sur le choix entre le bus et le métro, puis un second logit entre les modes transports collectifs et voiture. De plus, l'utilité du mode composite transports collectifs est différente de l'utilité de chacun des modes la composant (chapitre 9, p. 263).

2.2.5. Calage des modèles logit

Dans le cas le plus simple d'un choix binaire entre deux modes, la part de marché s'obtient à partir de l'équation 8.15 :

$$P\langle i | C_n \rangle = \frac{1}{1 + e^{-(V_{in} - V_{jn})}}$$

Dans ce cas, il n'est pas nécessaire de calculer l'utilité de chacun des modes. Le calcul de la différence d'utilité est suffisant pour utiliser le modèle en prévision. L'équation peut alors s'écrire :

$$e^{-(V_{in} - V_{jn})} = \frac{1 - P\langle i | C_n \rangle}{P\langle i | C_n \rangle} \quad (8.17)$$

$$\text{soit : } (V_{in} - V_{jn}) = -\log\left(\frac{1 - P\langle i | C_n \rangle}{P\langle i | C_n \rangle}\right) \quad (8.18)$$

Dans cette expression, la probabilité de choix de l'alternative i correspond à la part de marché de l'alternative i pour chacune des origines-destinations. Lorsque l'on dispose de données d'enquête, cette part de marché est connue. Il reste donc à formaliser l'expression de la différence d'utilité entre les deux modes. Lorsque l'expression de l'utilité est linéaire, certains auteurs (Ortuzar, Willumsen, 1994) proposent de résoudre l'équation à l'aide de la régression linéaire (le terme de droite correspond à la variable à expliquer et le terme de gauche comprend les variables explicatives qui composent la différence d'utilité).

De nombreuses cases de la matrice étant vides ou tout au moins avec un faible nombre de déplacements, il est généralement nécessaire de procéder à des regroupements de zones. Dans cet exercice, on privilégie évidemment le regroupement des zones dont les caractéristiques (relatives aux variables explicatives) sont proches. Une autre technique consiste à ordonner les origines-destinations par ordre croissant de différence d'utilité et ensuite à définir des classes par intervalle croissant de différence d'utilité. La largeur de chaque intervalle doit être définie de telle sorte qu'il comporte un nombre suffisant de données. Le nombre d'intervalles dépend donc de la taille de l'échantillon. On définit ensuite une différence moyenne d'utilité et une part de marché moyenne pour chacune des classes sur lesquelles est faite la régression.

Lorsque l'expression de l'utilité est non linéaire, d'autres logiciels sont nécessaires. Dans ce cas, il est préférable d'utiliser les logiciels du marché permettant le calage de logit (par exemple Alogit ou encore les principaux logiciels de statistiques généralistes).

L'utilisation de la régression linéaire avec l'estimation selon la méthode des moindres carrés n'est toutefois pas la plus pertinente du fait des hypothèses nécessaires sur la distribution des résidus pour le modèle linéaire. Nous verrons d'ailleurs que cette méthode n'est pas applicable dans le cas de l'approche désagrégée, puisque la probabilité observée est systématiquement 1 ou 0 selon que l'alternative est choisie ou non. Il est plutôt recommandé d'utiliser la méthode du maximum de vraisemblance (chapitre 9, p. 271). Domencich et McFadden (1975) ont montré que cette méthode est préférable tant d'un point de vue théorique que pratique. Nous revenons sur cette procédure d'estimation et d'une manière générale sur l'estimation des différentes familles de modèles logit dans le prochain chapitre (chapitre 9, p. 270).

2.3. Modèles combinés distribution – répartition modale

En reprenant l'analogie du modèle gravitaire présenté dans le chapitre précédent (chapitre 7, p. 199) et les résultats de la section précédente, nous aboutissons à la formulation générale d'un modèle combiné destination – choix du mode :

$$T_{ij}^1 = \alpha_i \beta_j E_i A_j e^{-bCg_{ij}^1} \quad (8.19)$$

$$\text{et } P(1) = \frac{e^{-bCg_{ij}^1}}{\sum_k e^{-bCg_{ij}^k}} \quad (8.20)$$

où T_{ij}^1 est le trafic sur l'origine-destination i/j avec le mode 1 et $P(1)$ la part de marché du mode 1 parmi l'ensemble des modes k . Pour les autres termes, nous reprenons les notations du chapitre précédent.

Il est intéressant de constater que ce modèle peut être déduit de l'approche par la maximisation de l'entropie. Reprenons les équations ayant conduit au modèle gravitaire (équations 7.39, 7.40 et 7.48, p. 208 et 209) :

$$E_i - \sum_{j,k} T_{ij}^k = 0 \quad (8.21)$$

$$A_j - \sum_{i,k} T_{ij}^k = 0 \quad (8.22)$$

$$\forall i, j \quad T_{ij}^k \geq 0 \quad (8.23)$$

$$Cg - \sum_{i,j,k} T_{ij}^k * Cg_{ij}^k = 0 \quad (8.24)$$

Ce système d'équations conduit au modèle gravitaire (équation 7.52, p. 210) :

$$T_{ij}^1 = \alpha_i \beta_j E_i A_j e^{-bCg_{ij}^1} \quad (8.25)$$

La part de marché du mode 1 s'écrit donc :

$$P(l) = \frac{T_{ij}^l}{T_{ij}} = \frac{e^{-bCg_{ij}^l}}{\sum_k e^{-bCg_{ij}^k}} \quad (8.26)$$

Si l'on remplace le coût généralisé par l'utilité, on retrouve la formulation logit traditionnelle que nous avons présentée (équation 8.16, p. 233). Un des problèmes de cette formulation (Ortuzar, Willumsen, 1994) est le rôle joué par le coefficient b . Il doit tout à la fois permettre de caler le modèle de distribution et celui du choix du mode de transport. C'est probablement un peu trop pour un seul coefficient. On utilise donc une forme plus générale :

$$T_{ij}^k = \alpha_i \beta_j E_i A_j e^{-bK_{ij}} \frac{e^{-\lambda Cg_{ij}^k}}{\sum_k e^{-\lambda Cg_{ij}^k}} \quad (8.27)$$

où K_{ij} est le coût généralisé composite relatif à l'ensemble des modes présents et λ un paramètre de calage supplémentaire pour le choix du mode. On retrouve le problème récurrent de la détermination de ce coût généralisé composite (chapitre 7, p. 216). On peut retenir l'une des deux solutions présentées dans cette section :

- le coût généralisé le plus faible ;
- la moyenne des coûts généralisés pondérés par le poids de chaque mode :

$$K_{ij} = \sum_k P(k) * Cg_{ij}^k \quad (8.28)$$

Nous avons vu qu'en fait aucune des deux solutions n'est réellement satisfaisante. Williams (1977) propose une autre solution tout en montrant qu'il s'agit de la seule solution qui soit cohérente avec le principe de rationalité des choix du consommateur (nous revenons sur cette formulation dans la démonstration de la formulation de l'utilité composite pour le logit emboîté, chapitre 9, p. 261) :

$$K_{ij} = \frac{-1}{\lambda} \log \left(\sum_k e^{-\lambda Cg_{ij}^k} \right), \text{ avec } b \leq \lambda \quad (8.29)$$

Ce coût composite présente les propriétés intéressantes suivantes :

$$\forall i, j \quad K_{ij} \leq \min_k \{ Cg_{ij}^k \}; \quad (8.30)$$

$$\lim_{\lambda \rightarrow \infty} K_{ij} = \min_k \{ Cg_{ij}^k \}; \quad (8.31)$$

$$\frac{dK_{ij}}{dC_{ij}^k} = P(k). \quad (8.32)$$

2.4. Modèles directs

Les modèles directs, c'est-à-dire regroupant les trois premières étapes des modèles à quatre étapes sont peu courants. Leur développement s'est surtout effectué dans le courant des années soixante et au début de la décennie suivante. A notre connaissance, ils ne sont plus du tout utilisés aujourd'hui en raison de leur difficulté de calage et de validation. Enfin, ils ont surtout été développés en interurbain où leurs résultats semblent meilleur qu'en urbain (Domencich et al., 1968 cité par Manheim, 1984).

Nous reprenons trois exemples de modèles cités par Manheim (1984). Le premier appelé, modèle Kraft-SARC (Kraft, 1963), propose l'équation suivante :

$$T_{ij}^k = \phi_0 (P_i P_j)^{\phi_1} (I_i I_j)^{\phi_2} \prod_l (t_{ij}^l)^{\theta_{1kl}} (c_{ij}^l)^{\theta_{2kl}} \quad (8.33)$$

Cette formulation mérite quelques explications :

- T_{ij}^k représente le trafic sur l'origine-destination i/j pour le mode k ;
- P_i et P_j sont la population de la zone origine i et de la zone de destination j ; I_i et I_j les revenus moyens de ces zones. ϕ_1 et ϕ_2 sont les élasticités à la population et au revenu. Ces termes représentent les facteurs générateurs de déplacements comme dans le modèle de génération. On pourrait remplacer population et revenu par d'autres variables explicatives de la génération (chapitre 6) ;
- t_{ij}^l et c_{ij}^l sont respectivement le temps et le coût sur l'origine-destination i/j à l'aide du mode l . θ_{1kl} et θ_{2kl} sont des coefficients représentant l'élasticité au temps et au coût pour chacun des modes. Lorsque k et l sont identiques, θ_{1kl} représente l'élasticité directe, tandis que lorsque k et l sont différents, θ_{1kl} représente l'élasticité croisée. Comme précédemment, on pourrait ajouter d'autres variables qualifiant le niveau de service des modes de transports.

Il s'agit ainsi d'une forme dérivée du modèle gravitaire. On retrouve une forme similaire à celle de l'équation 8.19 (p. 236). Toutefois, la fonction de résistance est de forme puissance et non exponentielle. De plus, elle ne tient pas seulement compte des caractéristiques du mode considéré, mais prend également en compte les modes concurrents à travers les élasticités croisées. Enfin, dans cette expression, les émissions et attractions sont remplacées par les variables explicatives de ces émissions et attractions, à savoir dans ce modèle population et revenu.

Le second, modèle de Baumol-Quandt (Quandt, Baumol, 1966), présente une forme similaire au précédent, mais seul le mode le plus rapide et le mode le moins cher (donc pas forcément le même mode) interviennent au niveau des élasticités croisées et non pas l'ensemble des modes concurrents :

$$T_{ij}^k = \phi_0 (P_i P_j)^{\phi_1} (I_i I_j)^{\phi_2} (t_{ij}^k)^{\theta_1} (c_{ij}^k)^{\theta_2} (t_{ij}^{l_1})^{\theta_3} (c_{ij}^{l_2})^{\theta_4} \quad (8.34)$$

où l_1 et l_2 représentent respectivement le mode le plus rapide et le mode le moins cher sur l'origine-destination i/j .

Le fait de faire référence à chacun des modes présentant la meilleure performance pour chacune des composantes du niveau de service a conduit certains auteurs (Géraud, 1977) à introduire la notion de mode abstrait. Le mode abstrait étant défini comme un mode fictif dont les performances relativement à chacune des composantes du niveau de service sont égales à celle du meilleur mode pour la composante considérée.

Le dernier, le modèle de McLynn (McLynn, Woronka, 1969), est de formulation un peu plus complexe :

$$T_{ij}^k = \phi_0 (P_i P_j)^{\phi_1} (I_i I_j)^{\phi_2} \frac{(t_{ij}^k)^{\theta_{1k}} (c_{ij}^k)^{\theta_{2k}}}{\sum_l (t_{ij}^l)^{\theta_{1l}} (c_{ij}^l)^{\theta_{2l}}} \left[\sum_l \left((t_{ij}^l)^{\theta_{1l}} (c_{ij}^l)^{\theta_{2l}} \right)^\delta \right] \quad (8.35)$$

On retrouve en fait une structure similaire à la précédente. Toutefois, les deux termes qui diffèrent des équations précédentes peuvent être assimilés pour le premier au choix du mode et pour le

second à la fonction de résistance. L'équation s'apparente donc plutôt à la forme de l'équation 8.27 (p. 237) plutôt qu'à l'équation 8.19 (p. 236) comme dans les deux modèles précédents. Décomposons le terme exprimant le rapport pour identifier l'expression du choix du mode :

$$\text{Posons : } L_k = (t_{ij}^k)^{\theta 1_k} (c_{ij}^k)^{\theta 2_k} \tag{8.36}$$

L_k représente en quelque sorte le niveau de service composite du mode k . Le rapport de l'équation 8.35 devient :

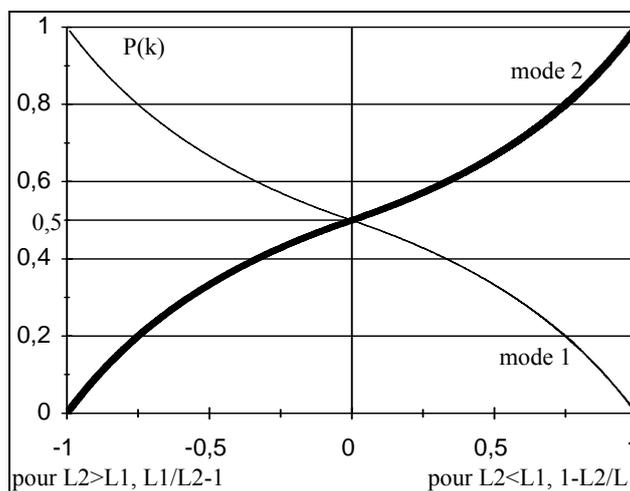
$$\frac{L_k}{\sum_m L_m} = \frac{1}{1 + \frac{\sum_{m \neq k} L_m}{L_k}} = P(k) \tag{8.37}$$

Cette équation représente la part de marché du mode k . Si l'on reprend l'équation 8.35, seul le rapport dépend du mode k , tous les autres termes sont identiques pour tous les modes. On retrouve donc bien la part de marché :

$$P(k) = \frac{T_{ij}^k}{\sum_m T_{ij}^m} = \frac{\frac{L_k}{\sum_m L_m}}{\sum_m \left(\frac{L_m}{\sum_m L_m} \right)}, \text{ le terme au dénominateur s'annulant } = \frac{L_k}{\sum_m L_m} \tag{8.38}$$

Pour simplifier considérons maintenant uniquement deux modes de transport, l'équation 8.38 devient :

$$P(1) = \frac{1}{1 + \frac{L_2}{L_1}} \text{ et } P(2) = \frac{1}{1 + \frac{L_1}{L_2}} \tag{8.39}$$



Graphique 8.5 : Part de marché des modes 1 et 2 en fonction des valeurs des niveaux de service composites des modes 1 et 2

On peut donc représenter la part de marché de chacun des modes en fonction de la valeur de L_1 et de L_2 . Afin d'avoir une représentation symétrique en fonction de la différence entre L_1 et L_2 ou

de leur rapport, il convient de les centrer et réduire. Cette transformation est effectuée par l'opération suivante :

$$\text{si } L_2 > L_1, \frac{L_1}{L_2} - 1 = \frac{L_1 - L_2}{L_2} ; \quad (8.40)$$

$$\text{si } L_2 < L_1, 1 - \frac{L_2}{L_1} = \frac{L_1 - L_2}{L_1} . \quad (8.41)$$

La représentation graphique de l'équation 8.39 figure dans le graphique 8.5. On retrouve bien une courbe en S qui correspond à l'allure générale des courbes de répartition modale. On pourrait toutefois s'attendre à une croissance plus modérée de la part de marché du mode 2 pour les valeurs faibles de L_1 par rapport à L_2 avec une croissance qui s'accélère à mesure que les valeurs de L_1 et L_2 s'égalisent comme pour la courbe du graphique 8.3.

3. PRINCIPALES LIMITES DE L'ETAPE DE REPARTITION MODALE

Nous conservons le plan des chapitres précédents qui s'appuie sur les trois conditions de l'opérationnalité des modèles (chapitre 2, p. 45) :

- pertinence : pertinence des variables explicatives du choix du mode (section 3.1), formalisation de la fonction d'utilité (section 3.3), « choice axiom » (section 3.4), stabilité des coefficients de calage (section 3.5) et segmentation selon les individus et/ou les motifs (section 3.6) ;
- cohérence d'objectifs : capacité à répondre aux problématiques de planification actuelles (section 3.2) ;
- cohérence interne : quelle valeur pour le coût généralisé (section 3.7) et le choix du mode succède ou précède le choix de la destination (ce point se pose dans des termes identiques à celui que nous avons traité pour la distribution : chapitre 7, p. 217) ;
- mesurabilité : qualité des données et notamment absence de données sur le stationnement (section 3.8).

Comme indiqué pour les chapitres précédents, les choix de classification proposés ci-dessus sont évidemment critiquables. Ils le sont par la définition même de la problématique des modèles puisque les trois conditions de l'opérationnalité des modèles sont contradictoires (chapitre 2, p. 48).

3.1. Pertinence des variables explicatives du choix du mode

La section 1.2 (p. 220) nous a permis d'aborder les principales variables explicatives du choix du mode. Force est de constater qu'en général seule une partie de ces variables est prise en compte comme nous l'avons mentionné. De plus, notre présentation a été volontairement limitée aux variables les plus couramment observées. En fait, l'analyse de la littérature propose un grand nombre d'approches de cette question mettant en évidence la multiplicité des facteurs à prendre en compte, certains d'entre eux de nature plus psycho-sociologique posant d'ailleurs de redoutables problèmes de mesurabilité.

Il est donc peu probable que la complexité du phénomène puisse être complètement traitée à l'aide des variables généralement utilisées à savoir coût généralisé, motorisation et stationnement. De plus, pour cette dernière dimension, les données sont rarement disponibles.

En théorie, rien n'empêche d'ajouter d'autres variables explicatives dans la fonction d'utilité. Il faut toutefois que les données soient disponibles. Sinon, les données doivent être produites. On recourt alors assez souvent aux méthodes d'enquêtes de préférences déclarées, notamment pour des variables de nature plus qualitative ou incluant une dimension perceptive par l'individu enquêté (valeur du

temps, temps d'attente...). Toutefois, si le nombre de variables est trop important, le calage peut s'avérer délicat avec des risques importants de coefficients peu significatifs tout comme pour la régression linéaire. L'additivité des variables dans la fonction d'utilité n'est pas non plus évident (section 3.3).

3.2. Prise en compte des problématiques de planification actuelles

L'évolution des problématiques de planification urbaine telle qu'elle transparaît des plans de déplacements urbains met en avant la volonté de maîtrise de la voiture au profit des autres modes de transport. Cet objectif passe par diverses mesures d'investissement dans les transports collectifs, mais aussi par des mesures diverses de trafic calmé (*traffic calming* en anglais) et de promotion des autres modes ou de la complémentarité modale. Il s'agit donc d'orienter le choix du mode en faveur des modes doux au détriment de la voiture. Il est donc nécessaire que le modèle soit capable de qualifier ces différentes mesures afin de simuler leur effet sur le choix du mode de transport.

En théorie rien n'empêche de prendre en compte ces dimensions dans le modèle afin de construire une fonction d'utilité qui permettent de tester la sensibilité à ces mesures. Dans la pratique, il faut reconnaître que l'exercice est assez délicat. Il passe généralement par une modification adéquate des réseaux et des paramètres de définition du coût généralisé ce qui ne permet pas toujours de différencier correctement les différentes mesures. Le problème spécifique du stationnement est évoqué ci-dessous (section 3.3).

3.3. Formalisation de la fonction d'utilité

En général, la fonction d'utilité présente une forme additive. Ce choix ne découle pas toujours d'une justification théorique relative à l'influence des différentes variables explicatives. En général, il s'agit au contraire d'une commodité analytique permettant un calage moins complexe et moins gourmand en temps de calcul.

Nous avons présenté d'autres formulations d'utilité non additives (équations 8.2 et 8.3, p. 229), mais il faut reconnaître que le choix de ces formulations est bien souvent plus heuristique que fondé sur des considérations théoriques.

Enfin, la spécification de la fonction rejoint les points abordés dans les deux sous-sections précédentes. Prenons l'exemple du stationnement pour l'aborder. L'absence de données pourrait conduire à ne pas prendre en compte cette dimension dans la spécification de la fonction d'utilité. On sait pourtant qu'il s'agit d'un levier d'action important sur le choix du mode (Massot, Orfeuill, 1989-91). Dans les villes de Berne ou de Zurich où le stationnement tant public que privé est fortement contrôlé, l'usage des transports collectifs est nettement plus important que dans des villes ayant également une offre de transports collectifs abondante (Bonnel, 1995b ; Kaufmann, Guidez, 1996, 2000). Même s'il ne s'agit pas du seul élément expliquant la réussite des villes suisses, il est reconnu qu'il s'agit d'un levier d'action efficace sur le choix modal. De ce fait, l'absence de variables exprimant la contrainte de stationnement, biaise forcément les coefficients des autres variables car l'estimation de leur valeur prend indirectement en compte l'effet du stationnement.

On utilise donc diverses techniques permettant d'approcher cette contrainte. Nous en avons vu deux exemples dans la section 2.2.1 (équations 8.2 et 8.3, p. 228) : variables muettes ou utilisation de la densité comme indicateur approchant la pression sur le stationnement. Ces techniques sont nécessaires dans le calage du choix du mode. En revanche, elles ne permettent pas de simuler l'impact de politiques contrastées de stationnement, sauf à le faire de manière plus ou moins approximative en jouant sur les variables ou la valeur des paramètres introduits pour prendre en compte le stationnement.

D'une manière plus générale, toute variable importante dans le processus de choix du mode des individus qui est oubliée contribue à biaiser l'estimation des coefficients des autres variables. C'est la raison pour laquelle l'analyse des facteurs explicatifs de ce choix est fondamentale, préalablement à la conception du modèle (schéma 2.1, p. 34).

3.4. « Choice axiom » et propriété IIA

Nous avons indiqué que le logit est la forme la plus courante de modèle utilisée pour le choix du mode dans une procédure agrégée. Le choix du logit découle de la simplicité de sa formulation et de son estimation comparativement à d'autres modèles. Toutefois, nous avons vu que la propriété IIA (indépendance vis-à-vis des autres alternatives) qui découle du « choice axiom » et qui fait la force du modèle, fait aussi sa faiblesse dans la mesure où cette propriété n'est pas toujours confortée par les données empiriques. Elle impose d'analyser la corrélation éventuelle entre les alternatives disponibles. La structure emboîtée permet de pallier partiellement les problèmes rencontrés dans ce cas.

3.5. Stabilité des coefficients de calage

La stabilité dans le temps des coefficients de calage est un problème récurrent qui est présent pour chacune des étapes des modèles à quatre étapes. Le calage étant souvent effectué sur des données correspondant à une seule date, l'influence du temps ne peut être prise en compte. Nous n'insistons pas sur cette question qui a déjà été traitée dans les chapitres 6 (p. 187) et 7 (p. 215).

3.6. Quelle segmentation de la demande

La question a déjà été abordée assez longuement dans l'introduction (section 1.4, p. 224).

3.7. Quelle valeur pour le coût généralisé ?

Le problème a lui aussi été évoqué dans l'introduction (section 1.3, p. 223). Nous ne faisons que resituer les enjeux :

- cohérence entre les étapes : avec la distribution, mais aussi avec l'étape suivante d'affectation ;
- définition du coût généralisé ou de l'utilité pour les modes composites dans les modèles emboîtés ;
- choix de la période temporelle (chapitre 8, p. 223).

Ce dernier point est particulièrement important dans la mesure où de plus en plus, la modélisation ne répond plus seulement à des objectifs de dimensionnement d'infrastructures, mais aussi (voire surtout) à des objectifs de simulation de politiques de déplacements (chapitre 1). Dès lors, ce qui nous intéresse n'est plus seulement le trafic pour chaque mode à l'heure ou la période de pointe, mais la répartition du trafic sur l'ensemble de la journée voire sur une plus longue période. Le choix d'une valeur représentative de la période de pointe ne sera pas pertinent pour analyser la répartition modale tout au long de la journée, car le coût généralisé pour la voiture est en général plus faible en dehors de cette période et à l'inverse celui des transports collectifs souvent plus élevé. Une solution consiste à segmenter la matrice de demande totale en autant de périodes que nécessaire pour obtenir des coûts généralisés relativement homogènes sur chacune des périodes. Cette solution présente toutefois l'inconvénient d'être plus gourmande en données pour la phase de calage et plus coûteuse en temps de calage. De ce fait, il est fréquent que la phase de choix modal se limite à la période de pointe avec la définition de coefficient de passage de la pointe à la journée pour chacun des modes pour obtenir une matrice jour par mode.

3.8. *Qualité des données*

Il s'agit d'un problème récurrent, commun à toutes les étapes de la modélisation. Il permet d'insister une nouvelle fois sur l'importance de la qualité de la production des données. Il permet de rappeler également l'importance de la taille de l'échantillon dans la définition des enquêtes.

La question de la taille des échantillons qui est en général trop faible pour obtenir des matrices origines-destinations suffisamment fiables se pose dans des termes identiques à ceux du chapitre précédent. Le problème est même accru puisque les matrices issues de l'étape précédente doivent être éclatées selon les modes. Nous avons évoqué les solutions possibles pour faire face à ce problème section 2.2.5 (p. 235).

Un problème complémentaire se pose en ce qui concerne le stationnement. Généralement, on ne dispose pas de données permettant d'estimer le nombre de places de stationnement. Dans le meilleur des cas, on peut approcher l'offre de stationnement publique. En revanche, l'offre de stationnement privé est totalement inconnue ne permettant pas une quantification, même très grossière. De plus, cette offre privée est probablement supérieure à l'offre publique au moins dans les zones où la pression sur le stationnement est forte. Nous avons abordé dans les sous-sections précédentes les conséquences de cette absence de données. On ne peut donc que recommander un effort de recherche et de production de données pour avancer sur cette question.

Toutefois, la principale critique à l'encontre de ces modèles porte sur le caractère agrégé de l'approche. En définissant un individu moyen, l'analyste introduit un individu fictif. Or, il n'est pas sûr que cet individu moyen se comportera toujours comme tel en situation de prévision, c'est-à-dire comme la moyenne des individus qu'il est censé représenter. De plus, disposant d'une information à niveau désagrégé, il paraît logique de vouloir exploiter toute la richesse de cette information. Il est ainsi possible de faire des hypothèses relatives au comportement de l'individu x doté de caractéristiques spécifiques et effectuant un déplacement déterminé. Le calage en exploitant les données individuelles permet de restituer cette richesse. Enfin, les modèles désagrégés « *sont basés sur une théorie du comportement individuel et ne constituent pas une analogie avec des théories issues de la physique* » (Ortuzar, Willumsen, 1994). Ce sont les principaux arguments avancés par les promoteurs de l'approche désagrégée pour justifier la plus grande pertinence de cette approche comparativement à l'approche agrégée. Ces critiques ne sont pas limitées au choix du mode. Elles s'appliquent à l'ensemble de la chaîne à quatre étapes. Si nous abordons cette question en conclusion de ce chapitre sur le choix du mode, c'est pour rappeler que cette étape a constitué le terrain de développement de l'approche désagrégée et constitue toujours son terrain de prédilection.

Chapitre 9 - Modèles désagrégés de choix discrets

D'après McFadden (2000), les modèles désagrégés dans la forme que nous leur connaissons aujourd'hui sont nés en 1970 sous l'impulsion de travaux menés principalement autour de Cambridge à Charles River Associates, M.I.T. et Cambridge Systematics sous la direction de John Meyer, Marvin Manheim et Paul Roberts, auxquels McFadden a participé. Ils résultent de programmes de recherche financés par le Ministère américain des transports autour du Corridor Nord-Est et du BART à San Francisco. Cette formalisation en 1970 ne signifie toutefois pas une absence de travaux dans les décennies précédentes. Parmi les précurseurs, on retrouve fréquemment cités (McFadden, 2000 ; Ben-Akiva, Lerman, 1985, Manheim, 1984) Luce (Luce, 1959 ; Luce, Suppes, 1965) pour ses travaux sur l'approche probabiliste de l'utilité qui débouche sur le modèle logit sous certaines hypothèses, mais aussi Thurstone (1927) pour le « *Thurstonian Random Utility Maximization model* ». Enfin, McFadden situe les premières applications du modèle logit binomial en biostatistiques dans les années cinquante, puis les premières applications dans le domaine des transports au début des années soixante par Warner en 1962 dans une étude du choix modal, puis dans des études sur la valeur du temps (Beesley, 1965 ; Lisco, 1967 ; Lave, 1970). Il serait toutefois probablement injuste de ne pas faire référence à certains auteurs français, et notamment Abraham et Coquand (1961) qui ont développé un modèle de choix entre itinéraires concurrents fondé sur une approche aléatoire des coûts généralisés de transport qui conduit à la formulation du modèle probit (équation 10.19, p. 338, le modèle probit est présenté section 2.3, p. 251), ainsi que les travaux menés par la SETEC en collaboration avec le CERAU (CERAU, SETEC, 1968) et le SAEI (Marche, Dumontier, 1974) mentionnés par Géraud (1977). McFadden et Ben-Akiva sont ensuite probablement les deux auteurs qui ont le plus contribué à la formalisation et aux développements théoriques et statistiques des différentes familles de modèles incluses dans la théorie de la maximisation de l'utilité aléatoire (*Random Utility Maximisation* (RUM)), objet de ce chapitre (McFadden, 2000 pour un historique de la contribution de cet auteur, qui a obtenu le Prix Nobel de l'économie pour ses travaux sur les choix discrets, en 2000). Ces travaux se poursuivent aujourd'hui principalement autour de recherches permettant de traiter les contextes où les hypothèses du logit et notamment la propriété IIA qui en découle ne sont pas suffisamment respectées (McFadden, 2000 ; Bhat, 1997, 2000, 2001 ; Brownstone, 2000). Si ces derniers travaux constituent des avancées théoriques indéniables, il n'est pas certain qu'ils contribuent fortement à une amélioration de l'opérationnalité des outils disponibles. Nous reviendrons en conclusion sur ces questions.

Mais avant, il est nécessaire de présenter les outils plus traditionnels de la modélisation désagrégée des choix discrets. Nous précisons tout d'abord les fondements théoriques en complément des apports du chapitre 5 (section 1), ce qui nous permettra de générer les différentes familles de modèles de choix discrets (section 2). L'estimation de ces modèles nécessite l'utilisation de la méthode du maximum de vraisemblance. Nous présentons cette méthode et les différents tests qui en découlent (section 3). Ces tests permettent d'estimer la qualité statistique du modèle selon un principe similaire à celui que nous avons présenté pour le modèle linéaire (chapitre 6, p. 174). Cette appréciation de la qualité statistique ne doit toutefois pas occulter l'analyse de la pertinence et de la cohérence d'objectifs du modèle. L'outil statistique ne mesure qu'une adéquation aux données. En aucun cas, il ne donne du sens à la relation étudiée. C'est à l'analyste de fonder sur le plan théorique les relations testées sur le plan statistique (schéma 2.1, p. 34). Nous examinons donc quelques questions importantes relatives à la spécification et à l'analyse des modèles de choix discrets (section 4). Enfin, si le modèle désagrégé permet une estimation du modèle au niveau de l'individu en conservant toute la richesse des données, cela ne doit pas occulter que la prévision ne s'effectue généralement pas avec le même niveau de désagrégation. Si les caractéristiques des individus sont connues en situation de référence grâce aux enquêtes, ces mêmes caractéristiques ne sont plus

disponibles au niveau de l'individu en prévision. Cela pose la question de l'agrégation des données en prévision (section 5).

1. FONDEMENTS THEORIQUES

Cette section complète la présentation du chapitre 5. L'approche de l'utilité n'est plus déterministe, mais aléatoire. Toutefois, à la différence du chapitre précédent (chapitre 8, p. 229), nous ne considérons plus que l'individu a un comportement aléatoire. Nous nous situons davantage dans la droite ligne de l'approche rationaliste de l'individu cherchant à maximiser son utilité. Toutefois, nous ne connaissons pas complètement cette utilité pour un certain nombre de raisons évoquées chapitre 5 (p. 158). Nous adoptons une expression aléatoire de l'utilité (section 1.1). Une seconde dimension de l'approche développée dans ce chapitre mérite d'être explicitée. L'univers de choix dans lequel l'individu est placé est fini. Il n'a qu'un nombre limité de choix disponibles au sein desquels il effectue ses choix, ce qui justifie le titre de ce chapitre : modèles désagrégés de choix discrets (section 1.2). L'approche théorique conduisant à une définition de l'utilité aléatoire ne permet plus, à elle seule, de déterminer le choix de l'individu. Il est nécessaire de faire des hypothèses complémentaires sur la forme de la distribution du terme aléatoire. C'est le choix de la fonction statistique qui donne naissance à différentes familles de modèles (section 2).

1.1. Utilité aléatoire

L'approche aléatoire de l'utilité permet de tenir compte d'une partie déterministe de l'utilité et d'une partie aléatoire. La partie déterministe correspond aux éléments connus influençant l'utilité de l'alternative considérée pour un individu donné. Toutefois, l'expérience montre que face à des situations identiques (au regard des paramètres pris en compte dans la partie déterministe de l'utilité) l'individu n'effectue pas toujours le même choix. De même, deux individus aux caractéristiques identiques placés face à des contextes de choix identiques (identiques au regard des paramètres pris en compte dans la partie déterministe de l'utilité) ne font pas toujours le même choix. Plutôt que de considérer que l'individu a un comportement probabiliste pour expliquer ces différences de choix (comme dans le chapitre 8, p. 229), nous considérons dans cette approche que l'individu choisit toujours l'alternative lui permettant de maximiser son utilité. Toutefois, notre connaissance des phénomènes ne nous permet pas d'identifier ou de mesurer exactement tous les éléments pris en compte par l'individu dans son évaluation de l'utilité de chacune des alternatives. Le terme aléatoire correspond à ces facteurs inconnus ou connus mais non ou mal mesurés. D'après Ben-Akiva et Lerman (1985) la formalisation de cette approche est due à Manski (1977). Nous l'illustrons rapidement dans cette section. On en trouvera une formalisation plus rigoureuse et surtout dans un contexte plus large non limité au champ transport dans McFadden (2000).

L'approche probabiliste de l'utilité aléatoire nous conduit à la formulation suivante de l'utilité (équation 5.13, p. 157) :

$$U_{in} = V_{in} + \varepsilon_{in} = f(X_{in}) + \varepsilon_{in} \quad (9.1)$$

où ε_{in} est un terme aléatoire, X_{in} un vecteur de paramètres explicatifs de l'utilité de l'alternative i pour l'individu n . Ce vecteur correspond en général aux caractéristiques socio-économiques de l'individu et aux caractéristiques du déplacement considéré.

Sous l'hypothèse de rationalité du consommateur qui cherche à maximiser son utilité, l'expression de la probabilité de choix de l'alternative i parmi l'ensemble C_n des alternatives disponibles pour l'individu n , (notée $P\langle i | C_n \rangle$ ou $P_n(i)$), s'écrit donc :

$$P\langle i | C_n \rangle = P(U_{in} \geq U_{jn}) \text{ pour tout } j \neq i \in C_n \quad (9.2)$$

qui peut encore s'écrire :

$$P(i|C_n) = P((V_{in} - V_{jn}) \geq (\varepsilon_{jn} - \varepsilon_{in})) \text{ pour tout } j \neq i \in C_n \quad (9.3)$$

Cette approche théorique est baptisée théorie de la maximisation de l'utilité aléatoire (*Random Utility Maximisation* ou RUM dans la littérature anglo-saxonne). Les initiales RUM sont parfois reprises dans la littérature française.

1.2. Choix discrets

Le contexte de choix dans lequel on se trouve correspond à un univers de choix fini. Le nombre d'alternatives disponibles pour un individu est toujours une quantité finie. Ce nombre est généralement assez faible dans le cas du choix du mode, mais peut être beaucoup plus important dans la modélisation du choix de la destination (modèle de distribution) ou du choix d'itinéraire (modèle d'affectation), voire dans une combinaison de ces choix, mais il est dans tous les cas fini. Nous allons illustrer les conséquences de cette situation en revenant à la présentation du chapitre 5 (nous empruntons pour cela certains développements de Ben-Akiva et Lerman (1985)).

La présentation du comportement du consommateur (chapitre 5, p. 145), introduit un individu qui cherche à maximiser son utilité sous contrainte de revenu. Cette utilité dépend de la consommation q_1, \dots, q_I des produits 1, ..., I disponibles (équation 5.4, p. 147) :

$$\begin{aligned} & \max_{q_i} (U(q_1, \dots, q_I)) \\ & \text{sous la contrainte } \sum_i p_i q_i \leq R \end{aligned} \quad (9.4)$$

La solution du problème consiste généralement à poser le Lagrangien du problème d'optimisation :

$$\max_{q_i, \lambda} \left[U(q_1, \dots, q_I) - \lambda \left(\sum_i (p_i q_i - R) \right) \right] \quad (9.5)$$

En annulant ensuite les dérivées du Lagrangien par rapport aux q_i et à λ , on obtient un système d'équations conduisant à la solution.

Pour pouvoir dériver le Lagrangien et donc la fonction d'utilité par rapport aux quantités q_1, \dots, q_I , il faut que la fonction d'utilité soit continue par rapport à ces variables.

Le type de problème auquel nous sommes confrontés se pose dans des termes un peu différents. Nous ne cherchons plus à déterminer la quantité de produits consommés, mais le choix entre différents produits ou plus précisément ici entre des alternatives. Dans ce processus de choix, l'individu ne choisit qu'une seule alternative :

$$q_i * q_j = 0 \text{ pour tout } i \neq j ;$$

de plus la quantité choisie, non divisible, est égale à 1

$$q_i = \begin{cases} 1 & \text{si le mode } i \text{ est choisi} \\ 0 & \text{si le mode } i \text{ n'est pas choisi} \end{cases} \text{ pour tout } i$$

Si l'on a trois modes, l'utilité ne peut prendre que trois valeurs :

$$U(1,0,0) ; U(0,1,0) \text{ ou } U(0,0,1)$$

La fonction U n'est donc plus différentiable par rapport aux variables q_i . « *Parce que dans cet exemple, on fait face à un choix discret, il n'est plus possible d'utiliser les techniques de maximisation consistant à calculer les dérivées de la fonction de demande. Ainsi, la représentation discrète de l'ensemble des alternatives nécessite une approche analytique différente. Nous conservons le concept de rationalité du consommateur. La seule différence entre la théorie des choix discrets et la théorie du comportement du consommateur est qu'au lieu de dériver des fonctions de demande, nous travaillons directement avec les fonctions d'utilité.* » (Ben-Akiva, Lerman, 1985).

Cette différence explique que l'on trouve parfois dans la littérature les fondements théoriques présentés sous la théorie des choix discrets et que les familles de modèles qui en découlent sont généralement dénommées *modèles de choix discrets*.

2. PRINCIPALES FAMILLES DE MODELES DESAGREGES

Les fondements théoriques étant rappelés, la résolution du problème de maximisation nécessite de faire des hypothèses sur le terme d'erreur. C'est le choix de la distribution qui conduit aux différentes familles de modèles.

Nous commençons par présenter une distribution linéaire qui malgré son manque de réalisme permet d'illustrer simplement le problème (section 2.1). Nous exposons ensuite la formulation générale d'estimation de la probabilité de choix (section 2.2).

Nous poursuivons avec la distribution normale des résidus qui présente le mérite d'être fondée sur le plan théorique sous des hypothèses assez simples (section 2.3). Cette distribution débouche sur les modèles probit qui possèdent l'inconvénient majeur d'être très difficiles à estimer dès que le nombre d'alternatives est supérieur à 3.

Pour faire face à cette difficulté, des hypothèses restrictives ont été posées sur la distribution des résidus et notamment une distribution identique et indépendante pour chacune des alternatives (IID selon les initiales anglaises : *Independent and Identically Distributed*) afin de simplifier la formalisation mathématique de la solution et son estimation. Parmi ces distributions, la distribution de Gumbel est la plus courante, car elle débouche sur un modèle très simple y compris face à des univers de choix multidimensionnels : le modèle logit multinomial (MNL pour *Multinomial Logit*) (section 2.4). Sa simplicité découle de ses hypothèses qui entraînent la propriété déjà évoquée d'indépendance vis-à-vis des autres alternatives (IIA).

Cette propriété fait toutefois la faiblesse de ces modèles car elle supporte difficilement la confrontation empirique lorsque le nombre d'alternatives s'accroît. Une première solution a été proposée avec une structure de choix hiérarchique qui limite les problèmes les plus aigus. Elle conduit aux modèles logit emboîtés (NL pour *Nested Logit*) encore appelés modèles hiérarchiques (HL pour *Hierarchical Logit*) (section 2.5). D'autres modèles ont également été développés qui sont inclus dans la famille de modèles plus large dénommée modèles de la valeur extrême généralisée (GEV *Generalized Extreme Value models*) (section 2.6).

La solution n'est toutefois que partielle car les hypothèses des modèles de la valeur extrême généralisée et à plus forte raison des modèles logit emboîtés et logit multinomiaux ne sont pas fondées sur un plan théorique. Différentes solutions ont été proposées pour faire face à certaines de ces hypothèses. Mais une nouvelle famille de modèles est en train de se développer qui permet de réduire fortement ces hypothèses. Cette famille est regroupée sous le terme de logit multinomial mixte (de l'anglais *Mixed Multinomial Logit* (MMNL)). Nous ne ferons toutefois qu'en présenter les principes car elle fait appel à des techniques d'estimation sophistiquées qui dépassent les ambitions de cet ouvrage (section 2.7).

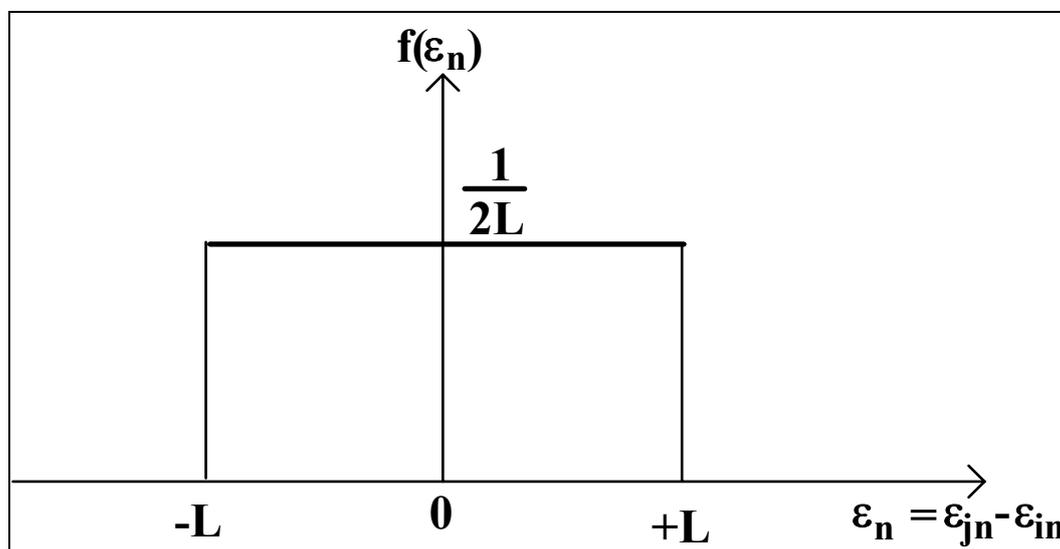
2.1. Modèle de probabilité linéaire

C'est le modèle le plus simple et le plus facile à développer. Toutefois, il n'est pas du tout réaliste et n'a donc d'autre intérêt que de présenter le principe de dérivation d'un modèle de choix discret à partir d'une distribution statistique particulière des résidus.

Pour simplifier la description et la représentation graphique, nous nous limitons au cas d'un choix binaire entre les deux modes voiture et transports collectifs par exemple. Dans ce cas, l'équation 9.3 devient :

$$P(i|C_n) = P((V_{in} - V_{jn}) \geq (\varepsilon_{jn} - \varepsilon_{in})) \quad (9.6)$$

La probabilité de choix de l'alternative i dépend donc de la distribution de la variable aléatoire $\varepsilon_{jn} - \varepsilon_{in}$. Nous choisissons donc une distribution des résidus uniforme entre deux valeurs fixées $-L$ et $+L$, avec L strictement positif. La fonction de densité est représentée dans le graphique 9.1.



Graphique 9.1 : Fonction de densité uniforme

La probabilité que l'individu n choisisse l'alternative i s'écrit donc :

$$P_n(i) = \begin{cases} 0 & \text{si } V_{in} - V_{jn} < -L, \\ \int_{-L}^{V_{in} - V_{jn}} f(\varepsilon_n) d\varepsilon_n = \frac{V_{in} - V_{jn} + L}{2L} & \text{si } -L \leq V_{in} - V_{jn} \leq L, \\ 1 & \text{si } V_{in} - V_{jn} > L. \end{cases} \quad (9.7)$$

La fonction de probabilité en fonction $V_{in} - V_{jn}$ est une droite entre $-L$ et $+L$ représentée sur le graphique 9.2. Cette formulation pose d'évidents problèmes de pertinence dans la mesure où les données empiriques conduisent généralement à une courbe sigmoïde (de la forme du graphique 8.3, p. 230). De plus, la courbe présente un problème de discontinuité de la dérivée en $-L$ et $+L$. Enfin, les données empiriques montrent en général une valeur non nulle, bien que faible, de la probabilité pour des valeurs inférieures à $-L$ ou supérieures à $+L$.

L'exemple met toutefois en évidence le principe de dérivation du modèle. Il permet également d'illustrer une autre caractéristique de la plupart des modèles qui est un problème d'échelle des coefficients. Si l'expression de la partie déterministe de l'utilité est également linéaire (de formulation

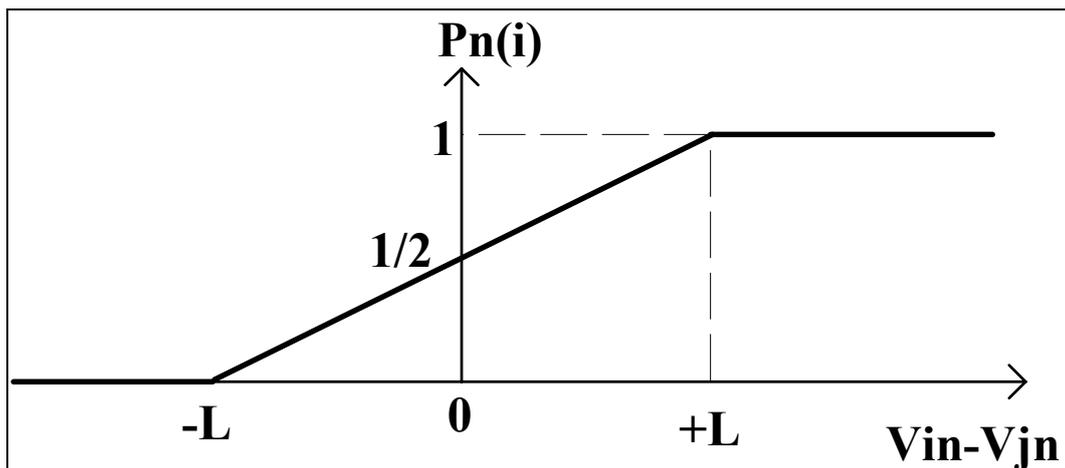
additive relativement aux différentes variables explicatives prises en compte (comme dans l'équation 5.14, p. 157)), la probabilité de choix de i , en reprenant l'équation 9.6 s'écrit :

$$P(i|C_n) = P\left(\sum_k \beta_k (X_{ikn} - X_{jkn}) \geq (\varepsilon_{jn} - \varepsilon_{in})\right) \quad (9.8)$$

ou X est le vecteur des caractéristiques du choix entrant dans la détermination de la fonction d'utilité et β le vecteur de paramètres de la fonction d'utilité.

$$\text{Donc } P_n(i) = \frac{\sum_k \beta_k (X_{ikn} - X_{jkn}) + L}{2L} \quad \text{si } -L \leq \sum_k \beta_k (X_{ikn} - X_{jkn}) \leq L \quad (9.9)$$

Si on multiplie le vecteur de paramètres β par un même coefficient α ainsi que L , on obtient la même probabilité. Donc l'étendue de la distribution des résidus est arbitraire et ne peut être déterminée. On retrouvera des indéterminations similaires dans les autres modèles. De ce fait, la comparaison des coefficients de la fonction d'utilité entre différents modèles doit tenir compte de la dispersion du vecteur de la partie aléatoire de l'utilité, c'est-à-dire de l'écart-type de la distribution. Les coefficients doivent donc être comparés en tenant compte du rapport d'écart-type des distributions du terme aléatoire. Nous retrouverons le problème dans la comparaison des fonctions d'utilité obtenues à l'aide d'un modèle probit ou logit (section 2.4.3, p. 255).



Graphique 9.2 : Fonction de probabilité du modèle de distribution linéaire

2.2. Formulation générale

Nous empruntons à Ben-Akiva et Lerman (1985) et à Ortuzar et Willumsen (1994) pour cette présentation. Nous reprenons l'équation 9.3 (p. 247) :

$$P(i|C_n) = P((V_{in} - V_{jn}) \geq (\varepsilon_{jn} - \varepsilon_{in})) \quad \text{pour tout } j \neq i \in C_n$$

Ne connaissant pas la distribution des termes d'erreur (c'est-à-dire la partie aléatoire de l'utilité), nous représentons leur distribution par la fonction de densité f_n :

$$f_n(\varepsilon) = f_n(\varepsilon_{1n}, \dots, \varepsilon_{J_n n}) \quad (9.10)$$

avec J_n le nombre d'alternatives disponibles dans C_n pour l'individu n .

La probabilité peut alors s'écrire :

$$P\langle i | C_n \rangle = \int_{R_{J_n}} f_n(\varepsilon) d\varepsilon \quad (9.11)$$

où R_{J_n} représente l'univers des valeurs de V_{in} , V_{jn} , ε_{in} et ε_{jn} respectant l'inégalité dans la parenthèse de l'équation 9.3 (p. 247) reprise ci-dessus, que l'on peut présenter à travers le système d'équations suivant :

$$R_{J_n} = \begin{cases} V_{in} - V_{jn} \geq \varepsilon_{jn} - \varepsilon_{in} & \text{pour tout } i, j \in C_n \\ V_{in} + \varepsilon_{in} \geq 0 & i \in C_n \end{cases}$$

L'équation 9.11 peut également s'écrire (sans aucune perte de généralité, nous posons $i = 1$ pour simplifier l'écriture de l'équation) :

$$P\langle 1 | C_n \rangle = \int_{\varepsilon_{1n} = -\infty}^{+\infty} \int_{\varepsilon_{2n} = -\infty}^{V_{1n} - V_{2n} + \varepsilon_{1n}} \dots \int_{\varepsilon_{J_n n} = -\infty}^{V_{1n} - V_{J_n n} + \varepsilon_{1n}} f_n(\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{J_n n}) d\varepsilon_{J_n n} d\varepsilon_{J_n - 1, n} \dots d\varepsilon_{2n} d\varepsilon_{1n} \quad (9.12)$$

En posant F_n la fonction de répartition cumulée de f_n et F_{ni} la dérivée partielle de F_n par rapport au $i^{\text{ième}}$ terme ε_{in} , l'équation 9.12 devient :

$$P\langle 1 | C_n \rangle = \int_{\varepsilon_{1n} = -\infty}^{+\infty} F_{n1}(\varepsilon_{1n}, V_{1n} - V_{2n} + \varepsilon_{1n}, \dots, V_{1n} - V_{J_n n} + \varepsilon_{1n}) d\varepsilon_{1n} \quad (9.13)$$

L'intégrale de l'équation 9.13 peut être interprétée de la manière suivante. Si l'on fixe l'erreur ε_{1n} à une valeur donnée, l'intégrale représente la probabilité que ε_{1n} égale cette valeur et que l'ensemble des autres termes d'erreur respecte la condition :

$$V_{1n} + \varepsilon_{1n} \geq V_{jn} + \varepsilon_{jn}, \text{ quel que } j \neq 1 \text{ appartenant } C_n.$$

En intégrant sur l'ensemble des valeurs possibles de ε_{1n} de $-\infty$ à $+\infty$, on obtient la probabilité que l'alternative 1 soit choisie. C'est le choix de la fonction de densité f_n et de la matrice de variance-covariance des termes d'erreur qui permet de dériver les différentes formes de modèles.

2.3. Modèle probit

La distribution linéaire n'étant pas réaliste, il convient de proposer une distribution qui puisse être justifiée sur le plan théorique. La distribution normale en fait partie. On interprète alors le terme d'erreur comme résultant de la variation d'un grand nombre de facteurs inconnus mais indépendants entre eux. Si ce nombre est suffisamment important (ce que l'on suppose), on peut alors assimiler la distribution à une distribution normale.

Une hypothèse supplémentaire est nécessaire pour dériver le modèle probit. On suppose que la matrice de variance-covariance du terme d'erreur entre les alternatives possède une structure identique entre les individus.

Considérons toujours un choix binaire. Nous supposons que les termes d'erreur ε_{in} et ε_{jn} ont tous deux une distribution normale de moyenne nulle et d'écart-type σ_i et σ_j . En n'imposant pas d'hypothèses supplémentaires sur les termes d'erreurs, nous avons une matrice de variance-covariance

qui a structure générale inconnue notée Σ_{ij} . Sous ces hypothèses, le terme $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ a aussi une distribution normale de moyenne nulle et d'écart-type :

$$\sigma = \sqrt{\sigma_i^2 + \sigma_j^2 - 2\sigma_{ij}} \quad (9.14)$$

avec σ_{ij} covariance des termes d'erreur ε_{in} et $\varepsilon_{jn} = \rho\sigma_i\sigma_j$ où ρ est le coefficient de corrélation entre ε_{in} et ε_{jn}

Dans ce cas, la probabilité devient :

$$P\langle i | C_n \rangle = \int_{-\infty}^{V_{in} - V_{jn}} \frac{1}{\sqrt{2\pi}\sigma} e^{\left[\frac{-1}{2} \left(\frac{\varepsilon_n}{\sigma} \right)^2 \right]} d\varepsilon_n \quad (9.15)$$

en posant $u_n = \frac{\varepsilon_n}{\sigma}$, l'équation 9.15 peut s'écrire :

$$P\langle i | C_n \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(V_{in} - V_{jn})/\sigma} e^{\left[\frac{-1}{2} (u_n)^2 \right]} du_n \quad (9.16)$$

qui s'écrit en reprenant la notation habituelle de la distribution cumulée de la loi normale standard Φ :

$$P\langle i | C_n \rangle = \Phi \left(\frac{V_{in} - V_{jn}}{\sigma} \right) \quad (9.17)$$

Ce modèle est appelé le modèle probit binomial. Ce modèle peut être généralisé à plusieurs alternatives prenant le nom de probit multinomial, sa formulation devient toutefois très rapidement extrêmement complexe.

Dans le cas d'une expression additive de la partie déterministe de l'utilité, l'équation devient :

$$P\langle i | C_n \rangle = \Phi \left(\frac{\sum_k \beta_k (X_{ikn} - X_{jkn})}{\sigma} \right) \quad (9.18)$$

Comme pour le modèle de probabilité linéaire, la valeur de σ peut être fixée arbitrairement, l'estimation de chacun des paramètres du vecteur β étant affectée d'une manière uniforme par un coefficient $1/\sigma$. Il est donc possible de choisir une valeur de $\sigma = 1$.

La fonction de distribution de la loi normale et la fonction de probabilité figurent dans les graphiques 9.3 et 9.4. Pour cette représentation, nous avons choisi une valeur de $\sigma = \frac{\pi}{\sqrt{3}}$ afin d'assurer une comparabilité avec la distribution logit (nous verrons pourquoi dans la section 2.4.3, p. 255).

2.4. Modèle logit

Nous présentons tout d'abord les hypothèses du modèle logit (section 2.4.1). Nous explicitons alors le modèle logit et la distribution de probabilité correspondante (section 2.4.2). Ce modèle découle d'une loi de distribution des résidus qui est la loi de Gumbel (section 2.4.3). Cette loi possède un certain nombre de propriétés qui permettent de démontrer la formulation du modèle logit

(section 2.4.4). Enfin nous examinons la propriété IIA d'indépendance vis-à-vis des autres alternatives (section 2.4.5).

2.4.1. Hypothèses du modèle logit

La complexité de la formulation probit à partir de 3 alternatives et plus encore au-delà a conduit à rechercher des distributions statistiques proches de la loi normale permettant de déboucher sur des formulations plus simples des lois de probabilité. Pour cela des hypothèses supplémentaires doivent être ajoutées. Nous reprenons ci-dessous, les trois hypothèses nécessaires présentées par Bhat (2000) :

- les parties aléatoires de l'utilité pour les différentes alternatives sont indépendantes entre elles et identiquement distribuées (noté IID). L'hypothèse d'indépendance signifie notamment qu'il n'y a pas de facteurs non observés qui soient communs à deux alternatives. L'hypothèse n'est évidemment pas toujours facile à vérifier dans la mesure où par nature les facteurs non pris en compte dans la partie déterministe de l'utilité sont pour la plupart inconnus. En général plutôt que d'être en mesure de valider l'hypothèse, on examine plutôt si l'on a de bonnes raisons de la rejeter. C'est généralement le cas lorsque l'on a plusieurs modes de transports collectifs. Ils peuvent partager entre eux des éléments de confort ou tout au moins le caractère collectif opposé au caractère individuel de la voiture. En général, on peut faire des observations identiques lorsque l'on subdivise le mode voiture entre voiture conducteur ou passager ou voiture possédée ou à disposition. L'hypothèse d'indépendance est donc souvent assez forte lorsque l'on a un grand nombre de choix. Le second élément de l'hypothèse concerne l'étendue de la variation autour de la partie déterministe de l'utilité. On considère que les termes d'erreurs sont distribués identiquement pour chacune des alternatives. Si l'on reprend l'exemple du confort qui n'est généralement pas inclus dans la partie déterministe de l'utilité, on fait l'hypothèse d'une variation similaire pour l'ensemble des modes. L'hypothèse est en général assez forte dans la mesure où les conditions de confort varient en général beaucoup pour les transports collectifs urbains et peu pour la voiture ;
- la seconde hypothèse concerne l'homogénéité des réponses aux différents attributs de la partie déterministe de l'utilité entre les individus. Plus précisément, le modèle n'autorise pas de variation de perception (*taste variation* en anglais) d'un attribut de la fonction d'utilité due à des caractéristiques non observées (c'est-à-dire non présentes dans la partie déterministe de l'utilité) de l'individu. Plus concrètement encore, les paramètres relatifs à chacune des variables présentes dans la fonction d'utilité sont identiques pour tous les individus. La perception du temps ou du coût du déplacement doit donc être identique pour tous les individus. Il faut retenir un coefficient unique, c'est-à-dire une valeur moyenne. Nous revenons sur cette hypothèse dans la conclusion pour examiner les conséquences sur la valeur du temps ou les temps d'accès aux transports collectifs ;
- la troisième hypothèse concerne la structure de la matrice de variance-covariance du terme d'erreur entre les alternatives. Comme pour le modèle probit, on fait une hypothèse de structure identique entre les individus. Concrètement cela signifie que la variation due aux facteurs non observés (c'est-à-dire non pris en compte dans la partie déterministe de l'utilité) est identique entre les individus. Si l'on reprend l'exemple du confort (variable généralement non prise en compte), il n'est pas certain que tous les individus soient soumis aux mêmes conditions et variations de confort y compris pour des déplacements différents. Un autre exemple concerne le choix du mode et de l'itinéraire. On peut raisonnablement penser que la variance risque d'augmenter à mesure que la distance s'accroît. En ce qui concerne l'homogénéité de la covariance (nous reprenons l'exemple cité par Koppelman et Sethi (2000)), l'hypothèse risque fort d'être contredite dans des situations où le degré de concurrence entre modes dépend des caractéristiques du déplacement ou de l'individu. Par exemple le fer et la voiture sont en situation de concurrence pour des déplacements courts, ce qui n'est pas le cas de l'avion. Inversement, pour des destinations plus longues, la concurrence peut être forte entre l'avion et le train.

Ces hypothèses sont donc assez fortes. C'est pour tenter d'en réduire la portée que d'autres modèles ont été développés comme nous le verrons dans les sections suivantes. Il doit être clair que ces hypothèses sont également présentes dans le cas du modèle agrégé à travers la fiction de l'individu moyen.

2.4.2. Formulation du modèle logit

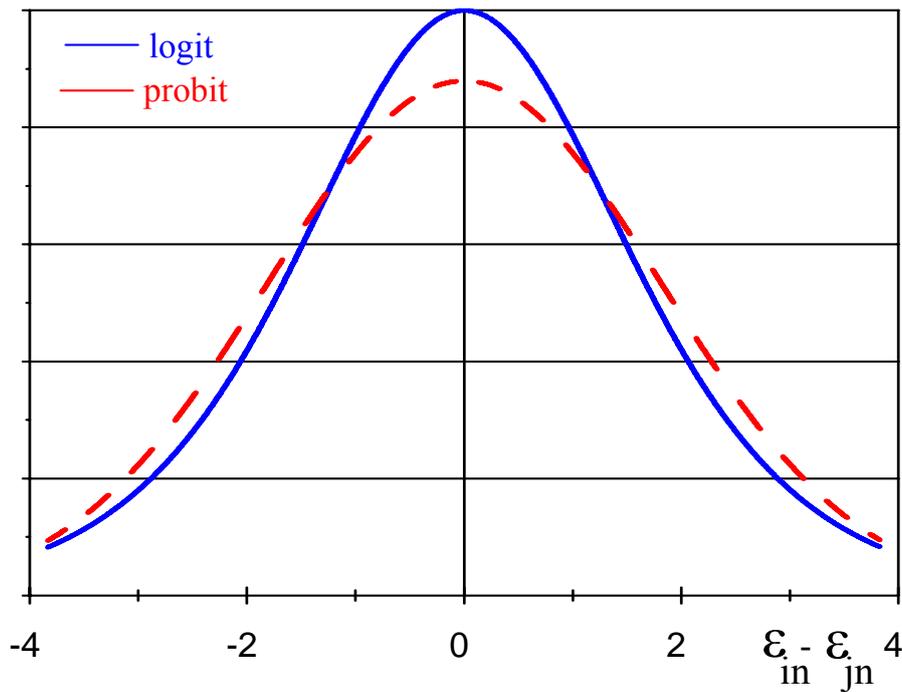
Sous ces hypothèses, le modèle le plus répandu est le modèle logit multinomial (MNL). Il a une formulation extrêmement simple, même dans un cas multidimensionnel :

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{\sum_{j \in C_n} e^{\mu V_{jn}}} \tag{9.19}$$

où μ est un paramètre strictement positif.

Dans le cas d'un choix binaire, le modèle binomial s'exprime par :

$$P(i|C_n) = \frac{e^{\mu V_{in}}}{e^{\mu V_{in}} + e^{\mu V_{jn}}} = \frac{1}{1 + e^{-\mu(V_{in} - V_{jn})}} \tag{9.20}$$



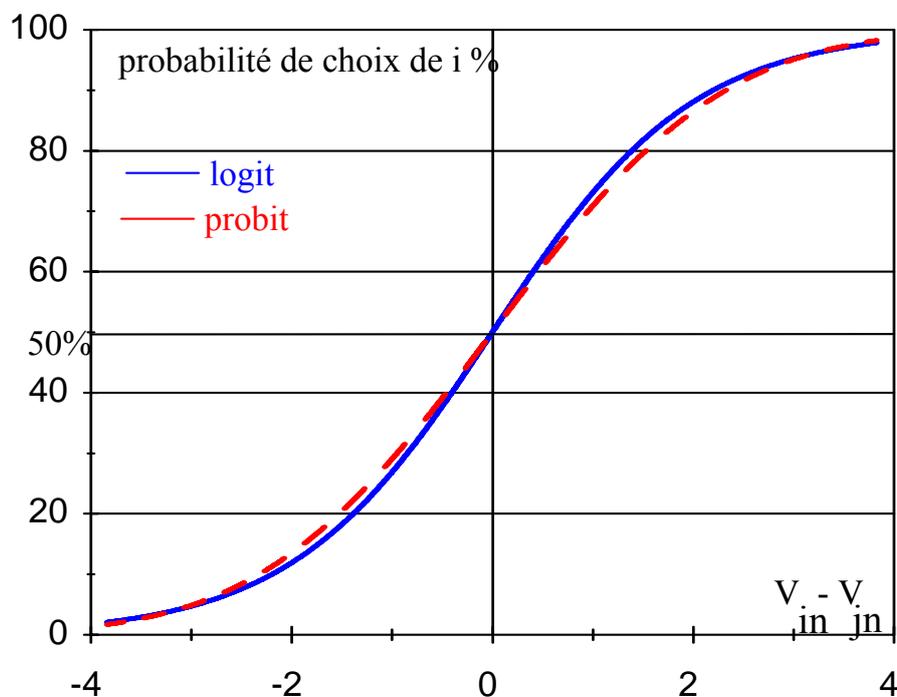
Graphique 9.3 : Fonction de distribution logit ($\mu = 1$) et probit ($\sigma = \frac{\pi}{\sqrt{3}}$)

et dans le cas d'une expression additive pour la partie déterministe de l'utilité :

$$P(i|C_n) = \frac{1}{1 + e^{-\mu \sum_k \beta_k (X_{ikn} - X_{jkn})}} \tag{9.21}$$

Comme pour les deux précédents modèles, le paramètre μ ne peut pas être identifié séparément du vecteur de coefficients β . Il est généralement pris égal à 1. Dans ce cas, la distribution du terme

d'erreur $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ figure dans le graphique 9.3. Elle conduit à la fonction de probabilité de choix qui figure graphique 9.4. Ces graphiques montrent que la distribution logistique de ε_n a un maximum plus élevé que la distribution probit et une queue de distribution légèrement plus grande (cela n'apparaît pas sur le graphique 9.3, car le phénomène n'est présent que pour des valeurs plus élevées de ε_n).



Graphique 9.4 : Fonction de probabilité logit ($\mu = 1$) et probit ($\sigma = \frac{\pi}{\sqrt{3}}$)

2.4.3. Loi de Gumbel

Le modèle logit dérive en fait d'une loi de distribution particulière des résidus qui est la distribution de Gumbel. Cette loi a les caractéristiques suivantes (Ben-Akiva, Lerman, 1985) :

$$\text{fonction de répartition } F(\varepsilon) = e^{-e^{-\mu(\varepsilon-\eta)}} \quad (9.22)$$

$$\text{fonction de densité } f(\varepsilon) = \mu e^{-\mu(\varepsilon-\eta)} * e^{-e^{-\mu(\varepsilon-\eta)}} \quad (9.23)$$

où η est un facteur de localisation et μ un facteur d'échelle strictement positif (graphique 9.5), qui constituent les deux paramètres d'une distribution de Gumbel.

Cette distribution a les propriétés suivantes (Ben-Akiva, Lerman, 1985) :

1. mode (valeur la plus fréquente donc la plus élevée de la distribution) = η ;
2. moyenne = $\eta + \gamma/\mu$ où γ est la constante de Euler ($\approx 0,577$) ;

$$3. \text{ variance} = \frac{\pi^2}{6\mu^2} ; \quad (9.24)$$

4. si ε est une variable distribuée selon la loi de Gumbel de paramètres (η, μ) et V et $\alpha > 0$ des constantes, alors $\alpha\varepsilon + V$ une variable distribuée selon la loi de Gumbel de paramètres $(\alpha\eta + V, \frac{\mu}{\alpha})$;

5. si ε_1 et ε_2 sont deux variables indépendantes distribuées selon la loi de Gumbel de paramètres (η_1, μ) et (η_2, μ) , alors $\varepsilon = \varepsilon_1 - \varepsilon_2$ est logistiquement distribué :

$$F(\varepsilon) = \frac{1}{1 + e^{\mu(\eta_2 - \eta_1 - \varepsilon)}} \quad (9.25)$$

6. si ε_1 et ε_2 sont deux variables indépendantes distribuées selon la loi de Gumbel de paramètres (η_1, μ) et (η_2, μ) , alors $\max(\varepsilon_1, \varepsilon_2)$ est une variable distribuée selon la loi de Gumbel de paramètres :

$$\left(\frac{1}{\mu} \ln(e^{\mu\eta_1} + e^{\mu\eta_2}), \mu \right) \quad (9.26)$$

Cette propriété peut être généralisée à J variables indépendantes distribuées selon la loi de Gumbel de paramètres (η_j, μ) , dont le maximum : $\max_j(\varepsilon_j)$ est une variable distribuée selon la loi de Gumbel de paramètres :

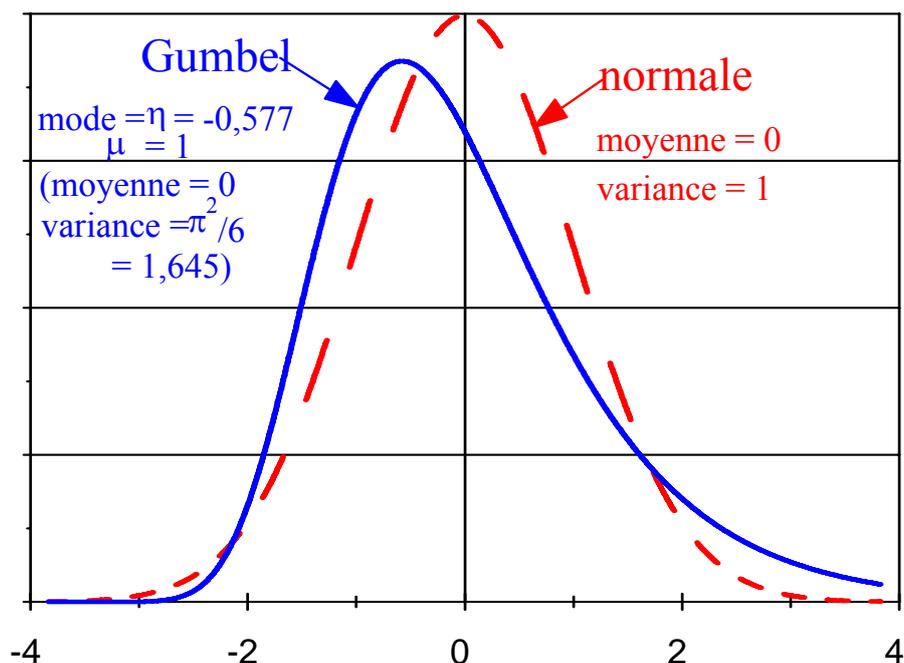
$$\left(\frac{1}{\mu} \ln\left(\sum_{j=1}^J e^{\mu\eta_j}\right), \mu \right) \quad (9.27)$$

Les propriétés 4 à 6 sont démontrées par Domencich et McFadden (1975). Elles permettent de démontrer la formulation des modèles logit multinomiaux et logit emboîtés.

Dans le calage du modèle logit binomial η_1 et η_2 ainsi que μ ne peuvent être déterminés séparément des autres coefficients du vecteur β . C'est la raison pour laquelle, on retient généralement une valeur de $\mu = 1$. De même, η_1 et η_2 sont généralement intégrés dans la constante de la fonction d'utilité de telle sorte que la moyenne des distributions soit nulle. Dans ce cas, la variance de chacune des distributions ε_{in} et ε_{jn} est égale à $\frac{\pi^2}{6}$ (propriété 3 ci-dessus) et la variance de $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ est égal à $\frac{\pi^2}{3}$, ce qui explique les choix qui ont été faits dans les graphiques 9.3 et 9.4 des paramètres $\mu = 1$ pour le modèle logit et $\sigma = \frac{\pi}{\sqrt{3}}$ pour le modèle probit afin d'avoir des distributions comparables pour $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ et la fonction de probabilité. Ces valeurs ont également été reprises pour le graphique 9.5 en centrant les distributions, ce qui nous donne $\eta = -0,577$.

Pour toute comparaison des coefficients de la fonction d'utilité entre un modèle logit et un modèle probit (retenant les valeurs habituelles : moyenne nulle et écart-type = 1 pour la distribution normale et moyenne nulle et $\mu = 1$ pour la distribution de Gumbel), les coefficients du modèle probit doivent être divisés par $\frac{\pi}{\sqrt{3}}$.

L'illustration des fonctions de densité de la loi de Gumbel et de la loi normale (graphique 9.5) met en évidence des formes assez différentes. Si les distributions logistiques et probits sont très proches, ce n'est pas tout à fait le cas des lois de distribution des résidus.



Graphique 9.5 : Fonctions de densité de la loi de Gumbel (centrée, $\mu = 1$) et de la loi normale (centrée réduite)

2.4.4. Démonstration du modèle logit

La propriété 5 permet de dériver le modèle logit binomial. Soit deux modes i et j dont l'utilité pour l'individu n est respectivement $U_{in} = V_{in} + \varepsilon_{in}$ et $U_{jn} = V_{jn} + \varepsilon_{jn}$ avec ε_{in} et ε_{jn} tous deux distribués selon une loi de Gumbel de paramètres $(0, \mu)$. L'hypothèse de l'individu maximisant son utilité individuelle, nous permet d'écrire :

$$P_n(i) = P(U_{in} \geq U_{jn}) = P((V_{in} + \varepsilon_{in}) \geq (V_{jn} + \varepsilon_{jn})) = P((\varepsilon_{jn} - \varepsilon_{in}) \leq (V_{in} - V_{jn}))$$

L'utilisation de la propriété 5, nous permet d'écrire :

$$P_n(i) = \frac{1}{1 + e^{\mu(V_{jn} - V_{in})}} = \frac{e^{\mu(V_{in})}}{e^{\mu(V_{in})} + e^{\mu(V_{jn})}}$$

Les propriétés 4 à 6 permettent d'établir la formulation du logit multinomial. Soit J_n alternatives dont l'utilité s'exprime : $U_{jn} = V_{jn} + \varepsilon_{jn}$, avec ε_{jn} distribués selon une loi de Gumbel de paramètres $(0, \mu)$, quel que soit $j = 1, \dots, J_n$. La probabilité que l'alternative i soit choisie, s'exprime :

$$P_n(i) = P\left(V_{in} + \varepsilon_{in} \geq \max_{j \neq i} (V_{jn} + \varepsilon_{jn})\right) \quad (9.28)$$

Si on définit : $U_n^* = \max_{j \neq i} (V_{jn} + \varepsilon_{jn})$, alors par application des propriétés 4 et 6, U_n^* est

distribué selon une loi de Gumbel de paramètres $\left(\frac{1}{\mu} \ln\left(\sum_{j \neq i} e^{\mu V_{jn}}\right), \mu\right)$.

En utilisant la propriété 4, on peut écrire : $U_n^* = V_n^* + \varepsilon_n^*$, avec ε_n^* distribué selon une loi de Gumbel de paramètres $(0, \mu)$ et $V_n^* = \frac{1}{\mu} \ln(\sum_{j \neq i} e^{\mu V_{jn}})$ (9.29)

L'équation 9.28 peut alors s'écrire : $P_n(i) = P(V_{in} + \varepsilon_{in} \geq V_n^* + \varepsilon_n^*)$. En reprenant la démonstration du logit binomial qui s'appuie sur la propriété 5, on obtient :

$$P_n(i) = \frac{e^{\mu(V_{in})}}{e^{\mu(V_{in})} + e^{\mu(V_n^*)}} = \frac{e^{\mu(V_{in})}}{e^{\mu(V_{in})} + e^{\mu(\frac{1}{\mu} \ln(\sum_{j \neq i} e^{\mu V_{jn}}))}} = \frac{e^{\mu(V_{in})}}{\sum_j e^{\mu(V_{jn})}} \quad (9.30)$$

La propriété 6 est utilisée dans la démonstration de la formulation du logit emboîté (section 2.5, p. 259) et justifie la formulation du logsum (équation 9.39, p. 261).

2.4.5. Propriété IIA d'indépendance vis-à-vis des autres alternatives

L'indépendance vis-à-vis des autres alternatives constitue la principale propriété du modèle logit. Elle s'exprime très simplement :

$$\frac{P\langle i | C_n \rangle}{P\langle j | C_n \rangle} = \frac{e^{V_{in}} / \sum_{k \in C_n} e^{V_{kn}}}{e^{V_{jn}} / \sum_{k \in C_n} e^{V_{kn}}} = \frac{e^{V_{in}}}{e^{V_{jn}}} = e^{V_{in} - V_{jn}} \quad (9.31)$$

Le rapport de probabilité de choix de deux alternatives ne dépend que de l'utilité des deux alternatives en question et est donc une constante (tant que les paramètres de l'utilité restent inchangés). En d'autres termes, on peut ajouter ou enlever une alternative, le rapport de probabilité des autres alternatives restera inchangé. Cette propriété est évidemment bien utile face à nouveau mode de transport. Il suffit de spécifier l'utilité de ce nouveau mode pour en déterminer la probabilité de choix, sachant que l'utilité des autres modes existants n'a pas à être redéfinie. Cette propriété n'est pas vérifiée pour d'autres modèles comme le modèle probit.

Il est important de comprendre que cette propriété est établie au niveau de l'individu et non pas au niveau de la population tout entière (sauf cas particulier). Prenons l'exemple cité par Ben-Akiva et Lerman (1985) pour l'expliquer. Soit une population constituée de deux groupes de tailles égales. Au sein de chacun des deux groupes, les individus ont des caractéristiques identiques au regard des variables prises en compte dans la fonction d'utilité déterministe. En revanche, ces caractéristiques sont différentes entre les deux groupes. Cette population dispose du choix entre un centre commercial périphérique et l'offre du centre ville. Les probabilités de choix de chacun des deux groupes sont reportées dans le tableau 9.1. Supposons qu'un nouveau centre commercial s'ouvre en périphérie dont les attributs (présents dans la partie déterministe de l'utilité) sont rigoureusement identiques à ceux du premier centre périphérique pour chacun des individus. L'application de la propriété IIA sur chacun des deux groupes d'individus conduit à conserver un rapport de probabilité identique entre le centre commercial périphérique et les commerces du centre ville. On obtient les probabilités qui figurent dans le tableau 9.1. Par contre pour l'ensemble de la population le rapport de probabilité n'est plus identique (tableau 9.1). La propriété IIA n'est pas respectée dans l'ensemble de la population (même si elle le reste pour chacun des deux groupes) car cette population n'est pas homogène contrairement à ce

que nous avons implicitement supposé dans l'illustration du paradoxe du bus bleu et du bus rouge (chapitre 8, p. 233).

De ce fait, si la fonction d'utilité ne fait pas intervenir de caractéristiques individuelles, la propriété IIA sera vérifiée au niveau de chacun des individus, mais aussi au niveau de l'ensemble de la population.

	Probabilité de choix			Rapport de probabilité
	Commerces du centre ville	Centre périphérique 1	Centre périphérique 2	commerces du centre ville / centre périphérique 1
Avant la construction du nouveau centre périphérique				
Groupe 1	0,95	0,05		0,0526
Groupe 2	0,05	0,95		19,0000
Ensemble de la population	0,50	0,50		1
Après la construction du nouveau centre périphérique				
Groupe 1	0,9048	0,0476	0,0476	0,0526
Groupe 2	0,0256	0,4872	0,4872	19,0000
Ensemble de la population	0,4652	0,2674	0,2674	0,5748

Tableau 9.1 : Illustration des résultats de la propriété IIA pour une population non homogène (Source : Ben-Akiva, Lerman, 1985)

Si la propriété IIA est très utile, elle constitue aussi la faiblesse du logit multinomial. Car si les rapports de probabilité ne restent pas constants à l'échelle de l'ensemble de la population, ils le sont au niveau de chacun des individus. Le paradoxe du bus bleu et du bus rouge (chapitre 8, p. 233) peut donc être illustré au niveau de chacun des individus. La propriété découle en fait directement de l'hypothèse IID qui est souvent prise en défaut. C'est notamment le cas en présence d'alternatives partageant certains attributs en commun (comme dans le cas du bus bleu et du bus rouge). Pour faire face à la non-indépendance des résidus, on peut utiliser le modèle logit hiérarchique ou emboîté. En revanche, pour faire face au non-respect des autres hypothèses, il faut recourir à des formes de modèles beaucoup plus complexes (modèle probit, section 2.3, p. 251 ou modèle mixte multinomial, section 2.7, p. 268).

2.5. Modèle logit emboîté (Nested Logit (NL))

Le modèle logit emboîté (qui est aussi parfois appelé logit hiérarchique pour Hierarchical Logit en anglais (HL)) permet de traiter le non-respect de l'hypothèse d'indépendance des résidus entre les alternatives. En revanche, les deux autres hypothèses que nous avons évoquées pour le modèle logit doivent être respectées. Nous développons tout d'abord les hypothèses du logit emboîté (section 2.5.1). Pour présenter le modèle logit emboîté, nous utilisons l'exemple souvent repris (Ben-Akiva, Lerman, 1985 ; Ortuzar, Willumsen, 1994) du choix combiné mode-destination (section 2.5.2). Enfin, nous explicitons sa formulation dans le cas du choix du mode, lorsque deux modes partagent des dimensions communes dans la formulation de l'utilité (section 2.5.3).

2.5.1. Hypothèses du logit emboîté

Prenons l'exemple d'un modèle de choix combiné mode-destination. Il est évident que parmi l'ensemble des alternatives croisant les destinations possibles et les modes disponibles, il y a des alternatives qui ne respectent pas l'hypothèse d'indépendance des résidus. Toutes les alternatives partageant la même destination et un mode différent ont une partie des facteurs entrant dans la partie

aléatoire de l'utilité qui est commune. Il en est de même des alternatives partageant le même mode, mais des destinations différentes. Le modèle logit multinomial ne pourra donc pas être utilisé directement.

Si l'on fait l'hypothèse d'une expression additive de l'utilité ou tout au moins d'une additivité des trois parties suivantes de l'utilité : relative au mode, relative à la destination, et relative aux facteurs communs mode-destination, on écrira (en abandonnant l'indice n relatif à l'individu pour simplifier les notations) :

$$U(d, m) = V(d, m) + \varepsilon(d, m) = (V_d + V_m + V_{dm}) + (\varepsilon_d + \varepsilon_m + \varepsilon_{dm}) \quad (9.32)$$

V_d, V_m, V_{dm} correspond à la partie déterministe de l'utilité relative respectivement de manière spécifique à la destination, au mode et à une combinaison de la destination et du mode ;

$\varepsilon_d, \varepsilon_m, \varepsilon_{dm}$ correspond à la partie aléatoire de l'utilité relative respectivement de manière spécifique à la destination, au mode et à une combinaison de la destination et du mode.

Par cette notation, on suppose que les termes d'erreur $\varepsilon_d, \varepsilon_m, \varepsilon_{dm}$ sont indépendants entre eux.

Si l'on a de bonnes raisons de considérer que le terme ε_{dm} est nettement prépondérant face aux deux autres termes, on pourra utiliser le modèle logit joint (*Joint logit model*) (Ben-Akiva, Lerman, 1985). Pour ce modèle, on fait l'hypothèse d'indépendance des résidus (on considère donc que l'hypothèse est quasiment respectée lorsque les résidus $\varepsilon_d, \varepsilon_m$ sont négligeables face à ε_{dm}). En application de l'équation 9.19 (p. 254) avec $\mu = 1$, on obtient :

$$P(d, m) = \frac{e^{V_m + V_d + V_{dm}}}{\sum_{m', d' \in C} e^{V_{m'} + V_{d'} + V_{d'm'}} \quad (9.33)$$

Toutefois, en général, surtout dans le cas du choix combiné mode-destination, il est peu probable que l'on puisse faire cette hypothèse. En d'autres termes la covariance entre deux alternatives partageant le même mode ou la même destination n'est pas négligeable :

$$\text{cov}(U_{dm}, U_{dm'}) = \text{cov}(\varepsilon_d + \varepsilon_m + \varepsilon_{dm}, \varepsilon_d + \varepsilon_{m'} + \varepsilon_{dm'}) \quad (9.34)$$

Chacune des composantes des termes d'erreur étant supposée indépendante des autres, l'équation 9.34 se réduit à :

$$\text{cov}(U_{dm}, U_{dm'}) = \text{var}(\varepsilon_d) \quad (9.35)$$

De même, on obtient :

$$\text{cov}(U_{dm}, U_{d'm}) = \text{var}(\varepsilon_m) \quad (9.36)$$

Donc si les termes d'erreurs ε_d et ε_m ne sont plus tous les deux négligeables face à ε_{dm} , on voit bien que le modèle logit joint ne peut plus s'appliquer car il n'est pas possible de postuler l'indépendance entre les alternatives partageant un mode ou une destination commune.

Dans le cas le plus général, où les deux termes d'erreur ε_d et ε_m sont importants, il est nécessaire de recourir au modèle probit (section 2.3, p. 251) ou au modèle mixte multinomial (section 2.7, p. 268) qui possèdent la forme la plus générale pour la structure de la matrice de variance-covariance des termes d'erreur entre les alternatives.

Toutefois, si l'on estime que l'on est en mesure de postuler que l'un des deux termes d'erreur est négligeable face aux deux autres, on pourra utiliser le logit emboîté. Nous considérons par la suite

que c'est le terme relatif au mode qui est négligeable (on pourrait tout aussi bien faire le choix du terme relatif à la destination, puisque dans cette présentation les deux termes sont interchangeables. Dans une application pratique, en revanche, le choix doit pouvoir être fondé sur des considérations comportementales). L'expression de l'utilité devient (équation 9.32) :

$$U(d, m) = V(d, m) + \varepsilon(d, m) = (V_d + V_m + V_{dm}) + (\varepsilon_d + \varepsilon_{dm}) \quad (9.37)$$

Cette hypothèse doit encore être complétée par quelques autres hypothèses :

1. ε_d et ε_{dm} sont indépendants entre deux alternatives quelle que soit la destination d appartenant à l'ensemble des destinations D et quel que soit le mode m appartenant à l'ensemble des modes M ;

2. les termes ε_{dm} sont indépendants et identiquement distribués selon une loi de Gumbel avec un paramètre μ^m ;

3. ε_d est distribué de telle sorte que $\max_m (U(d, m))$ (maximum de $U(d, m)$ sur l'ensemble des modes disponibles pour la destination d) soit aussi distribué selon une loi de Gumbel avec un paramètre μ^d , ceci quelle que soit la destination d appartenant à l'ensemble des destinations D .

Ces hypothèses signifient que l'hypothèse IID n'est plus maintenue sur la totalité du terme d'erreur, mais elle l'est sur la composante du terme d'erreur commun au choix du mode et de la destination (propriété 2) ainsi que sur le maximum de U_{dm} (sur l'ensemble des modes) pour chacune des destinations (propriété 3). De plus, les deux composantes du terme d'erreur doivent être indépendantes entre elles. Enfin, il faut rappeler l'hypothèse sur le terme d'erreur relatif spécifiquement au mode : ε_m que l'on considère négligeable face aux deux autres termes d'erreur. Ces quatre conditions sont nécessaires pour les calculs suivants (on en trouvera une démonstration dans l'analyse de la démonstration du modèle logit multinomial faite par Ben-Akiva et Lerman (1985)) :

- la première condition permet de décomposer la probabilité de choix combiné mode-destination en un produit de la probabilité marginale et de la probabilité conditionnelle (équation 9.41) ;
- la seconde condition est nécessaire pour pouvoir appliquer la formulation logit pour le calcul de la probabilité conditionnelle (choix du mode pour une destination donnée équation 9.40), elle est également nécessaire au calcul de l'utilité composite (équation 9.39) ;
- la troisième condition est nécessaire pour pouvoir appliquer le logit pour le calcul de la probabilité marginale (choix de la destination relativement à un ensemble de mode équation 9.38) ;
- la dernière condition est nécessaire pour pouvoir écrire l'équation 9.37 et donc négliger par la suite le terme ε_m .

2.5.2. Formulation du logit emboîté

Sous ces conditions, la probabilité marginale du choix de la destination d s'écrit (on en trouvera une démonstration complète dans Ben-Akiva, Lerman (1985)) :

$$P(d) = \frac{e^{\mu^d (V_d + V'_d)}}{\sum_{d' \in C} e^{\mu^d (V_{d'} + V'_{d'})}} \quad (9.38)$$

$$\text{avec } V'_d = \frac{1}{\mu^m} \ln \left(\sum_{m' \in C} e^{\mu^{m'} (V_{m'} + V_{dm'})} \right) \quad (9.39)$$

m' représentant l'ensemble des modes disponibles pour aller à la destination d .
sachant que la condition $\mu^d \leq \mu^m$ doit être respectée (voir Ben-Akiva et Lerman, 1985).

Cette expression (équation 9.38) correspond à une application d'un logit multinomial de choix entre les destinations dans laquelle l'utilité est définie par une partie spécifique à la destination et en quelque sorte une utilité composite relative à l'ensemble des modes de transport disponibles pour cette destination d . Cette utilité composite dérive en fait directement des propriétés 4 et 6 de la loi de Gumbel (section 2.4.3, p. 255). Si on définit :

$U'_d = \max_{m' \in C} (V_{m'} + V_{dm'} + \varepsilon_{dm'})$, alors par application des propriétés 4 et 6, U'_d est distribué

selon une loi de Gumbel de paramètres $\left(\frac{1}{\mu^m} \ln \left(\sum_{m' \in C} e^{\mu^m * (V_{m'} + V_{dm'})} \right), \mu \right)$.

En utilisant la propriété 4, on peut écrire : $U'_d = V'_d + \varepsilon'_d$, avec ε'_d distribué selon une loi de Gumbel de paramètres $(0, \mu)$ et $V'_d = \frac{1}{\mu^m} \ln \left(\sum_{m' \in C} e^{\mu^m * (V_{m'} + V_{dm'})} \right)$

Dans la pratique, il n'est pas possible de déterminer à la fois μ^d et μ^m . On fixe donc un des deux paramètres à 1 (généralement celui qui est le plus bas dans l'arbre de décision : schéma 9.1, niveau 2).

La probabilité conditionnelle du choix du mode pour une destination fixée s'écrit :

$$P(m|d) = \frac{e^{\mu^m (V_{dm} + V_m)}}{\sum_{m' \in C} e^{\mu^m (V_{dm'} + V_{m'})}} \quad (9.40)$$

Elle correspond à une application d'un logit multinomial entre les modes de transport dans laquelle on a enlevé la partie commune de l'utilité (la destination) afin de respecter la condition IID.

On en déduit donc la probabilité du choix combiné destination-mode :

$$P(m, d) = P(d) * P(m|d) = \frac{e^{\mu^d (V_d + V'_d)} * e^{\mu^m (V_{dm} + V_m)}}{\left(\sum_{d' \in C} e^{\mu^d (V_{d'} + V'_{d'})} \right) * \left(\sum_{m' \in C} e^{\mu^m (V_{dm'} + V_{m'})} \right)} \quad (9.41)$$

avec $\mu^d \leq \mu^m$. Cette inégalité est logique. Le paramètre μ^m est le facteur d'échelle de la distribution des résidus (ε_{dm}) relatifs au choix du mode pour une destination d donnée dont la variance est : $\frac{\pi^2}{6\mu_m^2}$. Le paramètre μ^d est le facteur d'échelle de la distribution des résidus relatifs

au choix de la destination pour l'ensemble des modes ($\varepsilon_d + \varepsilon_{dm}$) dont la variance est : $\frac{\pi^2}{6\mu_d^2}$. La variance de la seconde distribution étant forcément plus grande que celle de la première distribution,

on retrouve l'inégalité. De ce fait, lorsque μ^d est proche de μ^m , c'est que le terme ε_d est négligeable au regard de ε_{dm} . La structure emboîtée ne se justifie plus et il est préférable de retenir un logit multinomial (si $\mu^d = \mu^m$, on peut vérifier que l'équation 9.41 se ramène à l'équation du logit multinomial, équation 9.19 (p. 254)).

Cette présentation du logit emboîté peut être généralisée à un cas comportant un croisement d'un nombre plus important d'options. L'inégalité des paramètres μ doit toujours être respectée selon l'ordre de l'équation 9.41 (le paramètre le plus petit est celui situé le plus haut dans l'arbre de décision : schéma 9.1, niveau 1).

2.5.3. Application du logit emboîté au choix d'alternatives partageant des dimensions communes

L'exemple que nous avons évoqué ne s'applique pas qu'à des situations dans lesquelles plusieurs dimensions du choix sont croisées. Le modèle logit emboîté pourra être utilisé dans le cas où l'hypothèse IID ne peut être conservée pour l'ensemble des options relativement à une même dimension. Prenons le cas du choix du mode avec trois modes : voiture, métro, bus (schéma 9.1). Si l'on a de bonnes raisons de penser que le bus et le métro partagent des facteurs communs non pris en compte dans la partie déterministe de l'utilité, on pourra procéder comme ci-dessus.

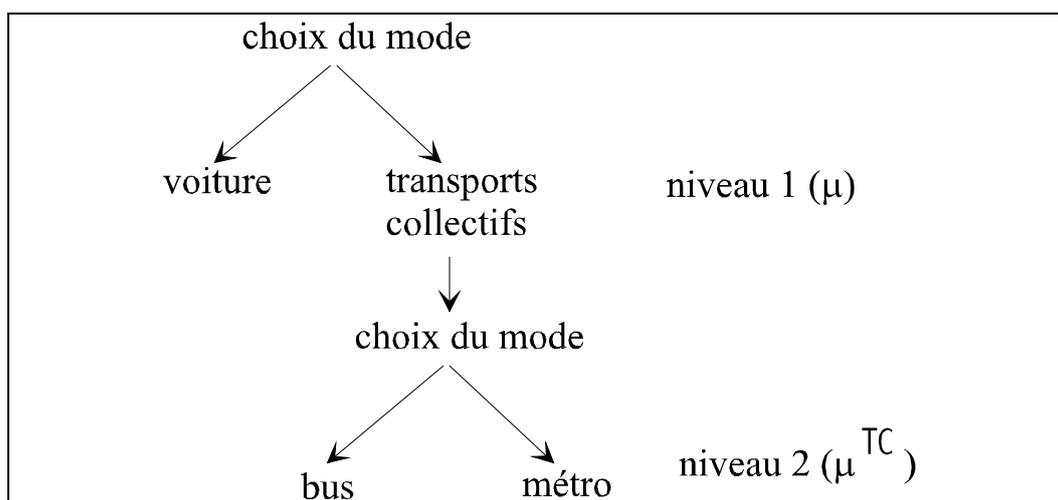


Schéma 9.1 : Structure de modèle logit emboîté

On considère dans un premier temps uniquement les deux modes corrélés bus et métro. Pour supprimer cette corrélation, on ne considère que la partie de l'utilité qui est spécifique à chacun des modes bus et métro de telle sorte que l'utilité de chacun des modes ne comporte pas de facteurs communs. On décompose donc l'utilité de la manière suivante :

$$U(\text{mode bus}) = V_{\text{bus}} + V_{\text{TC}} + \varepsilon_{\text{bus}} + \varepsilon_{\text{TC}}$$

$$U(\text{mode métro}) = V_{\text{métro}} + V_{\text{TC}} + \varepsilon_{\text{métro}} + \varepsilon_{\text{TC}} \quad (9.42)$$

où $U(\text{mode bus})$ représente la totalité de l'utilité du mode bus, de même pour le métro ;

V_{bus} représente la partie de l'utilité du mode bus qui est spécifique au mode bus, c'est-à-dire sans prendre en compte les facteurs communs au bus et au métro (elle correspond à la partie V_{dm} pour le mode bus de l'équation 9.37, p. 261) ;

$V_{\text{métro}}$ suit la même définition (elle correspond à la partie V_{dm} pour le mode métro de l'équation 9.37) ;

V_{TC} est la partie de l'utilité qui est commune aux deux modes de transports collectifs (elle correspond à la partie V_d de l'équation 9.37) ;

le terme V_m de l'équation 9.37 correspondrait à la partie de l'utilité spécifique au mode bus ou bien au mode métro relativement à l'ensemble des modes TC et VP, que l'on postule nulle ici tout comme le terme d'erreur correspondant ;

l'interprétation est identique pour les résidus.

Sous cette condition, on peut considérer que les termes d'erreur ε_{bus} et $\varepsilon_{m\acute{e}tro}$ sont indépendants et on fait l'hypothèse qu'ils sont identiquement distribués (condition 2 ci-dessus). On applique alors le logit binomial (équation 9.40, p. 262) :

$$P(\text{bus}|\text{TC}) = \frac{e^{\mu^{TC}(V_{bus})}}{e^{\mu^{TC}(V_{bus} + V_{m\acute{e}tro})}} \quad (9.43)$$

Dans un deuxième temps, on modélise le choix entre le mode TC et le mode voiture. On applique alors de nouveau un logit (équations 9.38 et 9.39, p. 261) :

$$P(\text{TC}) = \frac{e^{\mu(V_{TC} + V'_{TC})}}{e^{\mu(V_{TC} + V'_{TC})} + e^{\mu V_{VP}}} \quad (9.44)$$

avec V_{TC} qui comprend les facteurs de l'utilité communs aux modes bus et métro ;

$$\text{et } V'_{TC} = \frac{1}{\mu^{TC}} \ln \left(e^{\mu^{TC} V_{bus}} + e^{\mu^{TC} V_{m\acute{e}tro}} \right) \quad (9.45)$$

μ et μ^{TC} ne pouvant être conjointement déterminés. En général, on fixe le paramètre le plus bas dans l'arbre de décision à 1 (schéma 9.1, niveau 2, soit $\mu^{TC} = 1$). Dans ce cas, l'utilité du mode bus parmi l'ensemble des trois modes disponibles devient (équation 9.41, p. 262) :

$$P(\text{bus}) = P(\text{TC}) * P(\text{bus}|\text{TC}) = \frac{e^{\mu(V_{TC} + V'_{TC})} * e^{(V_{bus})}}{\left(e^{\mu(V_{TC} + V'_{TC})} + e^{\mu V_{VP}} \right) * \left(e^{(V_{bus} + V_{m\acute{e}tro})} \right)} \quad (9.46)$$

avec $0 < \mu \leq 1$ (pour respecter la décroissance du paramètre du logit lorsque l'on remonte dans l'arbre de décision).

Si $\mu = 1$ (et $\mu^{TC} = 1$), en remplaçant V'_{TC} par son expression dans 9.46, on retrouve le modèle logit multinomial. Dans ce cas, il n'est pas nécessaire d'utiliser une structure emboîtée ou hiérarchique qui comporte davantage de paramètres à estimer. Dans la pratique, quand μ est proche de 1, il est préférable d'abandonner la structure hiérarchique et de retenir un logit multinomial.

L'exemple avec ces trois modes étant un peu différent de celui comportant le choix combiné du mode et de la destination, il n'est probablement pas inutile de rappeler concrètement les quatre hypothèses nécessaires à l'application du logit emboîté :

1. la première condition s'écrit : ε_{bus} , ε_{TC} et $\varepsilon_{m\acute{e}tro}$ sont indépendants. Elle permet la décomposition de l'utilité des modes bus et métro en deux composantes et le calcul de la probabilité de chacun des deux modes collectifs en un produit de la probabilité marginale du mode TC et de la probabilité conditionnelle de chacun des deux modes collectifs :

$$P(\text{bus}) = P(\text{TC}) * P(\text{bus}|\text{TC}) \quad (\text{équation 9.46}) ;$$

2. la seconde condition s'écrit : les termes ε_{bus} et $\varepsilon_{\text{métro}}$ sont indépendants et identiquement distribués selon une loi de Gumbel de paramètre μ^{TC} , cette condition étant nécessaire pour obtenir les formulations des équations 9.43 et 9.45 ;

3. la troisième condition s'écrit : ε_{TC} est distribué de telle sorte que $U(\text{mode composite TC})$ soit aussi distribué selon une loi de Gumbel avec un paramètre μ .

Toutefois, dans cette traduction, le mode voiture a été oublié parmi l'ensemble des alternatives disponibles, car nous nous sommes concentrés sur les modes transports collectifs. Il faut donc ajouter :

1. pour la première condition le terme d'erreur de la voiture ε_{VP} ;

3. pour la troisième condition que ε_{VP} est aussi distribué selon une loi de Gumbel de paramètre μ .

Ces deux conditions sont nécessaires pour pouvoir écrire l'équation 9.44.

Enfin, on postule que le terme ε_m , comme dans l'équation 9.37, qui correspond au terme d'erreur spécifique au mode bus ou bien au mode métro relativement à l'ensemble des modes TC et VP est nul.

Ces conditions signifient que la propriété IIA d'indépendance vis-à-vis des autres alternatives n'est pas complètement abandonnée. La restriction (c'est-à-dire le maintien de la propriété IIA) est maintenue au niveau de l'ensemble des alternatives appartenant à une même branche de l'arbre de segmentation. Si l'on reprend la structure emboîtée présentée au schéma 9.1 (p. 263), cela signifie que la propriété IIA est maintenue entre les alternatives bus et métro d'une part et entre les alternatives voiture et transports collectifs d'autre part. En revanche, dans le modèle logit multinomial (structure horizontale du schéma 9.1), la propriété est valable entre les trois modes voiture, bus et métro.

Notre présentation du logit emboîté peut évidemment être généralisée à un plus grand nombre d'alternatives comportant plusieurs niveaux de hiérarchies et plusieurs modes au sein de chacune des branches. Cependant, l'inégalité des paramètres du logit doit toujours être vérifiée lorsque l'on remonte dans l'arbre de segmentation. De même, il est nécessaire de pouvoir séparer l'utilité de chaque alternative appartenant à un même niveau en termes communs et en termes spécifiques à chaque alternative et ceci pour chaque niveau hiérarchique. Enfin, les hypothèses doivent être respectées pour chacun des niveaux hiérarchiques.

Le modèle logit multinomial et le logit emboîté peuvent être inclus dans une famille de modèles plus large dénommée modèles de la valeur extrême généralisée.

2.6. Modèles de la valeur extrême généralisée (Generalised Extreme Value models, GEV)

Nous présentons tout d'abord le principe des modèles de la valeur extrême généralisée (section 2.6.1). Nous montrons ensuite que les modèles logit multinomiaux (section 2.6.2) et logit emboîtés (2.6.3) font partie de la famille des modèles de la valeur extrême généralisée. Enfin, nous indiquons d'autres modèles développés dans cette famille (2.6.4).

2.6.1. Principe, formalisation et démonstration

Le modèle de la valeur extrême généralisée a été formalisé par McFadden (1978) en le dérivant directement de la théorie de la maximisation de l'utilité aléatoire. Pour notre présentation, nous nous appuyons sur cet auteur ainsi que sur Ben-Akiva et Lerman (1985).

Posons $G(y_1, y_2, \dots, y_{J_n})$ une fonction des y_1, y_2, \dots, y_{J_n} tous ≥ 0 , qui possède les propriétés suivantes :

1. G est non négative ;
2. G est homogène de degré $\mu > 0$, c'est-à-dire $G(\alpha y_1, \alpha y_2, \dots, \alpha y_{J_n}) = \alpha^\mu G(y_1, y_2, \dots, y_{J_n})$;
3. $\lim_{y_i \rightarrow \infty} G(y_1, y_2, \dots, y_{J_n}) = \infty$, pour tout $i = 1, 2, \dots, J_n$;
4. la $k^{i\text{ème}}$ dérivée partielle de G relativement à toutes combinaisons de k y_i distincts, tel $i \in \{1, 2, \dots, J_n\}$ est non négative si k est impair et non positive si k est pair.

Si G satisfait ces conditions, en posant $y_i = e^{V_{in}}$ pour tout $i = 1, 2, \dots, J_n$, on obtient :

$$P\{i|C_n\} = \frac{e^{V_{in}} G_i(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}})}{\mu G(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}})} \tag{9.47}$$

avec $G_i(y_1, y_2, \dots, y_{J_n}) = \frac{\partial G}{\partial y_i}$ pour tout $i = 1, 2, \dots, J_n$.

L'équation 9.47 définit le modèle de la valeur extrême généralisée (*Generalised Extreme Value models, GEV*). McFadden (1978) a démontré que cette équation est cohérente avec théorie de la maximisation de l'utilité aléatoire. Dans sa démonstration McFadden suppose que G est homogène de degré 1.

Considérons la fonction F (McFadden (1978) a démontré que la fonction F est bien une fonction de répartition multivariée des termes d'erreurs $(\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{J_n n})$) :

$$F(\varepsilon_{1n}, \varepsilon_{2n}, \dots, \varepsilon_{J_n n}) = e^{\left\{ -G\left(e^{-\varepsilon_{1n}}, e^{-\varepsilon_{2n}}, \dots, e^{-\varepsilon_{J_n n}} \right) \right\}} \tag{9.48}$$

Reprenons l'équation 9.13 (p. 251). Nous supposons que la fonction F est identique entre les individus, ce qui explique l'abandon de l'indice n pour les fonctions F et G :

$$P\{1|C_n\} = \int_{\varepsilon_{1n} = -\infty}^{+\infty} F_1(\varepsilon_{1n}, V_{1n} - V_{2n} + \varepsilon_{1n}, \dots, V_{1n} - V_{J_n n} + \varepsilon_{1n}) d\varepsilon_{1n}$$

avec F_i la dérivée partielle de F par rapport au $i^{i\text{ème}}$ terme ε_{in} . En remplaçant F_1 par son expression obtenue en dérivant F par rapport à ε_{1n} , on obtient :

$$P\{1|C_n\} = \int_{\varepsilon_{1n} = -\infty}^{+\infty} e^{-\varepsilon_{1n}} G_1\left(e^{-\varepsilon_{1n}}, e^{-(V_{1n} - V_{2n} + \varepsilon_{1n})}, \dots, e^{-(V_{1n} - V_{J_n n} + \varepsilon_{1n})} \right) e^{\left\{ -G\left(e^{-\varepsilon_{1n}}, e^{-(V_{1n} - V_{2n} + \varepsilon_{1n})}, \dots, e^{-(V_{1n} - V_{J_n n} + \varepsilon_{1n})} \right) \right\}} d\varepsilon_{1n}$$

G étant une fonction homogène de degré 1, G_1 est donc une fonction homogène de degré $(1-1)=0$ (Ben-Akiva et François (1983) ont étendu la démonstration pour une fonction G homogène de degré μ), l'équation devient :

$$P\langle 1|C_n \rangle = \int_{\varepsilon_{1n} = -\infty}^{+\infty} e^{-\varepsilon_{1n}} G_1\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) \left\{ -\left(e^{-\varepsilon_{1n}}\right)\left(e^{-V_{1n}}\right) G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) \right\} d\varepsilon_{1n}$$

en posant $x = e^{-\varepsilon_{1n}}$, on a $dx = -e^{-\varepsilon_{1n}} d\varepsilon$
l'équation devient donc :

$$P\langle 1|C_n \rangle = -G_1\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) \int_{x=0}^{+\infty} e^{-\left(-V_{1n}\right)} G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) x \left\{ -\left(e^{-V_{1n}}\right) G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) \right\} dx$$

en faisant une nouvelle transformation $y = \left(e^{-V_{1n}}\right) G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) x$, on a
 $dy = \left(e^{-V_{1n}}\right) G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) dx$, l'équation devient :

$$P\langle 1|C_n \rangle = -\frac{G_1\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right)}{\left(e^{-V_{1n}}\right) G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right)} \int_{y=0}^{+\infty} e^{-y} dy$$

ce qui permet de retrouver l'équation 9.48 :

$$P\langle 1|C_n \rangle = \frac{e^{V_{1n}} G_1\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right)}{G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right)}$$

2.6.2. Dérivation du logit multinomial

En choisissant $G(y_1, y_2, \dots, y_{J_n}) = \sum_{j=1}^{j=J_n} y_j^\mu$, soit :

$$G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) = \sum_{j=1}^{j=J_n} e^{\mu V_{jn}} \quad (9.49)$$

l'équation 9.47 devient :

$$P\langle 1|C_n \rangle = \frac{\mu \left(e^{V_{1n}}\right)^{\mu-1} e^{V_{1n}}}{\sum_{j=1}^{j=J_n} e^{\mu V_{jn}}} = \frac{e^{\mu V_{1n}}}{\sum_{j=1}^{j=J_n} e^{\mu V_{jn}}}$$

ce qui permet de démontrer que le modèle logit multinomial appartient à la famille des modèles de la valeur extrême généralisée.

2.6.3. Dérivation du logit emboîté

De la même manière, on peut montrer que le logit emboîté appartient à la famille des modèles de la valeur extrême généralisée (McFadden, 1978 ; Ben-Akiva, Lerman, 1985). Dans ce cas, la fonction G prend l'une des deux formes suivantes :

$$G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) = \sum_{d=1}^D \left(\sum_{j \in M_d} e^{\mu^d V_{jn}} \right) \quad (9.50)$$

en segmentant en D ensembles disjoints, de taille M_d , l'ensemble des J_n alternatives. Dans ce cas, le paramètre d'échelle relatif à chacun des choix de plus bas niveau (choix des alternatives au sein de M_d pour chacun des ensembles $d \in \{1, \dots, D\}$) = 1 et $0 \leq \mu \leq 1$.

$$G\left(e^{V_{1n}}, e^{V_{2n}}, \dots, e^{V_{J_n n}}\right) = \sum_{d=1}^D \left(\sum_{j \in M_d} e^{\mu^m V_{jn}} \right)^{\mu^d / \mu^m} \quad (9.51)$$

avec les mêmes notations que ci-dessus. G est homogène de degré μ^d et $0 \leq \mu^d \leq \mu^m$.

2.6.4. Famille des modèles de la valeur extrême généralisée

D'autres modèles ont également été développés au sein de la famille des modèles de la valeur extrême généralisée (McFadden, 1978 ; Viauroux, 2001 ; Bhat, 1997 ; Koppelman, Sethi, 2000 ; Wen, Koppelman, 2001). Il s'agit principalement des *Ordered GEV* (OGEV), *Multinomial Logit-Ordered GEV* (MNL-OGEV), *Paired Combinatorial Logit* (PCL), *Cross-Nested Logit* (CNL), *Generalized Nested Logit* (GNL)... qui sont décrits dans les trois dernières citations ci-dessus. Deux éléments principaux distinguent ces différents modèles. Le premier concerne la définition de l'ensemble des alternatives de choix et l'arrangement de ces alternatives dans des classes à l'intérieur desquelles l'hypothèse d'une distribution indépendante et identique est généralement postulée (Swait, 2001). Ils se distinguent ensuite compte tenu de ces arrangements par la formalisation des élasticités croisées entre les alternatives qui sont plus ou moins flexibles selon les modèles.

Tous les modèles entrant dans la famille des modèles de la valeur extrême généralisée permettent de réduire voire supprimer l'hypothèse d'indépendance entre les résidus qui est nécessaire pour le modèle logit multinomial. En revanche, ils postulent tous l'identité de la distribution des résidus. Lorsque l'on ne peut raisonnablement postuler cette hypothèse, il faut se tourner vers d'autres formes de modèles que nous présentons dans la section suivante.

2.7. Développements récents : modèle logit multinomial mixte (*Mixed Multinomial Logit (MMNL)*)

Les hypothèses des deux modèles les plus couramment utilisés le logit multinomial et le logit emboîté sont assez fortes et ne supportent pas toujours la confrontation empirique. D'autres modèles sont alors nécessaires pour réduire ces hypothèses sur la structure des erreurs. Ils sont généralement organisés autour de leur capacité à réduire une ou plusieurs des trois hypothèses des modèles logit présentées précédemment (section 2.4.1, p. 253).

C'est la première hypothèse qui a reçu le plus d'attention car elle conduit à la propriété IIA qui pose le plus de problème. Deux directions ont été privilégiées. La première conduit à une réduction de la portée voire à l'abandon complet de l'hypothèse d'indépendance du terme d'erreur entre les alternatives. Elle débouche sur les modèles de la famille de la valeur extrême généralisée.

La seconde direction conduit à une réduction de la portée voire à l'abandon complet de l'hypothèse d'égalité de la variance du terme d'erreur entre les alternatives. Koppelman, Sethi (2000) et Bhat (1997), présentent plusieurs modèles répondant à cet objectif, et notamment le modèle *Heteroscedastic Multinomial Logit* (HMNL, logit multinomial hétéroscédastique).

La seconde hypothèse est relative à l'homogénéité des réponses aux différents attributs de la partie déterministe de l'utilité entre les individus. Plus précisément, le modèle n'autorise pas de variation de perception (*taste variation* en anglais) d'un attribut de la fonction d'utilité due à des caractéristiques non observées (c'est-à-dire non présentes dans la partie déterministe de l'utilité) de l'individu. Lorsque le problème est limité, une solution simple est évoquée plus loin (section 3.2.7, p. 284). Elle consiste à segmenter la population en sous groupes homogènes ou à introduire des variables spécifiques à certains groupes d'individus dans la fonction d'utilité. Lorsque le problème est plus important, une solution consiste à introduire des coefficients pouvant varier pour chacune des variables concernées. Dans ce cas, on conserve une seule fonction d'utilité. La variation des coefficients des variables incriminées peut être dépendante des caractéristiques socio-économiques des individus ou indépendante de ces caractéristiques. Dans ce dernier cas, il est alors nécessaire de faire une hypothèse sur la distribution des coefficients. Un tel modèle, appelé « *Random coefficient logit* » a été construit pour la première fois dans une analyse de la demande d'automobiles aux Etats-Unis par *Electric Power Research Institutes* (EPRI, 1977). Les distributions retenues étaient log-normales (Ben-Akiva, Lerman, 1985). Cette formulation a été reprise par d'autres auteurs (Bhat, 1997).

La troisième hypothèse concerne la structure de la matrice de variance-covariance du terme d'erreur entre les alternatives. On fait une hypothèse de structure identique entre les individus. Bhat (1997) propose différentes solutions pour traiter une variance ou bien une covariance non homogène entre les individus.

Toutefois, ces différents modèles ne permettent pas de traiter simultanément les différentes hypothèses. C'est la raison pour laquelle nous focalisons cette section sur les structures de modèles flexibles (de l'anglais *Flexible Structure Models* (Bhat, 2000)) qui permettent d'abandonner les deux premières hypothèses (IID et homogénéité de la réponse entre les individus (*taste variation*)). Les modèles probit répondent à cet objectif, mais leur estimation est extrêmement lourde dès que le nombre d'alternatives s'accroît. Nous nous centrons donc sur les modèles logit multinomiaux mixtes (MMNL). Pour cela, nous empruntons principalement à Bhat (2000).

Le principe de ces modèles est assez simple. En revanche, leur formulation l'est un peu moins et surtout leur estimation est assez complexe. Il est possible de présenter le modèle selon deux perspectives différentes, bien qu'elles soient équivalentes. La première dérive de la volonté de supprimer l'hypothèse d'indépendance et d'égalité de la variance du terme d'erreur de chacune des alternatives. Elle conduit à une formulation du terme d'erreur spécifique à chacune des alternatives dépendant des caractéristiques de cette alternative et de l'individu. La seconde permet d'introduire une hétérogénéité entre les individus dans leur perception des différents attributs de chacune des alternatives. Elle répond donc directement à la seconde hypothèse. Nous explicitons ces deux approches qui, malgré cette présentation, répondent bien aux deux hypothèses simultanément.

2.7.1. Structure spécifique du terme d'erreur

Les termes d'erreur de chacune des alternatives sont décomposés en deux parties. La première permet au terme d'erreur non observé de varier de manière non-identique et non-indépendante entre les alternatives. La seconde, en revanche, suit l'hypothèse IID et est distribuée selon une loi de Gumbel pour toutes les alternatives. En reprenant les notations de Bhat (2000), nous écrivons :

$$U_{ni} = \gamma Y_{ni} + \xi_{ni} = \gamma Y_{ni} + \mu Z_{ni} + \varepsilon_{ni} \quad (9.52)$$

Avec U_{ni} l'utilité de l'alternative i pour l'individu n ;

γY_{ni} la partie déterministe de l'utilité avec γ un vecteur de coefficients appliqué au vecteur de variables explicatives Y_{ni} ;

ξ_{ni} la partie aléatoire non observée de l'utilité qui est décomposée en deux termes :

- μZ_{ni} , avec Z_{ni} un vecteur de variables relatives à l'alternatives i (ces variables peuvent être communes ou non au vecteur Y_{ni}), et μ un vecteur aléatoire de coefficients de moyenne nulle ;
- ε_{ni} est le terme d'erreur avec une distribution IID Gumbel.

C'est le terme μZ_{ni} qui permet d'introduire l'hétéroscédasticité et la corrélation entre les termes d'erreurs de la partie aléatoire de l'utilité de chacune des alternatives. En définissant $\beta = (\gamma, \mu)$ et $X_{ni} = (Y_{ni}, Z_{ni})$, on obtient la structure du modèle logit multinomial mixte (MMNL) :

$$P_{ni}(\theta) = \int_{-\infty}^{+\infty} L_{ni}(\beta) f(\beta|\theta) d\beta \quad \text{où} \quad L_{ni}(\beta) = \frac{e^{\beta X_{ni}}}{\sum_j e^{\beta X_{nj}}} \quad (9.53)$$

Dans cette équation P_{ni} représente la probabilité que l'individu n choisisse l'alternative i ;

X_{ni} est un vecteur de variables explicatives observées spécifiques à l'alternative i et à l'individu n ;

β représente un vecteur de paramètres prenant des valeurs aléatoires suivant une distribution de densité $f(\cdot)$.

La spécification des variables composant le vecteur Z_{ni} est évidemment une tâche importante laissée à l'appréciation du modélisateur. Il est en de même du choix de chacune des distributions des paramètres composant le vecteur μ qui peuvent être spécifiques à chacune des variables et à chacune des alternatives dans les cas les plus complexes ou communes à certaines des alternatives (Bhat, 2000).

2.7.2. Distribution aléatoire des coefficients de l'utilité

Cette seconde présentation traduit l'hétérogénéité de la perception des attributs de chacune des alternatives entre les individus :

$$U_{ni} = \beta_n X_{ni} + \varepsilon_{ni} \quad (9.54)$$

X_{ni} est un vecteur de variables explicatives observées spécifiques à l'alternative i et à l'individu n ;

β_n représente un vecteur de paramètres qui varient entre les individus prenant des valeurs aléatoires suivant une distribution de densité $f(\cdot)$.

ε_{ni} est un terme d'erreur avec une distribution IID Gumbel.

Cette formulation de l'utilité conduit à la formulation de la probabilité de choix exposée sous l'équation 9.53. Selon Bhat (2000), la fonction de densité la plus couramment utilisée est une fonction de densité normale pour $f(\cdot)$. La fonction log-normale est également utilisée, notamment pour garantir un signe constant pour le coefficient.

Bhat (2000, 2001) propose différentes techniques d'estimation de ces modèles nécessitant toutes des procédures assez lourdes et complexes. Nous préférons ne pas les aborder ici en nous limitant aux structures de modèles les plus simples.

3. ESTIMATION DU MODELE DESAGREGE

A un niveau désagrégé, la valeur observée du choix prend systématiquement la valeur 0 ou 1. L'estimation statistique nécessite donc l'utilisation de la méthode du maximum de vraisemblance. Nous décrivons dans un premier temps cette méthode (section 3.1). Nous présentons ensuite différents

tests statistiques permettant d'estimer la qualité de l'estimation par le modèle de la réalité, ou plus exactement de la représentation de la réalité à travers les données d'enquête (section 3.2). Nous ferons dans la mesure du possible des analogies avec la présentation du modèle de régression linéaire (chapitre 6, p. 174) l'esprit des tests étant similaire, même si les méthodes de calcul diffèrent.

3.1. Maximum de vraisemblance

Comme pour le modèle linéaire, nous empruntons assez largement à Wonnacott et Wonnacott (1991) dont le mérite est de présenter très clairement les méthodes statistiques à des lecteurs non spécialistes sans pour autant abandonner le formalisme nécessaire.

Le principe de l'estimation par le maximum de vraisemblance est assez simple. Il consiste à trouver pour un modèle donné, la valeur de ses paramètres qui a le plus de chances de conduire à la situation observée (c'est-à-dire aux résultats des enquêtes). Prenons un exemple très simple emprunté à Wonnacott, Wonnacott (1991) pour illustrer le principe. « *Supposons que, pour contrôler la qualité d'un lot de radios, on y prélève un échantillon de 5, et que 3 d'entre elles soient défectueuses. C'est la proportion π de radios défectueuses qu'il faut estimer dans l'ensemble du lot. ... considérons toute une série de π possibles, ... puis essayons d'en tirer celle qui explique le mieux l'échantillon.*

Par exemple, $\pi = 0,1$ est-elle une valeur plausible pour la population ? Si $\pi = 0,1$, alors la probabilité de tirer $S = 3$ défectueuses dans un échantillon de $n = 5$ observations, serait donnée par la formule de la binomiale [que l'on peut obtenir sous un tableur ou par la formule suivante] :

$$\binom{n}{S} \pi^S (1-\pi)^{n-S} = 0,0081 \quad \text{»} \quad (9.55)$$

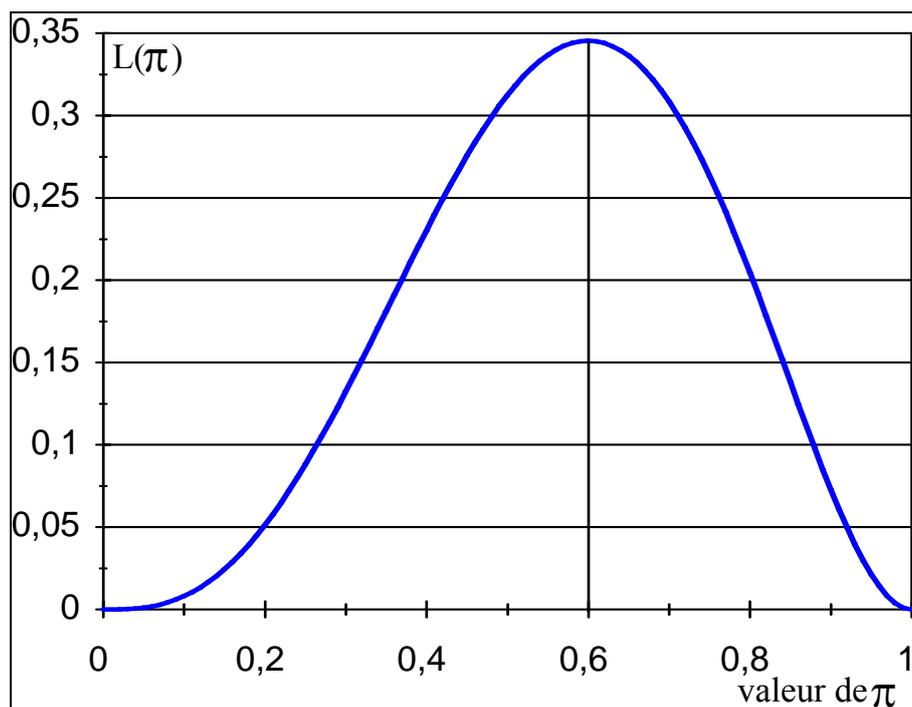
où $\binom{n}{S}$ est le nombre de combinaisons de S dans n , soit 10 dans notre exemple. Cette fonction est appelée la fonction de vraisemblance. Elle est notée $L(\pi)$. Il faut donc trouver le maximum de cette fonction pour trouver la valeur de π ayant la plus grande vraisemblance d'être trouvée.

Pour cet exemple, on peut représenter graphiquement la fonction de vraisemblance (graphique 9.6). On constate que la solution, qui correspond à la valeur de π qui maximise la fonction de vraisemblance est $\pi = 0,6$. On peut vérifier que cette valeur correspond au résultat intuitif de $3/5 = 0,6$.

Dans la pratique, on ne recherche pas la valeur qui maximise la fonction de vraisemblance à partir d'une représentation graphique. On recherche le maximum en annulant la dérivée de la fonction de vraisemblance :

$$\frac{\partial L(\pi)}{\partial \pi} = 0 \quad (9.56)$$

On pourra vérifier qu'en dérivant l'équation 9.55, on débouche effectivement sur la solution $\pi = 0,6$.



Graphique 9.6 : Illustration de la fonction de vraisemblance

Cet exemple illustre le principe de la méthode du maximum de vraisemblance que l'on applique à l'estimation des modèles de choix discrets. Dans ce cas, on dispose d'un échantillon d'observations (par exemple une enquête ménages déplacements) comportant N observations. Pour chacune de ces observations, on dispose d'un certain nombre de variables susceptibles d'expliquer le comportement de choix de l'individu et le choix que l'individu a effectué. Appelons f le modèle explicatif du choix Y_n en fonction de certaines variables X_n que nous considérons comme explicatives du comportement. Ce modèle prend la forme générale :

$$f(Y_n|X_n, \beta) \quad (9.57)$$

où β est un vecteur de paramètres appliqué au vecteur de variables explicatives X_n .

La méthode du maximum de vraisemblance consiste à trouver la valeur du vecteur β qui maximise la probabilité que le modèle reproduise la valeur observée des Y_n , compte tenu des valeurs X_n pour chacune des observations n . Pour cela, on pose la fonction de vraisemblance qui s'écrit :

$$L(\beta) = \prod_{n=1}^N f(Y_n|X_n, \beta) \quad (9.58)$$

Pour que le maximum de cette fonction corresponde à la solution optimale, il faut respecter l'hypothèse d'indépendance entre les observations obtenues par tirage aléatoire au sein de l'ensemble de la population. On retrouve une hypothèse très fréquente des modèles statistiques (cette hypothèse est également présente dans le modèle linéaire chapitre 6, p. 174) qui renvoie aux méthodes de production de données (chapitre 4).

La fonction de vraisemblance étant de forme produit, il est généralement plus facile de maximiser le logarithme de cette fonction, le maximum des deux fonctions étant atteint pour la même valeur de β :

$$\max_{\beta} L(\beta) = \max_{\beta} \ln L(\beta) = \max_{\beta} \sum_{n=1}^N f(Y_n | X_n, \beta) \quad (9.59)$$

La fonction de vraisemblance étant en général continue selon β , il est possible de la dériver et de l'annuler pour en trouver le maximum. Supposons que l'on ait K variables explicatives (noté N_k pour indiquer la dépendance à l'échantillon N) dans le vecteur X_n (en incluant la constante), on a également K paramètres à estimer dans le vecteur β , notés β_{N_k} . Le système d'équations à résoudre devient :

$$\frac{\partial \ln L(\beta)}{\partial \beta_{N_k}} = 0, \text{ pour } k = 1, \dots, K \quad (9.60)$$

Afin de s'assurer que le maximum obtenu, ne soit pas un maximum local, mais bien un maximum global, il faut que la dérivée seconde de la fonction de vraisemblance par rapport au vecteur β soit négative quel que soit β . La négativité de la dérivée seconde assure la concavité de la courbe ce qui permet de prouver que l'on est bien en présence d'un maximum global.

Les équations 9.57 à 9.60 sont fréquemment présentées d'une autre manière. La fonction f est remplacée par la fonction de probabilité de choix (modèle logit par exemple). L'équation 9.57 devient :

$$P_n(Y_{in} | X_{in}, \beta) \quad (9.61)$$

Définissons la variable suivante :

$$g_{ni} \begin{cases} 1 \text{ si l'individu } n \text{ choisit l'alternative } i \\ 0 \text{ sin on} \end{cases}$$

La fonction de vraisemblance devient (par application de l'équation 9.58) :

$$L(\beta) = \prod_{n=1}^N \prod_{i \in C_n} (P_n(i|\beta))^{g_{ni}} \quad (9.62)$$

On peut vérifier que le maximum de cette fonction correspond bien à la solution observée. Si le modèle $(P_n(i|\beta))$ fournit systématiquement une probabilité qui tend vers 1 lorsque l'alternative i est choisie et donc qui tend vers 0 pour les autres alternatives, alors $L(\beta)$ tend vers 1. Inversement, si le modèle donne systématiquement une probabilité qui tend vers 0 pour l'alternative choisie alors $L(\beta)$ tend vers 0. On notera qu'il est nécessaire d'avoir un modèle qui ne produise pas de probabilité = 0 (mais seulement qui tende vers 0). Dans ce cas, la fonction de vraisemblance serait nulle si une seule valeur de la probabilité était nulle, même si toutes les autres correspondent à l'observation. Toutefois, la nécessité pour la fonction de vraisemblance d'être dérivable selon β interdit en fait cette possibilité. Entre ces deux extrêmes, le terme $(P_n(i|\beta))^{g_{ni}}$ vaut 1 lorsque l'alternative i n'est pas choisie par l'individu n quelle que soit la valeur de $(P_n(i|\beta))$, et vaut $(P_n(i|\beta))$ si l'alternative i est choisie par l'individu n .

En prenant le logarithme de cette fonction (équation 9.62) comme dans l'équation 9.59, on obtient :

$$l(\beta) = \ln(L(\beta)) = \sum_{n=1}^N \sum_{i \in C_n} g_{ni} \ln[P_n(i|\beta)] \quad (9.63)$$

que l'on dérive comme dans l'équation 9.60 pour trouver la valeur de β qui maximise la fonction de vraisemblance. $L(\beta)$ variant entre 0 et 1, $l(\beta)$ varie entre $-\infty$ et 0. La valeur de $l(\beta)$ est donc toujours négative.

On peut montrer que sous les hypothèses du modèle linéaire, l'estimation par la méthode du maximum de vraisemblance est identique à la méthode de régression par les moindres carrés que nous avons présentée chapitre 6 (p. 174) (Wonnacott, Wonnacott, 1991).

Cette méthode, sous des hypothèses assez générales, possède des propriétés intéressantes. « *En particulier une estimation par cette méthode possède les propriétés (asymptotiques) des grands échantillons suivants : elle est :*

- non biaisée ;
- efficace avec la variance plus faible que tout autre estimateur non biaisé. Ce n'est vrai cependant que si la population-mère présente une forme spécifique – normale, uniforme ou autre – supposée connue quand on dérive le maximum de vraisemblance. Ainsi l'estimation peut manquer de robustesse dans le traitement d'une population dont la forme est inconnue ;
- de distribution d'échantillonnage normale, avec une moyenne et une variance calculées facilement. » (Wonnacott, Wonnacott, 1991)

Il est donc nécessaire de disposer d'effectifs suffisants pour bénéficier de ces propriétés. Compte tenu de ces propriétés, il est possible de dériver une série de tests présentant une forte analogie avec les tests développés pour la régression linéaire.

3.2. Tests statistiques

Nous ne limitons pas notre présentation aux tests statistiques, mais nous présentons l'ensemble des tests nécessaires pour examiner la qualité du modèle. Comme pour la régression, un des premiers examens concerne le signe des coefficients et l'ordre de grandeur des coefficients d'où l'on peut dériver des élasticités. Le calcul des élasticités n'étant pas aussi immédiat que dans le cas du modèle linéaire, nous nous y attardons quelque peu (section 3.2.1).

Comme pour la régression, nous examinons ensuite le test du t de student ou plus exactement dans le cas du maximum de vraisemblance le quasi-t test qui permet de tester la nullité des coefficients (section 3.2.2). Le t de student permet ensuite le calcul de l'intervalle de confiance (section 3.2.3). Nous présentons ensuite l'indicateur ρ qui correspond d'une certaine manière au R de la régression, même si les valeurs sont généralement plus faibles (section 3.2.4). Nous avons enfin la statistique du taux de vraisemblance qui permet de comparer différentes structures de modèles (section 3.2.5).

Pour conclure cette section, nous examinons des tests relatifs aux hypothèses nécessaires pour appliquer le modèle logit. Nous examinons deux de ces hypothèses. La première concerne l'indépendance des résidus qui conduit à la propriété IIA en raison de l'importance de cette propriété pour le choix de la famille de modèles et la structure de choix (section 3.2.6). La seconde porte sur la variation de la perception de certains attributs entre les individus (« *taste variation* ») (section 3.2.7).

3.2.1. Signe des coefficients et élasticités

Comme pour la régression, le premier examen concerne le signe des coefficients. Il est indispensable que le signe soit conforme à la logique économique. Prenons un exemple pour illustrer le propos. Dans un modèle de choix modal, l'équipement automobile est généralement une variable

importante. Il est donc généralement pris en compte dans la spécification de la fonction d'utilité du mode voiture. Dans ce cas, on est en droit d'attendre une valeur positive pour le coefficient correspondant. Plus un individu est motorisé, plus l'utilité du mode voiture a tendance à s'accroître.

Dans cet exemple, l'incidence de la variable explicative est évidente. Il arrive parfois que les effets des différentes variables soient plus complexes à identifier. Plusieurs cas peuvent alors se présenter. Nous les examinons dans la section suivante (tableau 9.4, p. 279). Cependant, l'analyste doit normalement être en mesure d'identifier la nature de la contribution de chacune des variables prises en compte dans le modèle. En effet, la réflexion sur le modèle théorique doit normalement précéder la spécification, puis l'estimation du modèle. Toutefois, dans la pratique, il est fréquent que le processus de construction du modèle soit constitué d'allers-retours entre la réflexion théorique, la spécification et le test d'hypothèses à l'aide de l'estimation de modèles alternatifs. Mais dans ce cas, la validation et le choix du modèle ne doit pas s'appuyer uniquement sur les résultats des tests statistiques. Il doit toujours s'accompagner d'une validation sur le plan théorique du mécanisme de réponse comportementale. Sinon, le risque est grand de construire un modèle qui reproduise bien le jeu de données dont on dispose, mais qui n'ait qu'un très faible pouvoir prédictif, c'est-à-dire un modèle non opérationnel faute d'une pertinence avérée (on retrouve la problématique du modèle, chapitre 2, p. 48).

Au-delà du signe du coefficient, l'ordre de grandeur des coefficients est également important. Plusieurs analyses sont possibles. La première concerne la comparaison de la valeur des différents paramètres au sein de la fonction d'utilité. Prenons de nouveau un exemple. Considérons une décomposition du temps de déplacement isolant pour les transports collectifs : le temps de marche à pied d'accès à l'arrêt, le temps d'attente du bus, puis le temps de circulation du bus, avec l'estimation d'un paramètre spécifique pour chacun de ces temps. On sait très bien que les usagers perçoivent les temps de marche et d'attente de manière beaucoup plus pénible que les temps en véhicule (SEMALY, 2000). Dans la fonction d'utilité du mode bus, on devra d'une part avoir des coefficients négatifs pour chacune de ces variables, mais de plus, la valeur des coefficients pour les temps de marche et d'attente devra être plus grande en valeur absolue que celle pour le temps en véhicule.

Dans la comparaison de la valeur des paramètres entre différents modèles, il ne faut pas oublier le facteur d'échelle qui peut être présent (sections 2.4.2, p. 254 et 2.4.3, p. 255 et notamment les valeurs retenues pour les graphiques 9.3 et 9.4 de telle sorte que les écart-types des distributions du terme d'erreur $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ soient identiques).

Souvent plutôt que de travailler directement sur la valeur du paramètre, il est préférable de travailler sur l'élasticité qui traduit mieux l'incidence de la variable sur l'objet d'étude et facilite la comparaison avec les résultats obtenus dans le cadre d'autres études. Le calcul de l'élasticité n'est toutefois pas aussi immédiat que dans le cas du modèle linéaire. Nous explicitons donc le principe du calcul et la formulation dans quelques cas simples.

L'élasticité est une grandeur très utile car elle permet de relativiser l'incidence de chacune des variables sur l'objet d'étude. Son principe est très simple, il consiste à estimer l'incidence de l'évolution d'une variable explicative de x % sur la variable à expliquer (mesurée aussi en pourcentage = y %). L'élasticité se formalise donc très simplement :

$$\frac{\Delta Y}{Y} = e \frac{\Delta X}{X} \quad (9.64)$$

où X est la variable explicative et Y la variable à expliquer ;
 e est l'élasticité de Y par rapport à X , qui est parfois notée $E_x(Y)$.

En fait, la définition n'est pas aussi simple. En effet, si la relation entre Y et X n'est pas de forme puissance, la valeur de l'élasticité ne sera pas la même selon la valeur initiale de X . De même, la

valeur ne sera pas la même pour une variation de X de 1 % ou pour une variation de 10 %. On définit donc en fait deux élasticités :

- élasticité point (*point elasticity* en anglais) :

$$\frac{\partial Y}{Y} = E_x(Y) \frac{\partial X}{X} \quad (9.65)$$

- élasticité arc (*arc elasticity* en anglais) :

$$\frac{\Delta Y}{Y} = E_{\bar{x}}(Y) \frac{\Delta X}{X} \quad (9.66)$$

Il est évident que quand ΔX tend vers 0, les deux définitions sont identiques.

On définit aussi deux types d'élasticités. L'élasticité directe qui correspond à la définition ci-dessus et l'élasticité croisée. Prenons l'exemple du choix du mode. L'élasticité de la probabilité du choix du mode voiture relativement à la variation d'une variable présente dans l'utilité du mode voiture sera appelée élasticité directe. Et l'élasticité de la probabilité du choix du mode voiture relativement à la variation d'une variable présente dans l'utilité du mode transports collectifs sera appelée élasticité croisée. Soit i et j deux modes, les définitions deviennent :

$$\text{élasticité directe : } \frac{\partial Y_i}{Y_i} = E_{x_i}(Y_i) \frac{\partial X_i}{X_i} ; \quad (9.67)$$

$$\text{élasticité croisée : } \frac{\partial Y_i}{Y_i} = E_{x_j}(Y_i) \frac{\partial X_j}{X_j} ; \quad (9.68)$$

Nous donnons quelques exemples d'élasticités directes pour des fonctions $f(Y = f(X))$ simples (tableau 9.2). Le calcul de l'élasticité nécessitant le calcul de la dérivée, nous rappelons le calcul de la dérivée (tableau 9.3) également pour des fonctions simples (Manheim, 1984 ; Ortuzar, Willumsen, 1994).

Le calcul des élasticités et des élasticités croisées permet d'examiner la sensibilité du modèle à la variation des variables explicatives du modèle. Pour certaines variables explicatives, on connaît l'intervalle généralement observé pour l'élasticité. Il est alors utile de comparer la valeur obtenue dans le modèle avec cet intervalle. Toutefois, il ne s'agit pas d'une règle absolue, car l'élasticité dépend de la valeur de la variable explicative comme on peut l'observer dans le tableau 9.2. Elle est aussi obtenue pour une certaine valeur de l'ensemble des autres paramètres. Elle est donc relative à un contexte donné. Sorti de ce contexte donné, il faut être très prudent sur la transférabilité de la valeur obtenue. Cette analyse constitue donc plutôt une aide à la compréhension du fonctionnement du modèle (ou plus précisément à la réponse du modèle face à une certaine modification des variables d'entrée toutes choses égales par ailleurs). Il faut alors s'assurer que le sens des évolutions et leur ordre de grandeur sont conformes à la logique économique donc au modèle théorique construit (schéma 2.1, p. 34).

	forme fonctionnelle de la fonction $y = f(x)$	dérivée de f par rapport à x : $\frac{\partial y}{\partial x}$	élasticité de y par rapport à x : $\frac{\partial y}{\partial x} \frac{y}{x}$
linéaire	$y = ax + b$	a	$a \frac{y}{x} = \frac{1}{1 + b/ax}$
puissance	ax^b	abx^{b-1}	b
exponentielle	ae^{bx}	abe^{bx}	bx

logistique	$\frac{a}{1 + ce^{bx}}$	$-by\left(1 - \frac{y}{a}\right) = \frac{-abce^{bx}}{\left(1 + ce^{bx}\right)^2}$	$-bx\left(1 - \frac{y}{a}\right) = -\frac{bcxe^{bx}}{1 + ce^{bx}}$
------------	-------------------------	---	--

Tableau 9.2 : Calcul de l'élasticité pour des fonctions courantes

	forme fonctionnelle de la fonction $y = f(x)$	dérivée de f par rapport à x : $\frac{\partial y}{\partial x}$
constante	a	0
produit par une constante	ax	a
puissance	x^b	bx^{b-1}
exponentielle	e^{bx}	be^{bx}
logarithme népérien	$\ln(ax)$	$\frac{1}{ax}$
fonction f(x)	af(x)	af'(x)
addition de fonctions	f(x) + g(x)	f'(x) + g'(x)
produit de fonctions	f(x) * g(x)	f'(x)g(x) + f(x)g'(x)
fonction de fonction	f(g(x))	f'(g(x))*g'(x)

Tableau 9.3 : Calcul de dérivée pour des fonctions courantes

A partir de la formule générale fournie dans le tableau 9.2, on peut exprimer l'élasticité directe et croisée du modèle logit multinomial sous une forme que l'on trouve habituellement dans la littérature (Ben-Akiva, Lerman, 1985, Koppelman, Sethi, 2000) :

$$\text{élasticité directe : } E_{x_{ink}}(P_n(i)) = (1 - P_n(i))X_{ink}\beta_{ik} ; \tag{9.69}$$

$$\text{élasticité croisée : } E_{x_{jnk}}(P_n(i)) = -P_n(j)X_{jnk}\beta_{jk} ; \tag{9.70}$$

avec $E_{x_{ink}}(P_n(i))$ l'élasticité directe de la probabilité $P_n(i)$ de choix de l'alternative i relative à X_{ink} , $k^{\text{ième}}$ variable de la partie déterministe de l'utilité de l'alternative i pour l'individu n ;

β_{ik} coefficient de la variable X_{ink} ;

β_{jk} coefficient de la variable X_{jnk} .

En analysant l'équation 9.70, on constate que l'élasticité croisée relative à la variable k est identique pour toutes les alternatives i différentes de j. Il s'agit d'une autre manière de présenter la propriété IIA d'indépendance vis-à-vis des autres alternatives.

Les deux formules ci-dessus (équations 9.69 et 9.70) ne sont toutefois valables que si la variable explicative X_{ink} ou X_{jnk} n'est pas elle-même la transformée d'une autre variable explicative Z_{ink} ou Z_{jnk} (comme dans les transformations de Box-Cox ou Box-Tukey, section 4.2.3, p. 289). Dans ce cas, les deux équations ci-dessus deviennent, en posant $X_{ink} = h^k(Z_{ink})$:

$$\text{élasticité directe : } E_{x_{ink}}(P_n(i)) = (1 - P_n(i))\frac{\partial h^k}{\partial z_{ink}}Z_{ink}\beta_{ik} ; \tag{9.71}$$

$$\text{élasticité croisée : } E_{x_{jnk}}(P_n(i)) = -P_n(j)\frac{\partial h^k}{\partial z_{jnk}}Z_{jnk}\beta_{jk} ; \tag{9.72}$$

Ces formulations peuvent être plus complexes dans certains cas (par exemple lorsque la variable explicative intervient dans l'utilité de plusieurs alternatives...). Il convient donc d'être prudent dans l'utilisation de telles formules (Ben-Akiva, Lerman, 1985). Koppelman et Sethi (2000) et (Wen, Koppelman, 2001) proposent différentes formulations des utilités pour des formes de modèles moins simples, comme les logit emboîtés ou certains modèles dérivant de la famille des modèles de la valeur extrême généralisée.

3.2.2. Quasi-t test

L'estimateur du maximum de vraisemblance a une distribution asymptotiquement normale (section 3.1, p. 271). Si la taille de l'échantillon est suffisante (selon Ortuzar et Willumsen (1994) un échantillon de 500 à 1000 observations est largement suffisant), il sera possible de faire des tests sur la significativité du coefficient. Le principe est identique à celui présenté pour la régression (chapitre 6, p. 176). Dans notre cas, l'estimateur ayant une distribution asymptotiquement normale (donc uniquement pour des effectifs suffisants), on parle de quasi-t test de student.

La valeur du t est fournie par les logiciels d'estimation ainsi que de plus en plus souvent la valeur de la probabilité critique. Elle permet de définir le seuil de probabilité auquel l'hypothèse H_0 « le coefficient β_{ik} de la variable X_{ik} est nul » ne peut plus être rejetée. Concrètement si la probabilité est supérieure à 5 % (qui est le seuil généralement retenu), le coefficient ne peut pas être considéré comme significativement différent de 0 et peut donc être rejeté car il n'est explicatif du processus de choix.

Dans la pratique, il arrive que l'on conserve un coefficient non significatif. Nous empruntons à Ortuzar et Willumsen (1994) le tableau 9.4 suivant qui propose une ligne de conduite quant à l'acceptation ou au rejet de la variable.

Il est évident qu'il ne s'agit là que de lignes directrices et qu'il convient de les adapter à chaque contexte d'étude. La procédure conduit à segmenter les variables entre les variables liées à la politique que l'on cherche à simuler, c'est-à-dire ce que l'on appelle en général les variables de commande, et les autres variables. Il est évident qu'il est nécessaire de disposer dans le modèle de ces variables de commande. Sinon c'est la cohérence d'objectifs (Cf. problématique des modèles, chapitre 2, p. 48) qui est prise en défaut. L'absence d'une variable de commande ne permet pas de tester la politique qui dépend de l'action sur cette variable. C'est ce qui explique que l'on soit souvent amené à conserver une telle variable même si la variable n'est pas significative au seuil de confiance choisi. Le problème est faible si le signe est correct, car l'absence de significativité peut être due à une taille d'échantillon insuffisante. Il devient, en revanche, plus important lorsque le signe n'est pas correct. Enfin, si le signe est incorrect et que la variable est significative, le problème ne peut plus être occulté. C'est le modèle qui doit être revu (dans ce cas les autres tests peuvent souvent aider à déterminer la nature du problème : mauvaise spécification de la fonction d'utilité, choix de la structure du modèle, hypothèses non respectées...).

signe du coefficient	test t de student	nature de la variable explicative	
		liée à la politique à simuler	autre
signe conforme à la logique économique	significatif (probabilité < 5 %)	acceptation	acceptation
	non significatif (probabilité > 5 %)	acceptation	peut être rejetée
signe non conforme à la	significatif (probabilité < 5 %)	problème important	rejet

logique économique	non significatif (probabilité > 5 %)	problème	rejet
--------------------	---	----------	-------

Tableau 9.4 : Procédure de sélection des variables

3.2.3. Intervalle de confiance sur les coefficients

Comme pour la régression linéaire (équation 6.7, p. 177), le t de student permet de calculer un intervalle de confiance (sous réserve d'un effectif suffisant). Dans ce cas, l'intervalle de confiance s'obtient à partir de l'expression de la probabilité suivante (Ben-Akiva, Lerman, 1985) :

$$\Pr \left[-t_{\alpha/2} \leq \frac{\hat{\beta}_k - \beta_k}{\sqrt{\text{var}(\hat{\beta}_k)}} \leq t_{\alpha/2} \right] = 1 - \alpha \quad (9.73)$$

où β_k est la valeur (inconnue) du paramètre k relatif à la variable X_k du modèle ;

$\hat{\beta}_k$ est la valeur estimée de β_k par le maximum de vraisemblance ;

$t_{\alpha/2}$ est la valeur du t de student (pour un nombre infini de degrés de liberté puisqu'il s'agit d'une propriété asymptotique valable pour de grands effectifs) pour un seuil de confiance $(1 - \alpha)$. α est le plus souvent pris égal à 5 % (parfois 10 %) ;

$\text{var}(\hat{\beta}_k)$ est la variance de $\hat{\beta}_k$.

On en déduit :

$$\Pr \left[\hat{\beta}_k - t_{\alpha/2} * \sqrt{\text{var}(\hat{\beta}_k)} \leq \beta_k \leq \hat{\beta}_k + t_{\alpha/2} * \sqrt{\text{var}(\hat{\beta}_k)} \right] = 1 - \alpha \quad (9.74)$$

3.2.4. Qualité de l'estimation : ρ^2

Dans la régression, le R^2 est une bonne mesure de la qualité de la régression en exprimant le pourcentage de la variance expliquée par la régression. La méthode du maximum de vraisemblance ne permet pas de calculer un indicateur identique. Il est, en revanche, possible de trouver un indicateur similaire dans l'esprit. Son principe de construction est très simple. Il s'agit de comparer le modèle obtenu avec le vecteur de coefficients maximisant la fonction de vraisemblance et le modèle dans lequel tous les coefficients sont nuls. Il s'agit du modèle le plus simple dans lequel aucune variable explicative n'est prise en compte. Il conduit à une équirépartition entre les alternatives. On définit ainsi l'index du taux de vraisemblance (rho-carré) :

$$\rho^2 = 1 - \frac{l(\hat{\beta})}{l(0)} \quad (9.75)$$

où $l(\hat{\beta})$ est la valeur du maximum du logarithme de la fonction de vraisemblance qui est obtenu pour le vecteur de paramètres $\hat{\beta}$;

$l(0)$ est la valeur du logarithme de la fonction de vraisemblance dans laquelle tous les coefficients du vecteur de paramètres sont nuls. Prenons un exemple pour l'illustrer. Prenons un échantillon de 1000 observations avec deux modes de transport. Par application de l'équation 9.63 (p. 274), on obtient :

$$l(0) = \sum_{n=1}^N \sum_{i \in C_n} g_{ni} \ln[P_n(i|0)] = 1000 \ln(0,5) = -693$$

car $N = 1000$, i prend la valeur 1 ou 2, g_{ni} vaut alternativement 0 ou 1 pour chaque valeur de n et la probabilité de choix est systématiquement égale à 0,5, puisque les utilités des deux modes sont égales et nulles dans le cas $\beta = 0$ (coefficients des variables explicatives et constantes nuls).

On vérifie que la valeur de ρ^2 varie effectivement entre 0 et 1 puisque par définition la valeur absolue de $l(\hat{\beta})$ est inférieure à la valeur absolue de $l(0)$ ($l(\hat{\beta})$ est supérieure à $l(0)$ puisque c'est le maximum de la fonction l , mais les deux valeurs sont négatives). Et la valeur maximum que puisse prendre $l(\hat{\beta})$ est 0 qui serait obtenue si la probabilité de choix était systématiquement = 1 lorsque l'alternative i est choisie.

Dans la pratique, la valeur de ρ^2 est souvent beaucoup plus faible que le R^2 .

On lui préfère souvent un autre indicateur (équation 9.76), car il est possible de proposer un modèle très simple qui peut conduire à des valeurs assez fortes de ρ^2 . C'est le modèle obtenu avec une simple constante (tous les autres coefficients relatifs aux variables explicatives sont toujours nuls). Ce modèle, dans le cas du logit, est celui qui reproduit en moyenne la part de marché observée pour chacun des modes. La valeur de la fonction de vraisemblance dans ce cas est notée $l(c)$. Dans le cas où la part de marché des deux modes est égale, elle conduit à une valeur $l(c) = l(0)$ car les deux modes se partagent le marché. En revanche, lorsque la part de marché de chacun des deux modes n'est pas identique, $l(c)$ peut être nettement supérieur à $l(0)$. Prenons un exemple toujours avec le même échantillon de 1000 observations. On a 300 observations qui utilisent la voiture et 700 les transports collectifs. En appliquant la formule de l'équation 9.63, on obtient :

$$l(c) = \sum_{n=1}^N \sum_{i \in C_n} g_{ni} \ln[P_n(i|c)] = 300 \ln(300/1000) + 700 \ln(700/1000) = -611$$

Dans ce cas, on obtient une valeur de :

$$\rho^2 = 1 - \frac{l(c)}{l(0)} = 1 - \frac{-611}{-693} = 0,119$$

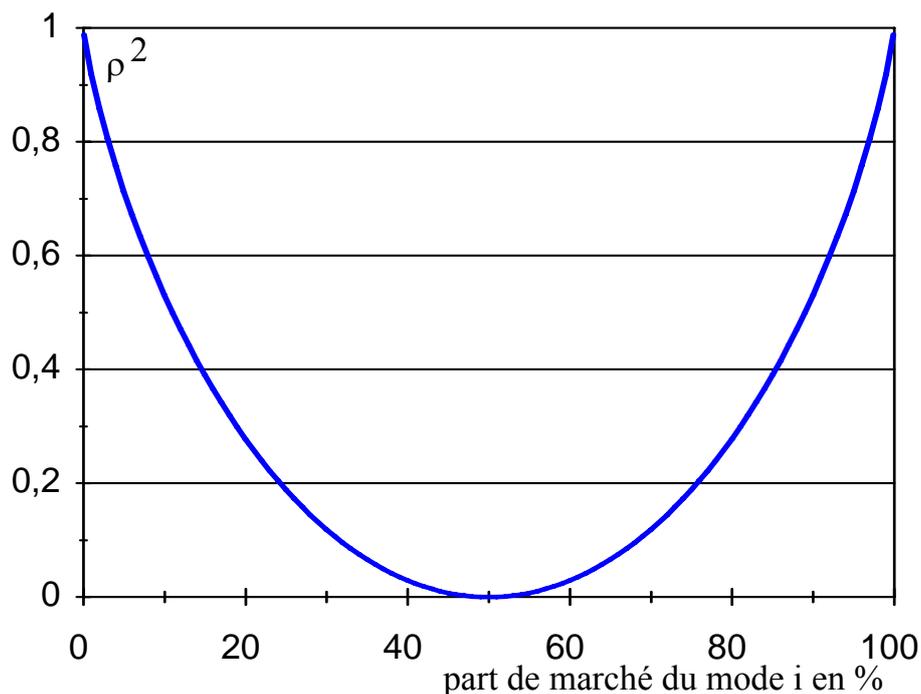
Il est possible de représenter graphiquement la valeur de $\rho^2 = 1 - \frac{l(c)}{l(0)}$ pour différentes parts de marché des deux modes. On obtient le graphique 9.7, qui fournit d'une certaine manière la valeur minimale de l'index (car elle correspond à un des modèles le plus simple dans lequel on n'a aucune variable explicative, seulement une constante).

Ce graphique ne dépend pas de la taille de l'échantillon initiale (il dépend par contre du nombre d'alternatives disponibles). La valeur que l'on obtiendra pour ρ^2 dépend donc de la répartition entre les modes. Une valeur de 0,3 pour une équirépartition des modes est sûrement meilleure qu'une valeur de 0,53 pour une répartition 10 %-90 % (valeur obtenue avec $l(c)$).

C'est pour faire face à ce problème que l'on propose souvent une autre formulation :

$$\rho^2 = 1 - \frac{l(\hat{\beta})}{l(c)} \tag{9.76}$$

qui permet de faire abstraction de la répartition observée dans l'échantillon. La valeur de cet index est toujours comprise entre 0 et 1, pour peu que l'on ait au moins une constante par mode dans la fonction d'utilité (ou plus exactement une constante par mode - 1 puisqu'il y a 1 indétermination entre l'ensemble des constantes modales).



Graphique 9.7 : Illustration de la valeur minimum de ρ^2 relativement au modèle avec constante seule pour le logit binomial

Comme pour le R^2 , cet index croît avec le nombre de variables incluses dans la fonction d'utilité. Comme pour le R^2 , on propose un ρ^2 ajusté obtenu par la formule suivante (Ben-Akiva, Lerman, 1985) :

$$\rho^2 = 1 - \frac{l(\hat{\beta}) - K}{l(c)} \quad (9.77)$$

où K est le nombre de variables de la fonction d'utilité en comptant les constantes.

Dans le cas du modèle logit emboîté, Ortuzar, Willumsen (1994) proposent d'utiliser la formulation suivante :

$$\rho^2 = 1 - \frac{l_1(\hat{\beta}) + l_2(\hat{\beta}) + \dots + l_s(\hat{\beta})}{l_1(c) + l_2(c) + \dots + l_s(c)} \quad (9.78)$$

où 1, 2 ... s représentent les différents niveaux du logit emboîté.

L'utilisation du ρ^2 ajusté permet donc de tester l'introduction de nouvelles variables dans la fonction d'utilité. Si sa valeur ne progresse pas avec de nouvelles variables, l'introduction de ces nouvelles variables ne permet pas d'apporter d'explications supplémentaires au comportement observé. Ces variables ne doivent donc pas être conservées. En revanche, si la valeur du ρ^2 ajusté s'accroît, on souhaite en général savoir si le gain « d'explication » est significatif ou non. On dispose pour cela du test présenté dans la section suivante.

3.2.5. Test du ratio de vraisemblance (Likelihood ratio Test ou LR)

Le ratio de vraisemblance correspond au F de Fisher de la régression. Comme pour le F, il permet de comparer deux modèles et de tester s'ils sont significativement différents.

Considérons deux modèles. Le premier modèle est celui pour lequel on veut faire un test. On l'appelle le modèle non restreint (c'est-à-dire de formulation générale). La valeur du logarithme de la fonction de vraisemblance est notée $l(\beta)$. Le second modèle est un modèle restrictif (par exemple avec des variables en moins ou avec l'égalité de certains coefficients...). La valeur du logarithme de la fonction de vraisemblance est notée $l(\beta_r)$. On peut montrer (Ben-Akiva, Lerman, 1985) que la statistique :

$$-2(l(\beta_r) - l(\beta)) \quad (9.79)$$

suit asymptotiquement une loi du χ^2 avec r degrés de liberté. r est le nombre de restrictions imposées sur le modèle restreint par rapport au modèle général. Si la valeur de la statistique est suffisamment grande (c'est-à-dire supérieure à la valeur du χ^2 avec r degrés de liberté au seuil de confiance choisi), il est possible de rejeter l'hypothèse « les deux modèles sont égaux ».

Par exemple, si l'on veut tester la nullité de tous les coefficients du modèle, la valeur du logarithme de la fonction de vraisemblance sera $l(0)$. La statistique devient :

$$-2(l(0) - l(\beta)) \quad (9.80)$$

avec K degrés de liberté (K étant le nombre de variables explicatives en comptant les constantes).

Cette statistique est généralement fournie par les logiciels d'estimation.

En général, on trouve aussi la statistique :

$$-2(l(c) - l(\beta)) \quad (9.81)$$

avec $(K - a)$ degrés de liberté où a est le nombre de constantes (en général le nombre d'alternatives $- 1$), qui permet de tester l'hypothèse « tous les coefficients sont nuls sauf les constantes modales ».

L'application de l'équation 9.79 permet en fait de faire beaucoup d'autres tests. En particulier, il est possible de tester l'égalité de coefficients. Si l'on a deux modes de transports collectifs, bus et métro, on pourra tester l'égalité des coefficients relatifs au temps de parcours ou à toutes autres variables communes aux deux modes. Il faut alors estimer le modèle sans restriction qui conduit à la valeur de $l(\beta)$. Il faut ensuite estimer le modèle dans lequel on introduit la restriction : égalité des deux coefficients relatifs au temps de parcours, qui conduit à la valeur de $l(\beta_r)$. On calcule ensuite la statistique : $-2(l(\beta_r) - l(\beta))$ qui possède 1 degré de liberté (puisque l'on a introduit une seule restriction).

On peut aussi tester une relation entre plusieurs coefficients....

Reprenons le cas cité à la fin de la section précédente. Il s'agit de l'ajout de nouvelles variables qui contribue à un accroissement du ρ^2 ajusté, mais pour lequel on souhaite savoir si le gain est statistiquement significatif ou non. Ce nouveau modèle n'est pas obtenu comme réduction du précédent mais par ajout. Il suffit donc d'inverser les rôles. On considère le nouveau modèle comme étant le modèle général et l'ancien modèle comme étant le modèle réduit. On calcule alors la statistique selon l'équation 9.79, le nombre de degrés de liberté étant le nombre de variables ajoutées (les restrictions étant formulées pour chacune des variables ajoutées : le coefficient de la variable = 0).

Ces tests imposent toutefois d'estimer deux modèles. Un premier modèle sans restriction, puis le second avec l'hypothèse que l'on souhaite tester.

Ce test permet ainsi de procéder à l'estimation du meilleur modèle pas à pas, en retenant ou excluant telle ou telle variable, en commençant en général par une formulation simple avec peu de variables.

3.2.6. Test de la propriété IIA

La propriété IIA d'indépendance vis-à-vis des autres alternatives est un des atouts du logit par la simplicité de formulation du modèle, elle constitue aussi sa faiblesse. Il est donc nécessaire de pouvoir tester la propriété lorsque l'on a des doutes sur sa pertinence.

L'utilisation des tests est toutefois assez complexe. En particulier, lorsque l'hypothèse doit être rejetée, l'analyse des causes et donc des solutions possibles n'est pas immédiate, la structure hiérarchique n'étant pas la seule ni forcément la meilleure solution. Nous ne ferons donc qu'ébaucher le test en renvoyant le lecteur vers l'ouvrage de Ben-Akiva et Lerman (1985) ou sur d'autres travaux cités par ces auteurs et notamment (McFadden et al., 1977 ; Small, Hsiao, 1982).

Le principe est assez simple. Si l'on pense que la propriété IIA n'est pas respectée en raison de la présence d'une alternative (notamment en raison de la non-indépendance entre cette alternative et une autre comme dans le cas de plusieurs modes collectifs), l'idée est d'estimer le modèle d'une part sur l'ensemble des alternatives et d'autre part sur un sous-ensemble d'alternatives (par exemple en excluant un ou plusieurs modes collectifs). Si la propriété IIA est respectée, les coefficients devraient rester inchangés pour les modes présents dans le sous-ensemble.

Plus concrètement, le test s'exécute de la manière suivante. Soit C_n l'ensemble complet des alternatives disponibles et \tilde{C}_n un sous-ensemble de C_n dans lequel on a enlevé les alternatives suspectes. Le modèle sur l'ensemble complet des alternatives, dans le cas d'une expression additive de l'utilité, s'écrit :

$$P\langle i | C_n \rangle = \frac{e^{\beta X_{in}}}{\sum_{j \in C_n} e^{\beta X_{jn}}} \quad (9.82)$$

De même, le modèle pour le sous-ensemble d'alternatives s'écrit :

$$P\langle i | \tilde{C}_n \rangle = \frac{e^{\beta X_{in}}}{\sum_{j \in \tilde{C}_n} e^{\beta X_{jn}}} \text{ avec } \tilde{C}_n \subset C_n \quad (9.83)$$

Dans la pratique, le second modèle ne peut être estimé que pour les observations ayant conduit à choisir une alternative appartenant à \tilde{C}_n . De même, seuls les paramètres relatifs à ces alternatives peuvent être estimés dans la fonction d'utilité. On note donc :

$\hat{\beta}_C$ le vecteur estimé pour l'ensemble des alternatives sur l'ensemble des données, mais ne comportant que les paramètres des variables relatives aux alternatives appartenant à \tilde{C}_n ;

$\hat{\beta}_{\tilde{C}}$ le vecteur estimé sur le sous-ensemble d'alternatives \tilde{C}_n avec uniquement les observations ayant conduit au choix d'une alternative appartenant à \tilde{C}_n . Ce vecteur ne comprend évidemment que les paramètres des variables relatives aux alternatives appartenant à \tilde{C}_n

$\Sigma \hat{\beta}_C$ la matrice de variance-covariance du vecteur $\hat{\beta}_C$;

$\Sigma \hat{\beta}_{\tilde{C}}$ la matrice de variance-covariance du vecteur $\hat{\beta}_{\tilde{C}}$.

Dans ce cas, la statistique suivante (en retenant une écriture matricielle) :

$$(\hat{\beta}_{\tilde{C}} - \hat{\beta}_C)' (\Sigma \hat{\beta}_{\tilde{C}} - \Sigma \hat{\beta}_C)^{-1} (\hat{\beta}_{\tilde{C}} - \hat{\beta}_C) \quad (9.84)$$

suit asymptotiquement une loi de χ^2 avec \tilde{K} degrés de liberté. Où \tilde{K} est le nombre de paramètres du vecteur $\hat{\beta}_{\tilde{C}}$ (Hausman, McFadden, 1984, cité par Ben-Akiva, Lerman, 1985).

Un autre test plus simple est obtenu à partir d'une approximation du test du ratio de vraisemblance (McFadden et al. 1977, cité par Ben-Akiva, Lerman, 1985) :

$$-2 \left[l_{\tilde{C}}(\hat{\beta}_C) - l_{\tilde{C}}(\hat{\beta}_{\tilde{C}}) \right] \quad (9.85)$$

qui suit également asymptotiquement une loi de χ^2 avec \tilde{K} degrés de liberté.

Le problème de ce test est qu'il peut être biaisé. On lui préfère donc la statistique suivante (Small, Hsiao, 1982 cité par Ben-Akiva, Lerman, 1985) :

$$\frac{1}{1 - N_1/\alpha N} \left\{ -2 \left[l_{\tilde{C}}(\hat{\beta}_C) - l_{\tilde{C}}(\hat{\beta}_{\tilde{C}}) \right] \right\} \quad (9.86)$$

où N_1 est le nombre d'observations sur lequel le modèle réduit est estimé, et N le nombre d'observations total.

α est un paramètre ≥ 1 . Cette statistique suit asymptotiquement une loi de χ^2 avec \tilde{K} degrés de liberté s'il est possible de faire l'hypothèse que les matrices de variance-covariance de $\hat{\beta}_{\tilde{C}}$ et de $\hat{\beta}_C$ diffèrent (asymptotiquement) d'au moins un facteur α , c'est-à-dire $\Sigma \hat{\beta}_C = \alpha^{-1} \Sigma \hat{\beta}_{\tilde{C}}$ (en retenant une écriture matricielle).

Pour ces tests, l'hypothèse nulle est celle de l'égalité des vecteurs $\beta_{\tilde{C}}$ et β_C . Si on ne peut rejeter l'hypothèse cela signifie que la structure logit est acceptable. Dans la pratique, il est possible de commencer le test avec $\alpha = 1$. Si l'on ne peut rejeter l'hypothèse avec $\alpha = 1$, on ne la rejettera pas non plus avec $\alpha > 1$. En revanche, si le test conduit au rejet de l'hypothèse, il faut affiner α .

Dans le cas où le test conduit au rejet de l'hypothèse avec la valeur adéquate de α , l'analyse n'est pas toujours facile, car la structure logit n'est pas forcément la seule cause de rejet (Ben-Akiva, Lerman, 1985). Ces auteurs proposent d'autres tests, mais ils sont en général plus lourds à mettre en œuvre.

3.2.7. Test de variation de perception (test of taste variation)

Il s'agit de la seconde hypothèse nécessaire à l'utilisation des modèles logit (section 2.4.1, p. 253). Elle stipule que tous les individus ont la même perception des différentes variables prises en compte dans la fonction d'utilité. Il est clair que cette hypothèse est parfois assez forte.

La variation de perception est souvent liée à des caractéristiques socio-économiques des individus (la perception des temps de transport peut par exemple dépendre de la motorisation (Lichère, Raux, 1997a, b) ou la perception des coûts dépendre du revenu...). L'idée de construction du test est très simple dans le principe. Si une variable socio-économique influence la perception des variables prises en compte dans la fonction d'utilité, une segmentation de la population selon cette variable doit conduire à des modèles dans lesquels les coefficients sont différents. Pour être plus précis, considérons la motorisation. Si l'on craint que la motorisation influence la perception des variables de la fonction

d'utilité, on segmente l'échantillon en trois sous-échantillons : déplacements réalisés par des ménages non motorisés ; ayant à disposition 1 voiture ; ayant à disposition 2 voitures ou plus. Ensuite, sur chacun des trois sous-échantillons, on estime le modèle. Si la motorisation n'influe pas sur la perception des autres variables, les vecteurs $\hat{\beta}$ de chacun des sous-échantillons doivent être identiques.

La construction du test s'effectue de la manière suivante (Ben-Akiva, Lerman, 1985). Considérons une variable de segmentation de la population. Cette variable comporte G classes d'effectifs N_g , $g = 1, 2, \dots, G$, avec :

$$\sum_g N_g = N \quad (9.87)$$

où N est la taille de l'échantillon.

L'hypothèse nulle « la variable considérée n'influe pas sur la perception » s'écrit :

$$\beta^1 = \beta^2 = \dots = \beta^G \quad (9.88)$$

où β^g est le vecteur de coefficients estimés sur l'échantillon N_g .

On peut construire la statistique du ratio de vraisemblance qui s'écrit :

$$-2 \left(\ln L_N(\hat{\beta}) - \sum_g \ln L_{N_g}(\hat{\beta}^g) \right) \quad (9.89)$$

Cette statistique suit un χ^2 avec un nombre de degrés de liberté égal au nombre de restrictions. Dans ce test, le modèle général est celui correspondant au marché segmenté. Le nombre de restrictions correspond à la différence entre le nombre de coefficients estimés dans le modèle général et le nombre de coefficients estimés dans le modèle réduit, c'est-à-dire celui sans segmentation de la population :

$$\left(\sum_g K_g \right) - K \quad (9.90)$$

avec K_g le nombre de coefficients de la fonction d'utilité du segment g et K le nombre de coefficients du modèle sur l'ensemble des observations. En principe $K_g = K$ sauf si, pour reprendre notre exemple, on introduit une constante pour les ménages sans voiture. Dans ce cas, le segment ménage sans voiture aura un coefficient de moins.

Le test est assez général, car il permet d'examiner l'égalité de l'ensemble du vecteur. Toutefois, en cas de rejet de l'hypothèse (c'est-à-dire un effet de perception), la différence de perception ne concerne pas forcément l'ensemble des variables de la fonction d'utilité. Dans ce cas, il est nécessaire d'examiner les coefficients individuellement. On utilise pour cela le test de student suivant :

$$\frac{\hat{\beta}_k^g - \hat{\beta}_k^h}{\sqrt{\text{var}(\hat{\beta}_k^g) + \text{var}(\hat{\beta}_k^h)}} \quad (9.91)$$

où $\hat{\beta}_k^g, \hat{\beta}_k^h$ est la valeur du coefficient de la fonction d'utilité de la variable k estimée respectivement sur le segment g, h ;

$\text{var}(\hat{\beta}_k^g)$ est la variance du coefficient $\hat{\beta}_k^g$.

Si le nombre de coefficients statistiquement différents est faible, la solution peut consister à diviser, au sein de la fonction d'utilité, la variable en autant de variables qu'il y a de segments conduisant à des valeurs différentes (par exemple séparation de la variable temps de parcours en deux variables l'une pour les ménages non motorisés, l'autre pour les ménages ayant au moins une voiture si les coefficients pour les ménages ayant 1 ou 2 et plus voitures ne sont pas statistiquement différents).

Si le nombre de variables concernées est important, il n'est généralement pas souhaitable de multiplier le nombre de variables du modèle. Dans ce cas, il peut être nécessaire de segmenter effectivement la population en sous groupes homogènes. Lorsque le problème est encore plus important, une segmentation de la population peut ne pas être suffisante. Dans ce cas, il peut être nécessaire d'utiliser un modèle probit. D'autres solutions moins complexes à mettre en œuvre sur le plan de l'estimation du modèle ont déjà été présentées lors de la présentation des développements méthodologiques récents (section 2.7, p. 268).

3.2.8. Conclusion

D'autres tests peuvent être construits pour tester certaines des hypothèses nécessaires à l'application de ces modèles ou pour tester d'autres spécifications de la fonction d'utilité (Ben-Akiva, Lerman, 1985). Une analyse exhaustive de l'ensemble des tests, lors d'une étude, est toutefois impossible car elle demanderait beaucoup trop de temps pour être menée. Les tests basiques sont évidemment indispensables (signes des coefficients, élasticité, t de student, ρ^2 et ratio de vraisemblance). Mais au-delà, le choix des tests doit être guidé par l'analyse du modèle, par l'analyse des données dont on dispose et enfin par des considérations théoriques. L'analyste doit garder en mémoire que les tests statistiques ne permettent que de confronter certaines hypothèses aux données empiriques. Ce n'est pas la statistique qui leur donnera du sens, c'est la réflexion de l'analyste. Sans cette réflexion le modèle n'est rien et n'a aucun pouvoir prédictif, aucune opérationnalité.

Cette réflexion passe par une analyse des variables explicatives de l'objet d'étude que l'on cherche à prédire ou à simuler et des relations qui les lient entre elles et avec les variables caractérisant cet objet. Cette analyse conduit à faire des choix menant vers certaines familles de modèles compte tenu des hypothèses que l'on peut accepter puis vers la spécification du modèle.

4. SPECIFICATION DES MODELES DESAGREGES

La spécification d'un modèle désagrégé passe par l'examen de plusieurs dimensions. On peut regrouper ces choix au sein de trois grands groupes. Nous considérons tout d'abord le choix des alternatives disponibles dans le processus de choix (section 4.1). Nous abordons ensuite la sélection des variables explicatives et leur spécification dans la fonction d'utilité (section 4.2). Enfin, nous considérons le choix d'une structure de modèle (section 4.3). Il est évident que ces trois grands groupes interagissent entre eux et que l'analyse n'est pas séquentielle.

Comme nous l'avons explicité ci-dessus, ces choix doivent découler de considérations théoriques relatives au comportement des individus. Dans la pratique, il s'agit d'un processus plus heuristique dans lequel l'analyse des données d'enquête, des sorties du modèle, des tests statistiques présentés dans la section précédente et l'analyse de la pertinence des hypothèses nécessaires à l'application des différents modèles permettent de tester et de confronter ces modèles. Le choix entre ces modèles découle de l'analyse de leur opérationnalité au sens où nous l'avons défini chapitre 2 (p. 45).

4.1. Choix des alternatives

La question du choix des alternatives est en apparence simple à résoudre. Dans la pratique, les choses ne sont pas aussi simples pour deux raisons principalement. La première relève de la précision recherchée dans la définition des alternatives. La seconde concerne la définition des alternatives disponibles pour chacun des individus parmi l'ensemble des modes considérés.

Dans l'approche agrégée, il est fréquent de ne considérer que deux modes de transport, voiture particulière d'une part et transports collectifs d'autre part. Dans ce cas, la répartition des déplacements en transports collectifs entre les différents moyens de transports composant ce mode (bus, tramway, métro, train...) est effectuée au cours de l'étape d'affectation. Dans l'approche désagrégée, la définition des modes de transport est généralement plus fine. Le choix du degré de finesse dépend bien évidemment des objectifs de l'étude. S'il s'agit d'étudier les potentialités du covoiturage (car-pool), une décomposition assez fine du mode voiture sera souhaitable. De même, s'il s'agit de simuler l'impact d'une ligne de tramway, on aura intérêt à segmenter le mode transports collectifs.

Cette segmentation ne permet plus toujours de maintenir l'hypothèse d'indépendance des termes d'erreur entre les alternatives conduisant à des structures de modèles pouvant être plus complexes. La considération des trois conditions de l'opérationnalité des modèles pourra aider dans la détermination du niveau de précision adéquat.

L'ensemble des modes ou plus généralement des alternatives considérées constitue l'univers de choix disponible. Toutefois, ces alternatives ne sont pas forcément disponibles ou considérées pour l'ensemble des individus. Un individu appartenant à un ménage non motorisé n'a généralement pas accès à la voiture en tant que conducteur, mais peut l'utiliser en tant que passager. Si l'arrêt de bus le plus proche est situé à une distance trop importante, l'individu ne pourra pas y accéder en marche à pied, mais éventuellement en voiture. Ces exemples renvoient donc d'une part à la définition des modes et d'autre part à la question de leur disponibilité pour les individus.

Par ailleurs, des modes apparemment disponibles peuvent très bien ne pas être considérés par certains individus. On sait en effet que l'hypothèse du consommateur complètement informé est une fiction et que les individus ont une appréhension de l'univers de choix pouvant s'apparenter à un processus de filtrage découlant du poids des habitudes. La prise en compte dans le modèle de modes appartenant à l'univers de choix mais non considérés par un nombre important d'individus peut biaiser l'estimation du modèle. Ortuzar et Willumsen (1994) proposent trois solutions pour traiter ce problème :

- l'utilisation de règles heuristiques (à partir des données disponibles) ou déterministes permettant d'exclure certaines alternatives de l'univers de choix, par exemple le bus n'est pas disponible si l'arrêt est situé à plus de telle distance ;
- le recueil de données permettant de déterminer l'univers de choix directement auprès de l'individu enquêté. Par exemple en interrogeant les enquêtés sur leur perception des alternatives disponibles (ou mieux sur les alternatives considérées comme non disponibles et sur la raison de leur indisponibilité) ;
- l'utilisation d'une procédure aléatoire de détermination du sous-échantillon d'alternatives disponibles. Les probabilités de choix résultent alors d'un processus en deux temps : premièrement détermination du sous-échantillon des alternatives disponibles à partir d'une fonction de distribution des alternatives disponibles, ensuite sur ce sous-échantillon d'alternatives, détermination d'une probabilité de choix pour chacune des alternatives (Lerman, 1984 ; Richardson, 1982 cités par Ortuzar, Willumsen, 1994).

4.2. Choix des variables, de leur spécification et de leur forme fonctionnelle

Le choix des variables doit d'abord découler d'une analyse du comportement de choix des individus. Comme toujours, c'est la construction théorique qui doit guider le choix des variables explicatives de l'objet d'étude. Toutefois, il existe des procédures, comme pour la régression linéaire, permettant de savoir si l'ajout d'une variable supplémentaire dans le modèle permet d'améliorer la reproduction par le modèle du jeu de données. Il s'agit du quasi-t test sur la nullité du coefficient (section 3.2.2, p. 278), puis du ρ^2 ajusté (section 3.2.4, p. 279) qui permet d'examiner l'accroissement de la valeur de la fonction de vraisemblance (il est nécessaire d'utiliser le ρ^2 ajusté, car la valeur du ρ^2 augmente automatiquement quand le nombre de variables s'accroît) et enfin le test du ratio de vraisemblance (section 3.2.5, p. 281) pour savoir si cet accroissement est statistiquement significatif ou non.

4.2.1. Détermination des constantes

Parmi l'ensemble des variables présentes dans la fonction d'utilité, on retient généralement une constante par mode de transport. Toutefois, l'estimation du modèle ne permet pas de calculer toutes les constantes. Il y a toujours une indétermination. Prenons l'exemple du logit binaire avec une formulation additive de l'utilité pour illustrer le propos. En reprenant l'équation 9.21 :

$$P(i|C_n) = \frac{1}{1 + e^{\mu \sum_k \beta(X_{in} - X_{jn})}}$$

L'estimation du modèle revient à estimer, la différence d'utilité entre les alternatives i et j , soit : $\sum_k \beta(X_{in} - X_{jn})$. Si l'on abandonne l'indice n pour simplifier les notations et que l'on considère deux variables explicatives pour chacune des alternatives, on obtient l'équation suivante :

$$\beta_1 X_{i1} + \beta_2 X_{i2} - \beta_3 X_{j3} - \beta_4 X_{j4} + (\beta_5 - \beta_6) \quad (9.92)$$

où β_5 et β_6 sont les constantes respectivement de l'alternative i et j .

L'estimation du modèle ne permettra donc pas d'estimer simultanément β_5 et β_6 , mais seulement leur différence. L'exemple peut en fait être généralisé à un plus grand nombre d'alternatives. Il est donc nécessaire de fixer l'une des constantes à 0. Toutes les autres constantes sont donc estimées relativement à celle de cette alternative.

Le problème est de même nature lorsque deux alternatives partagent une même variable explicative. Considérons l'équation 9.92, et supposons que les variables X_{i1} et X_{j3} sont en fait une seule et même variable (par exemple le tarif). Si cette variable prend la même valeur pour les deux alternatives (par exemple bus et métro), on se retrouve de nouveau face à un cas d'indétermination où l'on ne peut déterminer que la différence des coefficients : $(\beta_1 - \beta_3)$. Le problème n'est pas très fréquent lorsque l'on considère une variable explicative relative au déplacement, il sera par contre systématique pour des variables socio-économiques caractérisant l'individu, comme le revenu par exemple.

4.2.2. Variables spécifiques ou génériques

Une dimension importante de la spécification du modèle concerne le choix de variables spécifiques ou génériques :

- variables spécifiques : variables dont le coefficient est différent pour chacune des alternatives ;

- variables génériques : variables dont le coefficient est identique pour toutes les alternatives ou au moins pour un groupe d'alternatives.

La question se pose par exemple pour l'estimation des coefficients relatifs au temps passé dans le véhicule. Le coefficient est-il le même pour l'ensemble des modes ou tout au moins pour l'ensemble des modes collectifs ou doit-il être différent pour chacune des alternatives. Deux tests peuvent être utilisés pour savoir si une variable est générique (à l'ensemble des alternatives ou à un groupe d'alternatives) ou spécifique.

Le premier utilise la matrice de variance-covariance entre les alternatives qui est généralement fournie par le logiciel d'estimation. Pour cela, il faut estimer le modèle en considérant que la variable est spécifique, puis tester l'égalité des coefficients. Soit β_1 et β_2 , les coefficients d'une variable relativement aux alternatives 1 et 2.

$$\text{var}(\hat{\beta}_1 - \hat{\beta}_2) = \text{var}(\hat{\beta}_1) + \text{var}(\hat{\beta}_2) - 2 \text{cov}(\hat{\beta}_1, \hat{\beta}_2) \quad (9.93)$$

où $\text{var}()$ est la variance et $\text{cov}()$ est la covariance.

$$\text{Le rapport } \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{\text{var}(\hat{\beta}_1 - \hat{\beta}_2)}} \quad (9.94)$$

suit une loi de student permettant de tester l'hypothèse nulle $\beta_1 = \beta_2$.

Ce test doit être repris pour chacune des paires d'alternatives. Il est donc préférable de commencer par tester le caractère spécifique de la variable pour l'ensemble des alternatives. Il faut pour cela utiliser le test du ratio de vraisemblance. Ce test nécessite d'estimer deux modèles, le premier dans lequel la variable est spécifique et le second dans lequel la variable est générique. Le premier est considéré comme le modèle global et le second comme le modèle réduit. On utilise alors le test du ratio de vraisemblance (équation 9.79, le nombre de degrés de liberté étant le nombre d'alternatives pour lequel on impose l'égalité du coefficient – 1).

Si l'on ne peut rejeter l'hypothèse, on peut donc considérer que les deux modèles sont égaux. Dans ce cas la variable est considérée comme générique. Dans le cas contraire, il y a au moins un coefficient qui diffère statistiquement des autres.

Si une variable socio-économique apparaît comme générique pour l'ensemble des alternatives, il ne sera pas possible d'estimer son coefficient. Reprenons l'équation 9.92. Si β_1 et β_3 sont égaux dans ce cas $(\beta_1 - \beta_3)$ est systématiquement nul empêchant la détermination du coefficient commun.

4.2.3. Forme fonctionnelle et transformations de variables

Si l'expression de l'utilité est généralement additive, cela ne signifie pas pour autant que l'effet de chacune des variables explicatives est forcément linéaire. Par exemple, on considère fréquemment que le temps en véhicule n'est pas ressenti de la même manière par les individus pour les premières minutes et pour les minutes supplémentaires lorsque le temps passé est déjà élevé (CERTU, 1998b ; Ben-Akiva, Lerman, 1985). Une expression purement linéaire en fonction du temps en véhicule ne permet donc pas de représenter correctement la perception du temps par les individus. Dans ce cas, deux solutions sont disponibles. La première consiste à segmenter la variable par intervalles successifs et à estimer un coefficient pour chacun des intervalles. La seconde consiste à transformer la variable à l'aide d'une fonction non linéaire. Nous explicitons ces deux méthodes.

L'application de la première méthode est très simple sur le principe. Elle consiste à segmenter la variable explicative. Par exemple dans l'étude décrite dans l'ouvrage de Ben-Akiva et Lerman (1985), le temps de parcours a été scindé en 4 intervalles, conduisant à la définition des variables suivantes :

$$T_1 \text{ (temps de trajet de 0 à 20 minutes) : } \begin{cases} t & \text{si } t < 20 \\ 20 & \text{si } t \geq 20 \end{cases}$$

$$T_2 \text{ (temps de trajet de 20 à 50 minutes) : } \begin{cases} 0 & \text{si } t < 20 \\ t - 20 & \text{si } 20 \leq t < 50 \\ 50 & \text{si } t \geq 50 \end{cases}$$

$$T_3 \text{ (temps de trajet de 50 à 80 minutes) : } \begin{cases} 0 & \text{si } t < 50 \\ t - 50 & \text{si } 50 \leq t < 80 \\ 80 & \text{si } t \geq 80 \end{cases}$$

$$T_4 \text{ (temps de trajet } > 80 \text{ minutes) : } \begin{cases} 0 & \text{si } t < 80 \\ t - 80 & \text{si } t \geq 80 \end{cases}$$

L'estimation du modèle conduit à estimer 4 coefficients pour la variable temps de parcours. Cette méthode est assez simple. Elle pose toutefois quelques problèmes de mise en œuvre. La définition des segments nécessite une bonne connaissance de l'effet de la variable sur l'objet d'étude. La détermination des frontières est évidemment sujette à caution, car elle crée une discontinuité dans l'incidence d'une minute supplémentaire selon que l'on est juste avant ou après une frontière. En général, l'analyse des données empiriques permet d'aider à la définition du nombre de classes nécessaires et à leur frontière. Le dernier problème concerne le nombre de variables de la fonction d'utilité. Il est évident que le nombre de classes doit rester limité. De même, cette segmentation ne peut être effectuée que sur un très petit nombre de variables sous peine d'accroître fortement le nombre de variables explicatives. Dans ce cas, le risque est grand d'avoir des variables reprenant une partie de la variance sans avoir un réel pouvoir explicatif.

Il est donc nécessaire de tester la significativité de cette segmentation. La première étape comme toujours doit être d'examiner le signe des coefficients et la valeur relative des coefficients pour les différents segments. Les signes et les valeurs relatives doivent être conformes à la connaissance que l'on a de l'incidence de la variable. Dans l'exemple cité ci-dessus, on serait en droit d'attendre une incidence dégressive sur la fonction d'utilité (tout au moins dans un contexte urbain, pour des trajets longue distance, la relation est probablement différente). Pour compléter l'analyse, on dispose ensuite des tests statistiques déjà évoqués. Quasi-t test pour tester l'égalité des coefficients, ρ^2 ajusté pour déterminer le gain de qualité d'estimation et enfin le test du ratio de vraisemblance afin de savoir si ce gain est statistiquement significatif ou non.

La seconde solution consiste à transformer la variable à l'aide d'une fonction non linéaire adhoc. Si l'on a suffisamment de connaissance sur l'incidence de la variable sur l'objet d'étude, on peut déterminer cette fonction a priori. Il est alors souhaitable d'examiner les résultats et généralement de confronter les résultats obtenus avec cette transformation avec ceux obtenus sans transformation. Pour cela, on examine comme toujours le ρ^2 ajusté pour déterminer le gain de qualité d'estimation. En revanche, il n'est plus possible d'utiliser le test du ratio de vraisemblance pour savoir si la différence est statistiquement significative ou non, car un des deux modèles n'est pas une restriction de l'autre. Pour traiter cette situation, Ben-Akiva et Lerman proposent le test suivant :

$$\Pr((\bar{\rho}_2^2 - \bar{\rho}_1^2) > z) \leq \Phi\left\{-\sqrt{-2zI(0) + (K_2 - K_1)}\right\} \quad (9.95)$$

où z est un nombre strictement positif ;

$\bar{\rho}_1^2$ est l'index ajusté du taux de vraisemblance du modèle 1 ;

$\bar{\rho}_2^2$ est l'index ajusté du taux de vraisemblance du modèle 2 ;
 K_1 est le nombre de variables (en incluant les constantes) de la fonction d'utilité du modèle 1 ;
 K_2 est le nombre de variables (en incluant les constantes) de la fonction d'utilité du modèle 2 ;
 $l(0)$ la fonction de vraisemblance des modèles 1 et 2 quand tous les coefficients sont nuls ;
 Φ la fonction cumulative de la distribution normale standard.

Dans ce test, le modèle 1 est celui qui présente la plus faible valeur du $\bar{\rho}^2$. L'hypothèse nulle est donc : « le modèle 2 correspond à la meilleure spécification du modèle ».

Si tous les individus enquêtés disposent effectivement de l'ensemble des I alternatives appartenant à l'univers de choix C_n disponible, $l(0) = N \ln(I)$. Dans ce cas, l'équation 9.95 devient :

$$\Pr((\bar{\rho}_2^2 - \bar{\rho}_1^2) > z) \leq \Phi \left\{ -\sqrt{-2zN \ln(I) + (K_2 - K_1)} \right\} \quad (9.96)$$

avec N la taille de l'échantillon et I le nombre d'alternatives.

La connaissance de l'incidence de la variable sur l'objet d'étude n'est toutefois pas toujours suffisante pour déterminer a priori la forme fonctionnelle de la transformation. On utilise alors en général la transformation de Box-Cox, définie par :

$$\begin{cases} \frac{X^\lambda - 1}{\lambda}, & \text{pour } \lambda \neq 0 \\ \ln(X), & \text{pour } \lambda = 0 \end{cases} \quad (9.97)$$

On peut noter que pour $\lambda = 1$, il ne s'agit que d'une transformation linéaire. L'estimation du paramètre λ pour chacune des variables pour laquelle on souhaite faire une transformation s'effectue lors de l'estimation du modèle. Elle consiste à rechercher la valeur des paramètres $\hat{\lambda}_k$ et du vecteur de paramètres $\hat{\beta}$ qui maximise la fonction de vraisemblance $l(\hat{\beta}, \hat{\lambda})$. Si la valeur obtenue lors de l'estimation est proche de 1, il n'est pas souhaitable de conserver la transformation. Il est toujours possible de faire le test ci-dessus (équation 9.95) pour savoir si le gain sur l'index du taux de vraisemblance est statistiquement significatif ou non.

Cette transformation n'est définie que pour des valeurs de $X \geq 0$. Dans le cas où X peut prendre des valeurs négatives, on utilise la transformation de Box-Tukey définie par :

$$\begin{cases} \frac{(X + \alpha)^\lambda - 1}{\lambda}, & \text{pour } \lambda \neq 0 \\ \ln(X + \alpha), & \text{pour } \lambda = 0 \end{cases} \quad (9.98)$$

où α est choisi de telle sorte que $(X + \alpha)$ ne prenne que des valeurs positives. α n'intervient donc que comme une translation linéaire.

4.3. Structure du modèle

Le choix d'une structure de modèle ou, pour reprendre la terminologie retenue pour la première section, d'une famille de modèles dépend principalement de la nature des hypothèses que l'on considère comme acceptables. Ce choix est bien évidemment effectué sous contrainte de budget monétaire ou temps disponible. Il s'agit donc de choisir la structure de modèle conduisant à la plus grande opérationnalité selon la définition du chapitre 2.

Le modèle le plus couramment utilisé est le modèle logit multinomial. Deux raisons expliquent cette situation. Tout d'abord, ce modèle permet une formulation extrêmement simple de la probabilité de choix d'une alternative. Ensuite, les logiciels d'estimation sont facilement disponibles. Toutefois, il

repose sur des hypothèses restrictives qu'il convient de bien garder en mémoire. Elles sont au nombre de trois (section 2.4.1, p. 253) :

- les parties aléatoires de l'utilité pour les différentes alternatives sont indépendantes entre elles et identiquement distribuées (IID). L'hypothèse d'indépendance signifie qu'il n'y a pas de facteurs non observés qui soient communs à deux alternatives. Le second élément de l'hypothèse concerne l'étendue de la variation autour de la partie déterministe de l'utilité. On considère que les termes d'erreurs sont distribués identiquement pour chacune des alternatives. C'est la combinaison de ces deux hypothèses qui conduit à la propriété IIA d'indépendance vis-à-vis des autres alternatives ;
- la seconde hypothèse concerne l'homogénéité des réponses aux différents attributs de la partie déterministe de l'utilité entre les individus. Plus précisément, le modèle n'autorise pas de variation de perception (*taste variation* en anglais) d'un attribut de la fonction d'utilité due à des caractéristiques non observées (c'est-à-dire non présentes dans la partie déterministe de l'utilité) de l'individu. Plus concrètement encore, les paramètres relatifs à chacune des variables présentes dans la fonction d'utilité sont identiques pour tous les individus. La perception du temps ou du coût du déplacement doit donc être identique pour tous les individus ;
- la troisième hypothèse concerne la structure de la matrice de variance-covariance du terme d'erreur entre les alternatives. On fait une hypothèse de structure identique entre les individus. Concrètement cela signifie que la variation due aux facteurs non observés (c'est-à-dire non pris en compte dans la partie déterministe de l'utilité) est identique entre les individus.

Ces hypothèses ne sont bien évidemment pas toujours respectées. C'est la première hypothèse et surtout son corollaire la propriété IIA qui a bénéficié de la plus grande attention. L'expérience montre en effet que cette propriété n'est pas toujours crédible, notamment lorsque des alternatives sont assez proches ou partagent des attributs communs (comme dans le cas du choix combiné mode-destination). Des tests (section 3.2.6, p. 283) existent pour tester le non-respect de cette propriété.

Dans ce cas, le logit emboîté constitue une première réponse. Il permet de traiter le cas où les termes d'erreur de certaines alternatives ne sont pas indépendants entre eux. En revanche, il n'abandonne pas l'hypothèse d'identité des distributions (section 2.5, p. 259). Les deux autres hypothèses restent également inchangées. D'autres modèles existent également qui appartiennent à la classe des modèles de la valeur extrême généralisée « *Generalized Extreme Value* » (section 2.6, p. 265, Bhat, 1997 ; Koppelman, Sethi, 2000) pour traiter de cas plus généraux de non-indépendance du terme d'erreur entre les alternatives. Ces modèles ne sont pas encore trop complexes à estimer car ils bénéficient des mêmes algorithmes d'estimation que le logit multinomial. En revanche, la spécification du modèle n'est pas toujours simple. En particulier, le choix du nombre de niveaux hiérarchiques et de l'affectation des alternatives à chacun des niveaux ou du choix de la structure de variance-covariance entre les alternatives peut dans certains cas s'avérer problématique (Bhat, 1997).

Dans l'article précité, Bhat propose des modèles qui ne nécessitent pas l'hypothèse d'une distribution identique des termes d'erreur de chacune des alternatives, mais des distributions indépendantes entre elles. L'utilisation de tels modèles peut être nécessaire lorsque l'on est confronté à une situation où une ou plusieurs variables non prises en compte dans la partie déterministe de l'utilité varient de manière très différente selon les alternatives. L'exemple suivant correspond à une telle situation : si le confort apparaît comme une variable assez importante, mais non disponible et que les conditions de confort varient très fortement pour le mode collectif, mais peu pour la voiture, il est souhaitable de postuler des distributions non identiques du terme d'erreur. En revanche, dans ce cas, il est possible de maintenir l'hypothèse d'indépendance du terme d'erreur entre les deux alternatives. Ces modèles peuvent être une réponse dans le cas où la propriété IIA n'est pas respectée. Le problème principal dans l'utilisation de cette famille de modèle réside dans l'estimation car les outils traditionnellement disponibles dans les logiciels standards ne peuvent pas être utilisés.

Le cas le plus général de violation de la propriété IIA correspond au cas où les termes d'erreur ne sont ni indépendants, ni identiquement distribués. Dans ce cas, la première hypothèse n'est plus maintenue. Les formulations simples ne sont plus disponibles. Il est nécessaire de recourir au modèle multinomial probit (section 2.3, p. 251). En ayant une structure de matrice de variance-covariance entre les alternatives complètement libre, il permet également l'abandon de la seconde hypothèse. Le principal problème de ce modèle est sa complexité d'estimation. Il nécessite autant d'intégration qu'il y a d'alternatives dans l'univers de choix. « *Un autre problème de l'estimation du modèle MNP (multinomial probit) est le nombre de paramètres à estimer dans la matrice de variance-covariance. Une structure complètement libre ... génère de nombreux problèmes conceptuels, statistiques et pratiques, et notamment des difficultés d'interprétation, un comportement du modèle non-intuitif, et une faible précision de l'estimation des paramètres de la matrice de variance-covariance* » (Bhat, 2000, voir aussi Srinivasan, Mahmassani, 2000). Il n'est toutefois pas toujours nécessaire de postuler une structure de matrice complètement libre. Le plus souvent, il est possible d'introduire des restrictions.

Toutefois, même avec ces restrictions, l'estimation de modèles probit multinomiaux reste extrêmement lourde. On lui préfère donc le modèle logit multinomial mixte (MMNL, section 2.7, p. 268), qui présente le mérite de pouvoir être estimé avec un simulateur multinomial logit Kernel (Bhat, 2000). Ce modèle permet, tout comme le probit, d'abandonner les deux premières hypothèses des modèles logit (Bhat, 2000 ; McFadden, 2000 ; Brownstone, 2000).

Si l'on souhaite maintenir l'hypothèse d'une distribution identique et indépendante des termes d'erreur entre les alternatives, mais que la seconde hypothèse ne peut être maintenue, on dispose de réponses plus souples que le passage forcé par le modèle probit ou le modèle logit mixte multinomial. Le non-respect de cette hypothèse peut être dû à des variations de perception de certains attributs entre les individus, comme pour le temps en véhicule (section 2.4.1, p. 253). Des tests (section 3.2.7, p. 284) permettent de vérifier le respect de l'hypothèse. Dans le cas d'un non-respect, deux solutions assez simples sont disponibles (section 3.2.7, p. 284). La première consiste à segmenter la variable incriminée en autant de segments que nécessaires de telle sorte que l'hypothèse puisse être considérée comme acceptable au sein de chacun des segments. La seconde solution consiste à segmenter la population enquêtée en sous-échantillons pour lesquels l'hypothèse de constance de la perception pourrait être considérée comme acceptable. D'autres solutions plus complexes prennent en compte des variations plus importantes de la perception de certaines variables ou de toutes les variables. Ces variations peuvent être dépendantes des caractéristiques socio-économiques des individus ou non dépendantes. Dans ce dernier cas, on postule généralement une distribution normale ou bien log-normale (comme pour la valeur du temps) de la perception de la variable par les individus (ces différentes formes sont présentées dans Bhat, 1997).

La troisième hypothèse concerne l'homogénéité entre les individus de la structure de la matrice de variance-covariance du terme d'erreur entre les alternatives. Bhat (1997) présente différentes structures de modèles permettant de répondre à un non-respect de l'hypothèse.

5. PREVISION ET AGREGATION

L'estimation du modèle désagrégé est réalisée sur des données individuelles. En revanche, en prévision, le modèle n'est généralement plus utilisé sur des données individuelles. L'utilisation d'un modèle désagrégé en prévision nécessite alors un processus d'agrégation. Ce processus d'agrégation est synonyme de biais potentiels (section 5.1). Nous examinons donc les méthodes susceptibles de réduire ce biais (section 5.2).

5.1. Prévision et biais d'agrégation

Nous illustrons ce biais en reprenant l'exemple cité par Ortuzar et Willumsen (1994). Considérons un modèle logit multinomial avec une expression additive de la partie déterministe de l'utilité. Dans ce modèle, la variance de la distribution du terme d'erreur ne peut être déterminée indépendamment du paramètre μ (équation 9.24). De même, le vecteur de coefficients des variables explicatives de l'utilité ne peut être déterminé indépendamment du paramètre μ . C'est pourquoi, on pose en général $\mu = 1$, ce qui conduit à une variance $\sigma^2 = \pi^2/6$ pour le terme d'erreur (équation 9.24).

Supposons que les hypothèses du logit multinomial sont bien respectées et que l'on a estimé le modèle sur des données individuelles. Soit β le vecteur de solution pour l'alternative i . Supposons que nous cherchions à estimer un des paramètres (par exemple celui relatif à la variable X_{1i}) en utilisant des données agrégées pour cette variable. La variable X_{1i} est donc remplacée par la variable Z_{1i} , avec :

$$X_{1i} = Z_{1i} + \tau_1 \quad (9.99)$$

en supposant que τ_1 suit une distribution de Gumbel avec une moyenne nulle et un écart-type σ_τ .

En reprenant les notations habituelles pour la formulation additive de l'utilité, l'utilité de l'alternative i s'écrit :

$$U_i = \beta_1 Z_{1i} + \sum_{k \neq 1} \beta_k X_{ki} + \delta_i \quad (9.100)$$

avec δ_i qui possède toujours une distribution de Gumbel, mais avec une variance plus élevée pour tenir compte de la distribution de la variable X_{1i} autour de la valeur moyenne Z_{1i} : $(\beta_1^2 \sigma_\tau^2 + \sigma^2)$.

Dans ce cas l'estimation des coefficients pour l'alternative i , ne sera plus β_k , mais :

$$\psi_k = \frac{\sigma}{\sqrt{(\beta_1^2 \sigma_\tau^2 + \sigma^2)}} \beta_k \quad (9.101)$$

Le terme au dénominateur est forcément plus grand que σ . De ce fait, l'estimation conduit à une sous-estimation systématique des coefficients. C'est cette sous-estimation qui est généralement appelée le biais d'agrégation dans les modèles désagrégés. De ce fait, il est recommandé dans les modèles désagrégés de ne pas utiliser de variables agrégées, au niveau zonal par exemple (cette recommandation ne peut toutefois pas être complètement respectée. Le temps d'accès au réseau de transports collectifs est par exemple une variable zonale qui entre en général dans la fonction d'utilité à un niveau agrégé, car les localisations ne sont généralement pas connues au niveau des coordonnées x-y. Le choix du zonage influe donc forcément sur l'estimation des coefficients puisqu'il conduit en général à une variance s'accroissant à mesure que le nombre de zones diminue. Nous revenons sur cette question dans la section 6).

Ce biais a des conséquences en prévision (nous reprenons l'exemple de Ortuzar et Willumsen (1994), qui empruntent eux-mêmes à Gunn (1985)). Considérons le cas où la variable X_1 prend une valeur double pour toutes les alternatives comprenant cette variable, la valeur des autres variables restant inchangée. Une ré-estimation du modèle à un niveau désagrégé avec un nouveau jeu de

données aléatoires d'effectif suffisant conduira au même jeu de coefficients. Cela signifie que le modèle estimé sur le premier jeu de données sera capable de prévoir correctement l'effet de l'évolution de la variable X_1 .

En revanche, si l'estimation est effectuée en utilisant des données agrégées pour X_1 , la variance de δ_i (équation 9.100) deviendra $(\beta_1^2 4\sigma_\tau^2 + \sigma^2)$, puisque la nouvelle variance de τ_1 devient $4\sigma_\tau^2$. L'application de l'équation 9.101 fournit à de nouveaux coefficients :

$$\psi'_k = \frac{\sigma}{\sqrt{(\beta_1^2 4\sigma_\tau^2 + \sigma^2)}} \beta_k$$

Donc ψ_k est supérieur à ψ'_k . L'utilisation du modèle avec le premier jeu de données agrégées conduira donc à une surestimation de l'effet du doublement de la variable X_1 . Une réduction de la valeur de la variable X_1 conduirait à une sous-estimation.

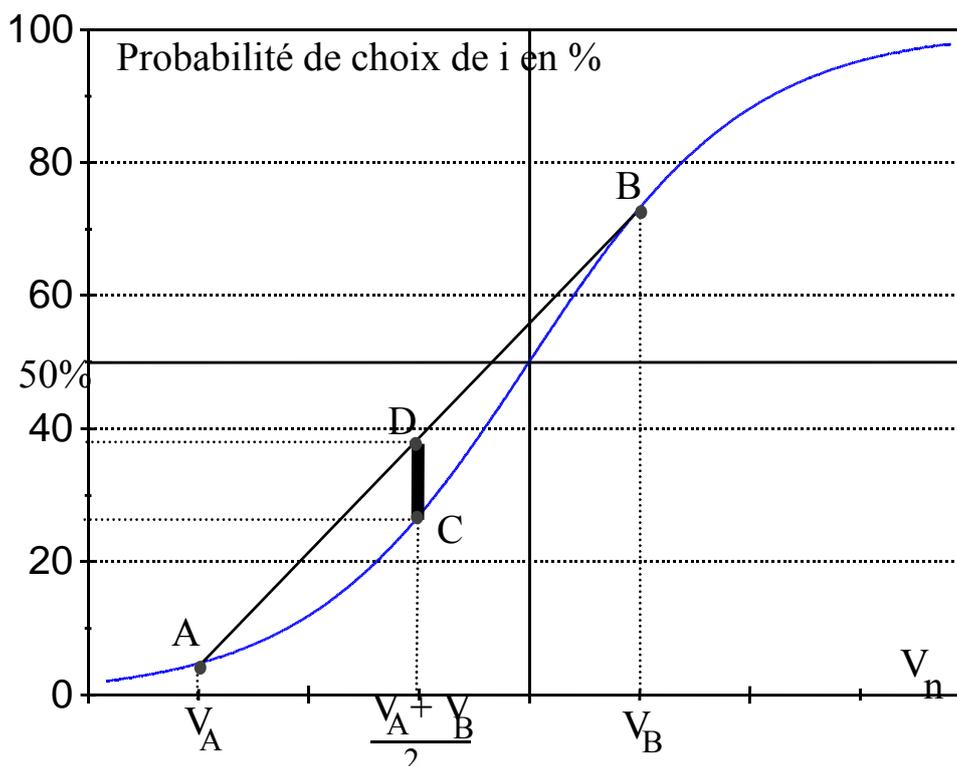
5.2. Méthodes d'agrégation

Dans un modèle désagrégé, la probabilité de choix d'une alternative i par un individu n dépend des caractéristiques de cet individu. Pour une population N , la probabilité agrégée sera fournie par l'équation suivante :

$$P_N(i) = \frac{1}{N} \sum_n P_n(i) \quad (9.102)$$

La probabilité agrégée est tout simplement la moyenne des probabilités de choix de l'alternative i par chacun des individus composant la population N . Dans un modèle linéaire, cette probabilité pourrait être estimée à partir d'un individu moyen doté des caractéristiques moyennes de la population étudiée. En revanche, avec un modèle logit, l'utilisation d'un individu moyen a toutes les chances de conduire à une solution différente de celle obtenue à partir d'une application de l'équation 9.102. Un exemple avec deux individus seulement (exemple repris de Ortuzar, Willumsen, 1994) permet d'illustrer graphiquement le problème (graphique 9.8). Soit deux individus A et B. L'estimation de la probabilité moyenne à partir de l'équation 9.102 conduira à la valeur de la probabilité correspondant au point D, tandis que l'estimation à partir de l'utilité moyenne des deux individus conduit à la valeur de la probabilité correspondant au point C. Le trait épais schématise le biais d'agrégation lié à l'individu moyen.

Toutefois, dans un exemple concret de prévision, il n'est généralement pas possible d'énumérer l'ensemble des situations individuelles pour appliquer l'équation 9.102, car l'information n'est pas disponible. Une agrégation est donc obligatoire. Plusieurs solutions sont disponibles (Ortuzar, Willumsen, 1994).



Graphique 9.8 : Illustration du problème de l'agrégation en prévision dans les modèles désagrégés

Dans la première, si l'on considère que l'échantillon d'estimation constitue une bonne approximation de la population étudiée, on pourra utiliser cet échantillon pour appliquer le principe d'énumération (équation 9.102). Chaque individu de cette base est alors affecté d'un poids pour obtenir la population totale. Il est évident que la base de données utilisée pour l'estimation doit avoir un effectif suffisamment important d'une part et être représentative de la population étudiée d'autre part, pour que l'approximation soit acceptable, notamment relativement aux objectifs de la prévision. Le domaine de pertinence de cette méthode reste toutefois limité. Si l'approximation de la population future par les données en référence peut être acceptable dans des environnements peu évolutifs, il est clair que la précision décroît à mesure que l'horizon temporel s'accroît et que le contexte évolue. En effet, cette méthode suppose une distribution inchangée de la population relativement aux variables prises en compte dans la fonction d'utilité. De plus, si cette méthodologie est acceptable pour des données relativement agrégées, elle ne l'est plus pour des données nécessitant un plus grand niveau de désagrégation. Par exemple, la prévision d'une répartition modale à court terme à l'échelle de l'agglomération sera probablement suffisamment précise. En revanche, la production d'une matrice origines-destinations de flux automobile pour estimer le trafic sur un lien sera très probablement beaucoup plus sujette à caution (car la matrice en référence contient en général beaucoup de 0).

Une seconde solution consiste à construire un échantillon. Cet échantillon est ensuite utilisé pour l'énumération en appliquant l'équation 9.102. La construction de cet échantillon nécessite la connaissance de la distribution de la population relativement à chacune des variables prises en compte dans la fonction d'utilité à l'horizon d'étude. L'idéal est évidemment de disposer de ces distributions pour chacune des zones de l'aire d'étude relativement au découpage retenu pour la modélisation. L'information est toutefois rarement disponible en référence à un niveau spatial très fin, il est donc peu probable qu'elle le soit à l'horizon d'étude. La procédure est ensuite assez simple, puisqu'il s'agit d'estimer la matrice la plus probable satisfaisant aux marges constituées par les distributions spatiales relativement à chacune des variables prises en compte dans la fonction d'utilité. Certaines des méthodes présentées dans le chapitre sur la distribution des déplacements peuvent être utilisées. Le

plus simple étant de partir de la distribution en référence et d'appliquer des facteurs de croissance pour obtenir la matrice à l'horizon. La méthode sera d'autant plus précise que les données en référence sont précises et que les distributions à l'horizon d'étude (c'est-à-dire les marges à l'horizon d'études) sont connues avec précision.

Lorsque les méthodes précédentes ne sont pas applicables, on pourra segmenter la population selon certaines variables socio-économiques et calculer la part de marché moyenne de l'alternative i sur chacun des segments. La part de marché agrégée est alors obtenue en faisant une moyenne pondérée sur chacun des segments. La précision de la méthode dépendra bien évidemment de la pertinence des variables de segmentation relativement au problème étudié et du nombre de segments ainsi défini. Lorsque ce nombre augmente et que l'on peut procéder également à une segmentation spatiale, cette méthode s'apparente à la précédente.

6. PRINCIPALES LIMITES DE L'APPROCHE DESAGREGEE

Nous adoptons un plan différent de celui des chapitres précédents. Une bonne partie des critiques formulées pour l'approche agrégée du choix modal peut en fait être reprise pour l'approche désagrégée. Il est donc plus intéressant de positionner les deux approches par rapport à ces différentes dimensions (section 6.1). Nous souhaitons toutefois insister sur certaines dimensions fréquemment problématiques dans la formalisation ou l'estimation des modèles désagrégés : la perception des attributs du déplacement (*taste variation* section 6.2), l'implication du zonage sur la définition du temps de déplacement (section 6.3), le choix d'une structure de modèle (section 6.4) et enfin, le problème spécifique de la prévision avec le modèle désagrégé (section 6.5).

6.1. Approche agrégée ou désagrégée

Nous reprenons les dimensions du chapitre précédent qui s'organisent autour des trois conditions de l'opérationnalité des modèles :

- pertinence : pertinence des variables explicatives du choix du mode (section 6.1.1), formalisation de la fonction d'utilité (section 6.1.3), « choice axiom » (section 6.1.4), stabilité des coefficients de calage (section 6.1.5) et segmentation selon les individus et/ou les motifs (section 6.1.6) ;
- cohérence d'objectifs : capacité à répondre aux problématiques de planification actuelles (section 6.1.2) ;
- cohérence interne : quelle valeur pour le coût généralisé (section 6.1.7) ;
- mesurabilité : qualité des données et notamment absence de données sur le stationnement (section 6.1.8).

6.1.1. Pertinence des variables explicatives du choix du mode

Il est indéniable que le modèle désagrégé présente une plus grande pertinence que le modèle agrégé. En s'appuyant sur les caractéristiques spécifiques de l'individu, il n'a pas besoin de faire référence à un individu moyen fictif doté de caractéristiques moyennes. L'approche désagrégée utilise donc la totalité de l'information disponible dans la procédure d'estimation. Elle permet de tirer profit de l'ensemble de la variabilité présente dans les données d'enquête.

L'estimation du modèle désagrégé étant moins sujette à des problèmes de corrélation entre les variables que l'estimation du modèle agrégé (Ortuzar, Willumsen, 1994), il est généralement possible d'avoir une définition plus fine des variables. Par exemple, dans un modèle agrégé, les attributs du déplacements sont souvent réunis dans une même variable, le coût généralisé qui comporte des coefficients qui ne sont pas estimés en même temps que la fonction d'utilité. A l'inverse, l'estimation du modèle désagrégé autorise une décomposition des éléments constitutifs du temps généralisé afin

que les coefficients permettant de les pondérer soient estimés en même temps que l'ensemble des paramètres entrant dans la fonction d'utilité.

Toutefois, si le modèle désagrégé offre une plus grande pertinence relativement aux variables prises en compte dans la fonction d'utilité, les limites évoquées pour l'approche agrégée (chapitre 8, p. 240) restent entières. Seule une partie des variables identifiées comme influant sur le choix du mode (chapitre 8, p. 220) sont prises en compte dans la fonction d'utilité. De ce point de vue, l'approche désagrégée offre peu d'avantages vis-à-vis de l'approche agrégée.

6.1.2. Prise en compte des problématiques de planification actuelles

Le modèle désagrégé permet de définir de manière plus fine les différents attributs du déplacement dans la fonction d'utilité. De plus, les coefficients peuvent être estimés sur les données disponibles. De ce fait, la fonction d'utilité est beaucoup plus sensible à des politiques d'offre diversifiées. Les modèles désagrégés sont donc en principe beaucoup plus souples et sensibles pour répondre aux problématiques actuelles de planification. C'est d'ailleurs l'un des principaux arguments avancés par les promoteurs de l'approche désagrégée (Ortuzar, Willumsen, 1994).

6.1.3. Formalisation de la fonction d'utilité

Les remarques formulées pour l'approche agrégée s'appliquent de manière similaire à l'approche désagrégée (chapitre 8, p 241).

6.1.4. « Choice axiom » et propriété IIA

Nous avons voulu conserver le titre du chapitre 8, même si le « *choice axiom* » est spécifique à l'approche agrégée. Sa traduction dans l'approche désagrégée correspond à l'hypothèse IID (distribution indépendante et identique des résidus) sur les résidus. En revanche, la propriété IIA est bien commune. La solution à la violation de cette propriété passe bien souvent par la structure emboîtée comme dans l'approche agrégée. Les hypothèses nécessaires ne sont toutefois pas toujours respectées. Il est alors nécessaire d'avoir recours à des structures de modèles beaucoup plus complexes qui rendent l'estimation du modèle tout aussi complexe.

S'il est indéniable que l'approche désagrégée offre des réponses à ces problèmes, ces développements restent toutefois, pour l'instant, principalement confinés dans les milieux de recherche, probablement faute de logiciels commerciaux adaptés.

6.1.5. Stabilité des coefficients de calage

La stabilité dans le temps des coefficients de calage est un problème récurrent qui est présent pour chacune des étapes des modèles à quatre étapes, tant pour l'approche agrégée que désagrégée. Les réponses possibles sont tout à fait similaires entre les deux approches.

6.1.6. Quelle segmentation de la demande ?

Le problème se pose de manière très différente pour les deux approches. En prenant en compte, dans la fonction d'utilité, les caractéristiques de chacun des individus, la segmentation sur les individus n'a a priori pas lieu d'être dans le modèle désagrégé. A l'inverse, dans le modèle agrégé, la segmentation permet de réduire les biais induits par la fiction de l'individu moyen. En segmentant la population, on cherche à définir des classes au comportement plus homogène relativement à l'objet d'étude.

En fait, cette réponse n'est que partielle. Même si les caractéristiques socio-économiques de l'individu sont prises en compte, on suppose que tous les individus perçoivent de manière identique les différents attributs pris en compte dans la fonction d'utilité. On définit donc une perception moyenne. Il est pourtant évident que pour certaines variables comme le prix par exemple, la perception n'est pas uniforme. Ce problème est identifié dans la littérature anglo-saxonne comme « *taste variation* ». Nous revenons sur cette question en section 6.2. La segmentation de la population par exemple par classe de revenus peut constituer une solution dans ce cas (section 3.2.7, p. 284).

6.1.7. Quelle valeur pour le coût généralisé ?

Le problème évoqué dans le chapitre précédent se pose dans des termes identiques pour l'approche désagrégée (chapitre 8, page 242).

6.1.8. Qualité des données

La qualité des données reste un problème crucial quelle que soit l'approche développée et l'étape traitée. Elle nécessite donc une attention toute particulière dans tout modèle (chapitre 8, p 243).

6.2. « *Taste variation* », *variation de perception*

Nous revenons sur cette question, bien que nous l'ayons déjà traitée dans ce chapitre, car elle constitue une des limites importantes de certaines familles de modèles désagrégés. Elle découle de la seconde hypothèse nécessaire à l'utilisation des modèles logit (section 2.4.1, p. 253). Elle stipule que tous les individus ont la même perception des différentes variables prises en compte dans la fonction d'utilité. Il est clair que cette hypothèse est souvent assez forte. Les temps ou les prix par exemple ne sont pas perçus par tous les individus de manière identique.

Une première solution consiste à segmenter la population, par exemple sur les revenus pour tenir compte de la perception des prix (section 6.1.6, p. 298). Cette solution peut toutefois s'avérer insuffisante. Il est alors nécessaire de recourir à des formulations plus complexes de l'utilité conduisant à la famille des logit multinomiaux mixtes (MMNL, section 2.7, p. 268) ou aux modèles probit (section 2.3, p. 251). Il est alors possible de postuler une distribution de la perception de certains attributs au sein de la population. On utilise alors le plus souvent des distributions normales ou log-normales (notamment pour la perception de la valeur du temps ou des prix).

Cette solution conduit toutefois à des modèles beaucoup plus complexes à estimer et à valider.

6.3. Approche désagrégée et choix du zonage

L'avantage du modèle désagrégé tient dans la possibilité de caler le modèle sur des données individuelles. Il est donc possible de prendre en compte les caractéristiques spécifiques de l'individu considéré et du déplacement qu'il réalise. Cette affirmation que l'on retrouve chez tous les promoteurs des modèles désagrégés néglige toutefois l'indétermination relative au temps d'accès aux modes de transport. Le choix d'un zonage ne permet pas de connaître précisément la localisation de l'origine et de la destination du déplacement. Il est donc nécessaire de définir un temps d'accès moyen aux réseaux. Si l'approximation est acceptable pour la voiture compte tenu de la capillarité des réseaux (encore que le temps de marche dépendant du lieu de stationnement est en général inconnu), elle ne l'est plus pour les transports collectifs. Le temps de marche à pied pour accéder aux arrêts de transports collectifs est d'autant plus important qu'il est ressenti très négativement (le coefficient affectant le temps de marche est fréquemment au moins double de celui du temps de circulation en véhicule).

Chapleau et De Cea (1983) ont ainsi montré que les erreurs commises à un niveau désagrégé lors de l'affectation de la demande sur les réseaux de transports collectifs sont extrêmement grandes y compris avec un zonage très fin (tableau 4.2, p. 128), même si à un niveau agrégé les résultats peuvent être corrects. Une information en coordonnées x-y permet une nettement meilleure reconstitution des comportements individuels. Il est très probable que l'erreur ne se limite pas à l'affectation, mais concerne également le choix du mode. La connaissance des aires de chalandise des arrêts de transports collectifs indique un rayon de 300 à 500 mètres (pour du bus, du tramway ou du métro). Si le zonage n'est pas très fin, il est très probable que dans la réalité des individus doivent considérer qu'ils n'ont pas accès aux transports collectifs (en grande couronne par exemple), même si la distance de marche moyenne pour la zone prise en compte par l'analyste dans le modèle stipule le contraire. Le temps de marche est donc biaisé, conduisant à une estimation biaisée de l'utilité pour l'individu et donc à une estimation biaisée du partage modal. C'est la raison pour laquelle Chapleau (1991, 1993, 1997) qualifie son approche de totalement désagrégée par opposition aux modèles désagrégés qui ne le sont pas pour la définition du temps d'accès aux réseaux.

Cette limite est évidemment d'autant plus importante que les zones sont étendues, car l'indétermination sur le temps d'accès croît avec la taille de la zone. D'autres paramètres sont toutefois à prendre en compte. La répartition de la population et des activités dans la zone ainsi que sa position par rapport aux arrêts de transport en commun, influent sur le temps d'accès en marche. Le maillage du réseau augmente les possibilités d'itinéraires concurrents et donc de mauvaise affectation lorsque le temps d'accès est biaisé... La prise en compte de ces éléments permet d'estimer une distribution de la distance d'accès au réseau au sein de la population des déplacements ayant une origine ou une destination dans la zone concernée. La connaissance de cette information permet de proposer une piste pour réduire l'incidence du biais. Cette solution passe par une structure de modèle plus complexe que le logit, dans laquelle le coefficient relatif au temps d'accès n'est plus unique, mais suit une certaine distribution. Cela nous conduit vers des structures de modèles similaires aux logit multinomiaux mixtes (MMNL, section 2.7, p. 268). Comme toujours ce gain en pertinence se paie au prix d'une estimation et d'une validation plus complexes.

Nous revenons sur cette question en l'élargissant à l'estimation du coût généralisé dans le chapitre suivant (p. 325).

6.4. *Quelle structure de modèle ?*

Nous avons déjà largement évoqué cette question dans ce chapitre. Nous y revenons, car elle constitue une bonne illustration de la problématique des modèles.

Le modèle logit désagrégé est très couramment utilisé (dans les pays anglo-saxons, car en France, la pratique des modèles désagrégés est encore assez limitée) en raison de sa très grande simplicité. Il repose toutefois sur des hypothèses fortement simplificatrices sur les résidus dont la pertinence n'est pas assurée. Le logit emboîté constitue une première réponse à cette pertinence douteuse. Cependant, si les hypothèses sont légèrement assouplies, elles restent encore souvent assez fortes. L'abandon des hypothèses les plus fortes pour répondre au problème de pertinence conduit alors à des structures de modèles beaucoup plus complexes comme les modèles probit ou les modèles logit multinomiaux mixtes (sections 2.3, p. 251 et 2.7, p. 268). Toutefois, ce gain en pertinence conduit à des problèmes redoutables de cohérence et de mesurabilité. Pour éviter une structure complètement libre des résidus qui conduit à une trop grande complexité du modèle, on pose généralement certaines hypothèses sur cette structure (hypothèses qui sont toutefois moins fortes que pour le logit). Toutefois, pour valider ces hypothèses, on est souvent confronté à des problèmes de mesurabilité, par exemple lorsque l'on postule une distribution normale ou log-normale d'un coefficient de la fonction d'utilité. La complexité de l'estimation peut également conduire à des coefficients de signe ou d'ordre de grandeur non attendus désignant souvent des problèmes de cohérence interne (ou bien de pertinence de certaines hypothèses).

C'est la raison pour laquelle, l'avancé théorique constituée par les modèles logit multinomiaux mixtes n'a pas encore débordé du cercle des chercheurs, anglo-saxons principalement. La complexité de l'estimation et de l'interprétation sont évidemment un frein important à leur diffusion. Mais il n'est également pas certain qu'ils constituent une avancée importante en termes d'opérationnalité des modèles. La facilité de validation et la robustesse d'un modèle sont des atouts importants en prévision, même si elles peuvent conduire à une moindre sensibilité face à certaines dimensions des politiques de transport.

6.5. Prévision

Nous avons également abordé cette question en section 5 (p. 293). Le modèle désagrégé peut être estimé sur des données individuelles. En revanche, son utilisation en prévision ne peut généralement plus s'appuyer sur des données individuelles. La prévision conduit donc toujours à un processus d'agrégation. Cet aspect est important lorsqu'il s'agit de comparer l'opérationnalité relative des modèles agrégés et désagrégés.

Chapitre 10 – Affectation, zonage et codification des réseaux

Avec ce chapitre, nous abordons la dernière étape des modèles à quatre étapes. Après avoir calculé la matrice origines-destinations des déplacements pour chacun des modes, cette étape correspond à l'affectation de la demande de déplacements sur les réseaux, essentiellement viaire et de transports collectifs. Dans la séquence des choix de l'individu, cette étape correspond au choix de l'itinéraire pour se rendre d'une origine à une destination données et avec un mode de déplacement donné.

Dans ce chapitre, nous présentons le zonage et la description des réseaux. Il peut paraître surprenant d'aborder cette question seulement dans le dernier chapitre de cet ouvrage, alors que la lecture du schéma de procédure de prévision (schéma 2.1, p. 34) positionne ces éléments en amont des quatre étapes des modèles. Il est d'ailleurs assez fréquent que cette question soit présentée dans un chapitre commun avec les données. Notre choix mérite donc justification. Il s'explique par la séquence des étapes nécessaires lors du calage du modèle qui positionne normalement l'affectation en premier (section 1). Nous abordons ensuite les grands principes présidant au choix du zonage (section 2) et à la description des réseaux viaires et de transports collectifs (section 3). Nous pouvons alors revenir à la structure des chapitres précédents. La présentation des concepts de base occupe ainsi la section 4. Nous examinons ensuite les principales familles de modèles (section 5), pour conclure avec les principales limites de l'affectation de la demande (section 6).

1. POSITIONNEMENT DE L'AFFECTION DANS LA CONSTRUCTION ET L'UTILISATION D'UN MODELE A QUATRE ETAPES

La principale sortie du modèle d'affectation consiste à produire la charge de trafic sur les différents réseaux, principalement viaire et transports collectifs. Comme nous le mentionnons plus loin dans ce chapitre (section 4.2, p. 321), les variables de description de réseaux permettent de déterminer l'incidence de la charge des réseaux sur le temps de parcours par l'intermédiaire de courbes débit-vitesse ou débit-coût. Ces courbes permettent de prendre en compte l'incidence de la congestion sur les temps de parcours et donc sur les choix d'itinéraires. Ce calcul des temps avec les réseaux chargés effectué par le modèle d'affectation permet donc de calculer les matrices origines-destinations de coûts généralisés par mode.

Dans la présentation des modèles de distribution et de choix du mode, nous avons pourtant utilisé ces matrices comme données d'entrée des modèles. Dans le même temps, le modèle d'affectation utilise comme données d'entrée, les matrices de demande produites en sortie du modèle de choix modal. Il y a donc une incohérence évidente dans l'enchaînement des étapes, car les matrices de coûts généralisés ne peuvent être disponibles en entrée du modèle de distribution, car elles ne sont produites qu'après l'utilisation du modèle d'affectation.

Pour lever cette incohérence, dans la pratique, on commence le calage du modèle à quatre étapes par le modèle d'affectation en utilisant comme matrices de demande en entrée du modèle, non pas les matrices issues des étapes de génération-distribution-choix modal, mais les matrices de référence issues des données d'enquête sur les déplacements. Ces matrices pour chacun des modes sont affectées sur les réseaux, ce qui permet de construire une première matrice origines-destinations de coûts généralisés pour chacun des modes. Ces matrices sont utilisées en entrée des modèles de distribution et

de choix du mode. En sortie de ces deux modèles, on dispose de nouvelles matrices estimées par le modèle (de génération-distribution-choix modal). Il est alors possible de réaffecter ces matrices sur les réseaux. Cette procédure est schématisée dans le schéma 10.1.

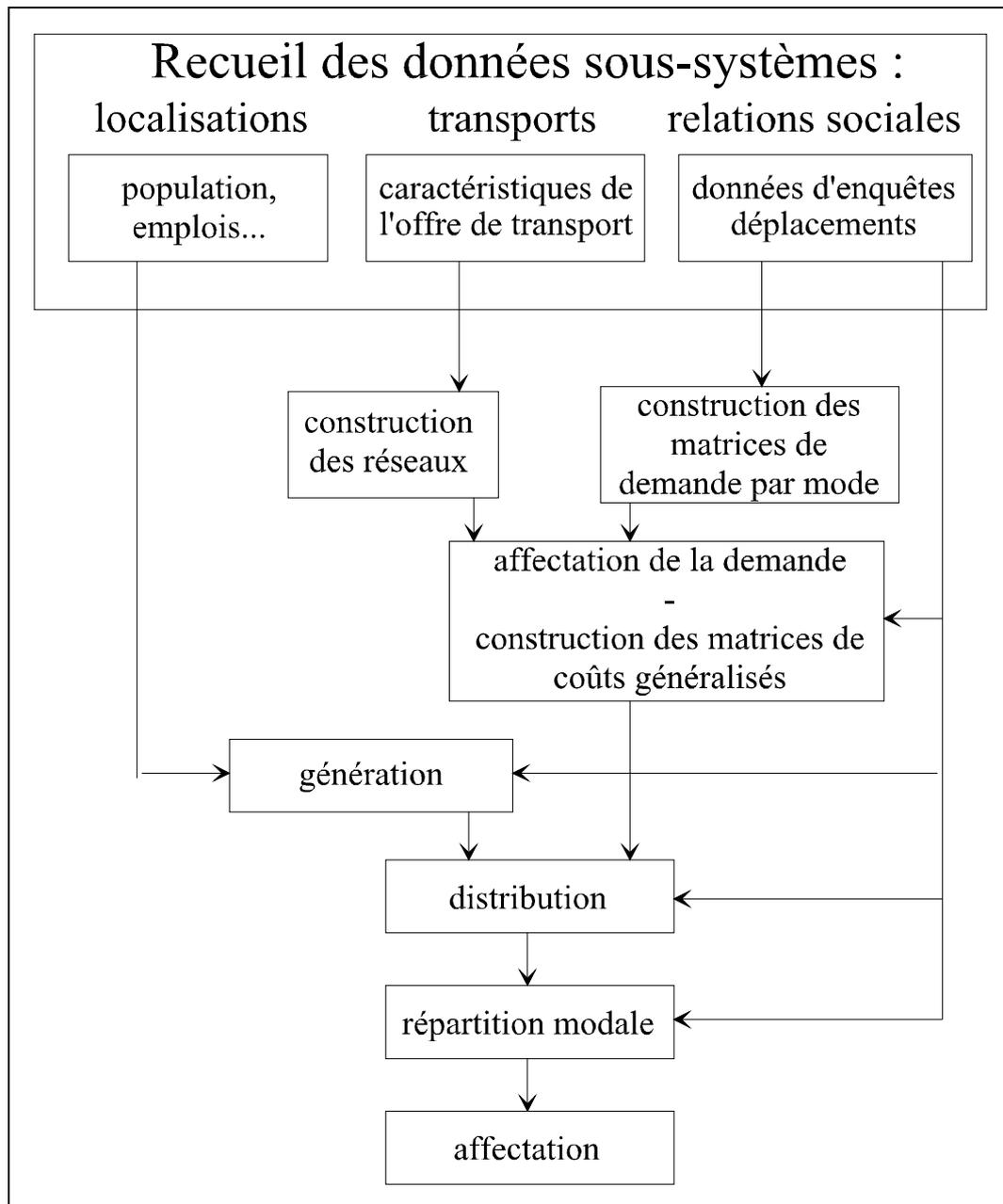


Schéma 10.1 : Séquence des étapes lors du calage du modèle à quatre étapes

L'utilisation en prévision pose un problème similaire. La modification des réseaux conduit à modifier les matrices de coûts généralisés et donc les matrices de demande par mode. De même, la modification du sous-système des localisations conduit à modifier les matrices de demande et donc les matrices de coûts généralisés. Mais, comme précédemment, les matrices de coûts généralisés sont produites par l'étape d'affectation qui nécessite en entrée les matrices de demande par mode qui elles-mêmes nécessitent en entrée les matrices de coûts généralisés.

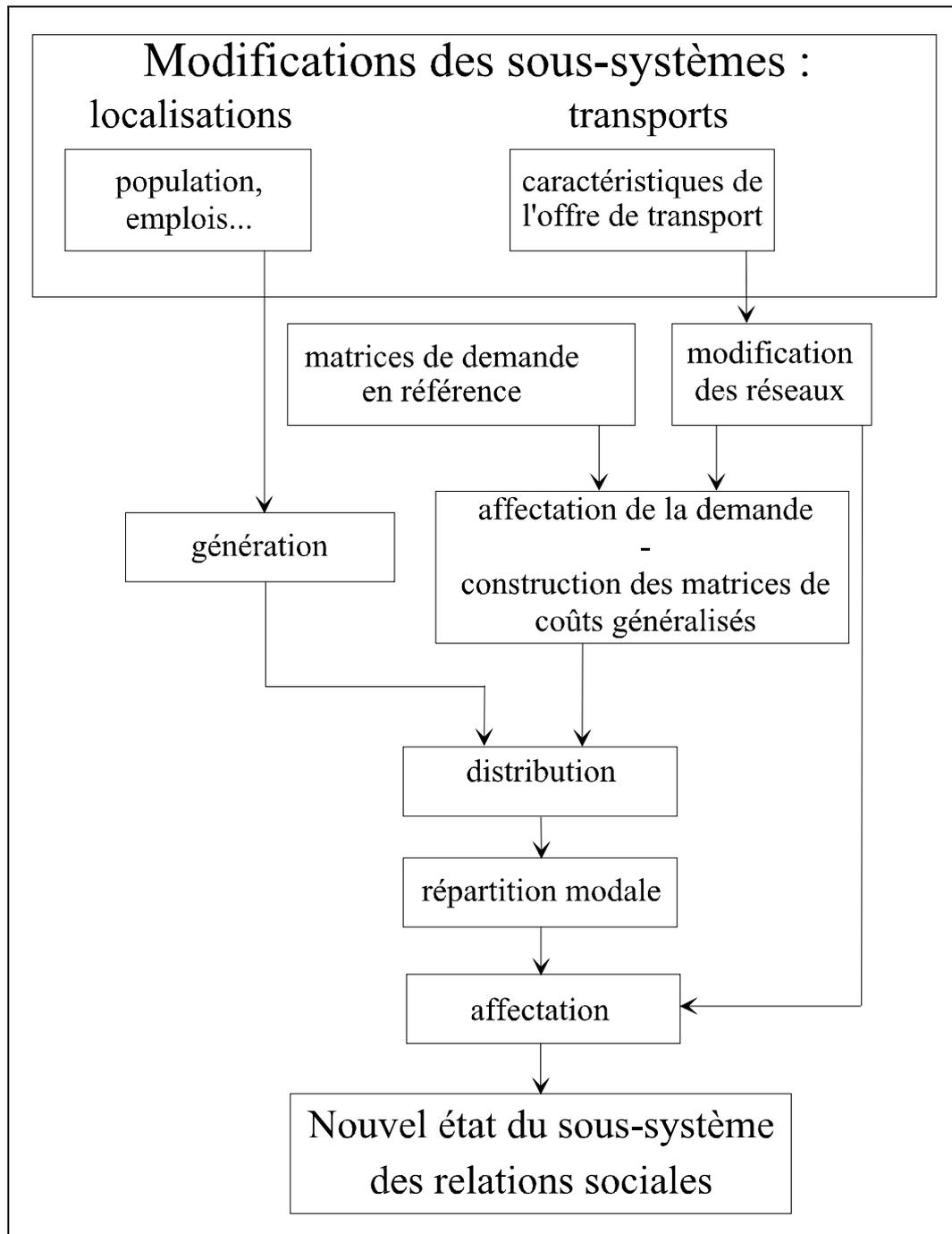


Schéma 10.2 : Séquence des étapes lors de l'utilisation du modèle à quatre étapes en prévision

Pour solutionner le problème comme précédemment, on commence par une première utilisation du modèle d'affectation pour produire de nouvelles matrices de coûts généralisés. Le calage du modèle a permis de produire des matrices de demande origines-destinations par mode relatives à la situation de référence (c'est-à-dire qui correspond à la situation repérée par les enquêtes déplacements). Ce sont ces matrices qui sont utilisées en entrée pour la première utilisation du modèle d'affectation. On peut ensuite dérouler les quatre étapes successives et prendre en compte les modifications en entrée des sous-systèmes de localisations et/ou de transports pour simuler leurs effets sur le sous-système des relations sociales. Cette procédure qui correspond au schéma 1.2 (p. 31) peut être reprise dans le schéma 10.2 en intégrant le modèle à 4 étapes.

Une solution alternative pourrait être de modifier les matrices de demande pour tenir compte de la modification du sous-système des localisations en retenant dans un premier temps les matrices de coûts généralisés estimées lors de la phase de calage. Il est alors nécessaire de dérouler les étapes de génération-distribution-répartition modale pour produire ces matrices, puis l'étape d'affectation pour estimer les nouvelles matrices de coûts généralisés qui tiennent également compte de la modification du sous-système de transports. Ces nouvelles matrices de coûts généralisés sont alors utilisées en entrée d'une nouvelle séquence des 4 étapes. Cette seconde solution étant plus lourde à mettre en œuvre, on lui préfère en général la première décrite dans le schéma 10.2.

Cette procédure permet donc de lever l'incohérence évoquée ci-dessus. En fait, elle ne la lève pas complètement, car il est improbable que les matrices de coûts généralisés utilisées en entrée du modèle de distribution restent stables à l'issue du modèle d'affectation. Le problème peut être négligé lorsque les matrices de coûts généralisés évoluent peu ainsi que les réseaux en prévision. Il ne peut, en revanche, être négligé dans le cas contraire. Dans ce cas, il est nécessaire de reprendre de manière itérative l'ensemble de la chaîne distribution-répartition modale-affectation avec les matrices de coûts généralisés calculées à l'issue de l'étape d'affectation de l'itération précédente jusqu'à une stabilisation des matrices des coûts généralisés.

Tant le calage du modèle que son utilisation en prévision nécessitent une procédure itérative pour s'assurer de la cohérence interne du modèle au niveau des matrices de coûts généralisés. L'expérience montre que la convergence de cette procédure itérative n'est pas toujours très rapide. Il est fréquent d'avoir plusieurs dizaines d'itérations pour obtenir une convergence suffisante. De plus, la convergence n'est pas garantie. Dans certaines configurations, on peut avoir un processus d'oscillations des résultats sans convergence. Dans ce cas, il est nécessaire de mettre en place des procédures « forçant » la convergence. Certains algorithmes que nous présentons dans ce chapitre peuvent être utilisés à cet effet, notamment l'algorithme des moyennes successives (section 5.5, p. 354).

Cette procédure n'est toutefois pas toujours mise en place pour plusieurs raisons :

- lors du calage du modèle, la mise en place de la procédure itérative conduit à une modification des matrices de coûts généralisés en entrée des modèles de distribution et de répartition modale et des matrices de demande en entrée des modèles de répartition modale et d'affectation. Si l'on conserve les coefficients de calage de l'itération précédente, l'utilisation de ces nouvelles matrices conduit à des résultats qui s'éloignent de la situation de référence. La procédure itérative impose donc de remettre en cause les coefficients de calage lors de chacune des itérations (dans la pratique, la remise en cause se limite généralement aux premières étapes de la procédure itérative tant que les matrices évoluent fortement). Cette procédure peut donc être rapidement très lourde à mettre en place et donc très coûteuse ;
- l'utilisation de procédure itérative en prévision signifie que l'on simule une situation à l'équilibre. Mais la notion même d'équilibre est critiquée par certains auteurs (Goodwin, 1984, chapitre 5, p. 159) qui estiment que le système urbain n'est jamais en situation d'équilibre stabilisée, mais au contraire que le système urbain est un système en perpétuelle dynamique. Le problème concerne plus particulièrement l'étape de distribution. Des changements importants des matrices de coûts généralisés produisent fréquemment des évolutions fortes des matrices de demande à l'issue de l'étape de distribution. Pourtant, une partie des choix de localisation est relativement contraint, tout au moins à court ou moyen terme (pour les emplois par exemple). Une déformation trop forte de la matrice de demande n'est donc pas très crédible. Mais, la mise en place d'un processus itératif risque d'amplifier ces évolutions de la matrice de demande lors de l'étape de distribution. La conduite du processus itératif nécessite donc un examen attentif de l'évolution des matrices dont la maîtrise est loin d'être évidente.

L'analyse des schémas 10.1 et 10.2 montre que la définition des réseaux et le choix du zonage sont des étapes qui précèdent l'affectation tout en étant fortement liées à l'affectation. Le calage du modèle d'affectation consiste en fait principalement à vérifier la cohérence de la codification du réseau, comme nous le verrons dans ce chapitre. Il nous a donc semblé préférable d'étudier ces dimensions dans un même chapitre.

2. PERIMETRE ET ZONAGE DE L'AIRE D'ETUDE

L'utilisation d'un modèle nécessite de définir un périmètre d'étude au sein duquel sera développé le modèle. Le choix d'un périmètre revient en quelque sorte à définir un isolat (section 2.1).

Le zonage (section 2.2) est utilisé pour agréger certaines données socio-économiques (du ménage ou des emplois...) ou de transport (temps d'accès au réseau...) utilisées par le modèle. Parallèlement à la définition du zonage, on définit pour chaque zone un centroïde de zone. Ce centroïde est un point fictif de la zone (qui n'a pas besoin d'être physiquement localisé, si ce n'est pour des commodités de représentation graphique) qui concentre l'ensemble des caractéristiques de la zone. En définissant un zonage, on abandonne en fait toute information désagrégée sur le plan spatial à l'intérieur de la zone. Cela signifie qu'un individu ou un emploi... n'est plus localisé en un point précis de la zone, mais est situé au centroïde de la zone. Tout est fictivement concentré en un seul point appelé centroïde. Ce mode de représentation a des conséquences importantes sur la définition des temps d'accès aux réseaux. Le temps d'accès réel ne peut plus être défini pour chacun des déplacements. On définit donc un temps d'accès moyen identique pour tous les déplacements de la zone. Nous reviendrons plus loin sur cette simplification qui est commune aux modèles agrégés et désagrégés.

Le choix du périmètre d'étude et du zonage est une opération importante, car ils servent de base à l'ensemble des calculs opérés dans les quatre étapes. Il répond à deux impératifs principaux qui sont contradictoires. La précision des calculs incite à une codification la plus fine possible et à une aire étendue (la précision ne se propage toutefois pas à l'ensemble de la chaîne à 4 étapes, car dans ce cas, la matrice de déplacements devient moins précise avec davantage de zéros). En revanche, les considérations de coûts (temps de travail nécessaire et temps de calcul des ordinateurs) et de disponibilité des données militent pour une précision limitée. Le choix du périmètre d'étude et du zonage est donc un compromis entre des impératifs contradictoires.

2.1. Périmètre d'étude

La question du périmètre d'étude a été abordée rapidement dans le premier chapitre (p. 27). Le choix d'un espace-temps d'analyse conduit vers une famille d'outils de prévision. Il conditionne la nature des interactions prises en compte dans le modèle. Le choix du périmètre conduit en fait à considérer que ce qui est externe à ce périmètre est exogène. Toutes les évolutions de cet « extérieur » sont donc considérées comme étant indépendantes des évolutions qui prennent place à l'intérieur du périmètre. On définit donc un isolat qui n'interagit pas avec l'extérieur (ou pour être plus précis, l'action de l'extérieur sur ce qui est interne au périmètre peut être pris en compte dans le modèle, en revanche, la rétroaction n'est pas prise en compte). Cette définition est évidemment une approximation, puisque dans la réalité les interactions sont permanentes.

Le choix de cet isolat dépend des objectifs de l'étude d'une part et de l'analyse des interactions d'autre part. L'étude d'un projet limité dans l'espace (par exemple un carrefour ou une ligne de transports collectifs de quartier) pourra s'appuyer sur un périmètre limité. En revanche, l'analyse d'un axe d'agglomération ou d'une rocade importante conduira généralement à retenir l'ensemble ou tout au moins une grande partie de l'agglomération. Un autre exemple permet de préciser la portée du choix qui doit être effectué. Considérons une rocade qui supporte un important trafic d'échange et de transit. Si l'on considère que les caractéristiques de ce trafic sont totalement (ou tout au moins

fortement) indépendantes de la nature des déplacements réalisés par les résidents d'un certain périmètre, ce dernier peut être considéré comme périmètre d'étude. En revanche, si l'on estime que le choix de l'itinéraire de ce trafic d'échange et de transit peut dépendre des conditions de trafic régnant sur les infrastructures à l'intérieur d'un certain périmètre, dans ce cas, le choix du périmètre d'étude devrait intégrer l'ensemble des itinéraires concurrents. On peut reprendre l'exemple du viaduc de Millau (chapitre 2, p. 36). Dans cet exemple, il est évident que l'agglomération ne peut être considérée comme un isolat. La réalisation du viaduc influe directement sur les choix d'itinéraires d'une partie du transit Nord-Sud. Il y a donc interaction forte entre ce transit et le système de transport de l'agglomération. Plus généralement encore, si le choix du mode ou de la destination des individus effectuant le trafic d'échange peut être fortement dépendant des conditions de circulation, il pourra être nécessaire d'élargir le périmètre pour intégrer les résidents des zones concernées. L'analyse de la nature des interactions en présence est donc un préalable indispensable au choix du périmètre.

Ce choix s'effectue toutefois sous contrainte de coûts et de disponibilité des données. L'élargissement du périmètre s'accompagne souvent d'un alourdissement du travail nécessaire d'une part pour collecter les données, d'autre part pour mettre en place le modèle. De plus, certaines données peuvent être disponibles sur un espace restreint par rapport aux ambitions que l'on peut avoir. L'enquête ménages déplacements constitue un bon exemple. Pour certaines études ayant une portée d'agglomération, on pourra souhaiter avoir un périmètre correspondant au bassin d'attraction de l'agglomération (par exemple l'aire urbaine au sens de l'INSEE). Ce périmètre est toutefois généralement plus large que celui de l'enquête ménages déplacements. On doit alors choisir entre la restriction de l'aire d'étude au périmètre de disponibilité des données et la nécessité de poser des hypothèses relatives à l'espace pour lequel on ne dispose pas de données (par exemple un comportement similaire à celui des résidents des franges du périmètre de l'enquête ménages déplacements...).

L'espace considéré comme externe au périmètre d'étude est généralement segmenté en zones. Ce zonage est en principe beaucoup plus agrégé que celui retenu pour le périmètre d'étude. Dans la plupart des études, on se contente de partitionner cet espace en fonction des infrastructures structurantes permettant de pénétrer au sein du périmètre d'étude. Chaque zone représente un quadrant de l'espace correspondant à la zone d'attraction de l'axe de pénétration considéré. Ces zones sont ensuite utilisées pour prendre en compte le trafic d'échange et de transit. Dans certains cas, le zonage peut être plus fin, par exemple si une gare ou un aéroport est situé à l'extérieur de la zone d'étude et qu'il peut être important de l'isoler au regard des objectifs de l'étude, ou si l'on veut distinguer le trafic de courte et de longue distance....

2.2. Zonage de l'aire d'étude

Le choix du zonage doit permettre de répondre aux objectifs de l'étude avec le niveau de précision souhaité tout en tenant compte des contraintes de coûts et de disponibilité des données. Le compromis entre ces deux dimensions contradictoires s'effectue généralement en prenant en compte les aspects suivants :

- cohérence avec les objectifs de l'étude (section 2.2.1) ;
- cohérence interne : homogénéité des zones, prises en compte des frontières « naturelles »... (section 2.2.2) ;
- cohérence avec la définition des réseaux (section 2.2.3) ;
- mesurabilité : disponibilité des données, compatibilité des découpages... (section 2.2.4) ;
- coût et temps de calcul (section 2.2.5).

2.2.1. Prises en compte des objectifs de l'étude

Le choix du zonage dépend fortement des objectifs de l'étude. Il doit permettre de répondre aux questions que l'on se pose. Prenons un exemple pour préciser le propos. Si l'on souhaite prévoir la

fréquentation d'une nouvelle ligne de TCSP (Transports Collectifs en Site Propre), il est souhaitable de disposer d'un zonage permettant d'isoler la zone de chalandise du TCSP. Si l'étude doit également permettre de dimensionner chacune des stations, il est souhaitable d'avoir une zone par arrêt avec pour chacune des zones une étendue proche de la zone de chalandise de l'arrêt. Si l'étude porte également sur les correspondances, le zonage devra permettre d'isoler chacune des lignes en correspondance. Si l'on veut isoler l'apport de clientèle de certains générateurs spécifiques, il est souhaitable de pouvoir les isoler dans le choix du zonage... La modélisation d'une voirie rapide d'agglomération s'appuiera plutôt sur les échangeurs pour déterminer les zones de chalandises. Si l'on retient uniquement l'objectif d'étude comme critère, le degré de précision du zonage dépend donc des sorties attendues du modèle. L'identification précise des objectifs de l'étude est donc un préalable indispensable.

2.2.2. Cohérence interne : homogénéité des zones, prises en compte des frontières « naturelles »...

L'homogénéité des zones est un critère généralement important principalement pour les modèles agrégés. La nécessité de définir un individu moyen d'une zone conduit à une approximation d'autant plus forte que la zone est hétérogène. L'homogénéité doit être recherchée par rapport aux variables les plus explicatives des comportements. C'est généralement le cas de la motorisation, ce qui conduit à rechercher une certaine homogénéité des revenus ou des PCS (position et catégorie socio-professionnelle). C'est également le cas des activités présentes dans la zone (si les données le permettent, il n'est généralement pas souhaitable d'inclure dans une même zone une importante zone d'activités et une importante zone d'habitat).

En ce qui concerne les variables socio-économiques, le problème n'est pas de même nature pour un modèle désagrégé puisque le calage s'opère sur les caractéristiques de l'individu. Il ne disparaît toutefois pas complètement, puisque la prévision nécessite une procédure d'agrégation.

Les frontières naturelles sont généralement prises en compte dans le choix du zonage, car elles constituent en général un frein à la mobilité entre les espaces situés de chaque côté. Les fleuves, voies de chemin de fer, autoroutes urbaines ou autres frontières géographiques sont généralement utilisées comme frontières de zones. Il est généralement fortement déconseillé d'inclure ces frontières à l'intérieur d'une zone.

2.2.3. Prise en compte des réseaux

Le choix du zonage est généralement conditionné par les réseaux de transports collectifs. Nous l'avons déjà évoqué dans la section relative aux objectifs de la modélisation. Le choix du zonage doit ainsi être compatible avec les sorties attendues du modèle. Par exemple, le dimensionnement de stations de TCSP privilégiera le choix de zones correspondant aux zones de chalandise de ces stations.

Mais plus généralement, le zonage doit tenir compte de la structure des réseaux. D'une manière générale, pour les transports collectifs, il est souhaitable d'avoir des zones pour lesquelles le temps d'accès aux arrêts des lignes de transports collectifs est relativement uniforme. Dans les formules de coût généralisé, ce temps de marche à l'arrêt est généralement affecté d'un coefficient au moins égal à deux pour tenir compte de la pénibilité perçue de la marche par les individus. La prise en compte de ce temps à l'origine et à la destination représente ainsi une proportion non négligeable de la totalité du coût généralisé des déplacements en transports collectifs. Une trop forte dispersion des temps de marche au sein d'une même zone conduira à une distorsion pouvant être assez forte de la concurrence entre les transports collectifs et la voiture. Ce problème est commun aux approches agrégées et désagrégées puisque la désagrégation porte sur les caractéristiques socio-économiques individuelles et non sur les localisations au sein de la zone. Seule la prise en compte explicite d'une loi de dispersion des temps d'accès pour chacune des zones permet de contourner le problème, sauf à disposer d'une information sur les localisations en coordonnées x-y (Chapleau, 1991, 1993, 1997 ; Dorland et al., 1997 ; Chapleau, De Cea, 1983) dans le cas d'une approche désagrégée ou totalement désagrégée pour

reprendre l'expression de Chapleau. Le respect de cette règle n'est généralement pas possible sur l'ensemble de l'agglomération, car il conduirait à un nombre de zones beaucoup trop important. L'analyste doit donc déterminer les zones ou arrêts devant être privilégiés quant au respect de cette règle.

Le même principe doit guider le zonage en ce qui concerne le réseau viaire. Cependant, l'irrigation des destinations par le réseau viaire de proximité rend ce problème du temps d'accès beaucoup moins problématique que pour les transports collectifs.

Il faut généralement éviter des zones émettant ou attirant un trop grand nombre de déplacements au regard des infrastructures d'accès à la zone. Le chargement de l'ensemble de la demande sur un seul connecteur de centroïde (la définition de ce terme est donnée à la section 3.2.1, p. 313) peut conduire à des perturbations dans les choix d'itinéraires lors de l'affectation. Dans ce cas une division de la zone considérée peut être souhaitable.

Compte tenu de la densité des réseaux viaire ou de transports collectifs et de la répartition de la population et des emplois dans l'espace, les zones sont généralement plus réduites dans le centre et leur surface s'accroît à mesure que l'on s'en éloigne.

2.2.4. Disponibilité des données, compatibilité des découpages...

Les considérations évoquées ci-dessus militent pour un nombre de zones assez élevé. Celles que nous abordons maintenant conduisent généralement à réduire les ambitions.

Les données disponibles ne permettent pas toujours d'envisager un découpage très fin de l'agglomération. Pour pouvoir caler les modèles de génération, il faut disposer d'un effectif suffisant pour chacune des zones. La précision théoriquement accrue par un zonage plus fin serait totalement illusoire si le nombre de déplacements émis ou reçus par chacune des zones ne pouvait être estimé correctement. Lorsque l'on utilise les données d'une enquête ménages déplacements, la taille de l'échantillon conditionne bien souvent la précision du zonage (à moins de procéder à des agrégations de zones pour la génération par exemple).

Les variables socio-économiques caractérisant la population et celles relatives à l'occupation de l'espace (emplois, nombre de scolaires...) doivent être également disponibles à ce niveau de précision. Cela ne pose généralement pas trop de difficulté en situation de référence pour les données de population. Les données de recensement peuvent être disponibles à l'îlot qui constitue généralement une entité suffisamment fine pour les besoins de modélisation. Ce n'est généralement plus le cas pour les autres sources de données qui ne sont pas disponibles à un niveau aussi désagrégé. Il est enfin évident que la disponibilité des données à un niveau très fin est totalement illusoire à l'horizon d'études. L'expérience montre de plus que les intentions des planificateurs ne se traduisent pas toujours sur le terrain 10 ou 20 ans plus tard. Les données disponibles contraignent ainsi fréquemment à des regroupements de zones ou bien à des hypothèses quant à la répartition de certaines données entre plusieurs zones.

La compatibilité des zonages utilisés pour les différentes production de données ou études de planification est un problème incontournable. Chaque production ou étude produit son propre découpage répondant à des besoins spécifiques. De ce fait, les découpages sont différents sans possibilité de passer aisément de l'un à l'autre. Le découpage commun le plus fin est en général un découpage communal. Mais si ce découpage peut convenir pour les communes les plus périphériques, il n'est généralement pas compatible avec les besoins d'études pour les communes plus centrales.

Les données de recensement constituant le plus souvent une base indispensable, il est généralement recommandé d'opter pour un zonage compatible avec les îlots de l'INSEE. Enfin, il est généralement recommandé de respecter le découpage communal.

Cette contradiction entre les objectifs d'étude et les données disponibles conduit parfois à considérer différents niveaux de zonage lors d'un exercice de modélisation. Le zonage peut être ainsi « relativement grossier » pour l'étape de génération pour laquelle on ne dispose pas forcément des données socio-économiques à un niveau suffisamment fin, voire également pour celles de distribution et de répartition modale. En revanche, il sera beaucoup plus fin pour l'étape d'affectation pour laquelle une plus grande précision du zonage est souvent nécessaire. Mais dans ce cas, l'éclatement de la matrice de demande estimée sur un zonage « grossier » en une matrice de demande correspondant au zonage le plus fin est évidemment problématique.

2.2.5. Coût et temps de calcul.

Agglomération	Population (en million d'habitants)	Nombre de zones	Commentaires
Londres (1972)	7,2	≈2252	Niveau le plus fin
		≈1000	Zonage normal du <i>Greater London Transport Study</i>
		≈230	District du <i>Greater London Transport Study</i>
		52	<i>Traffic boroughs</i>
Ile de Montréal (1980)	2	≈1260	Zonage le plus fin
Ottawa (1978)	0,5	≈120	Zonage « normal »
Santiago (Chili) (1986)	4,5	≈260	Etude stratégique
Washington (1973)	2,5	≈1075	Zonage « normal »
		≈134	Au niveau du district
West Yorkshire (1977)	1,4	≈1500	Zonage le plus fin
		≈463	Zonage grossier
Lyon	≈1,4	400 à 500	Zonage CETE de Lyon étude routière
	≈1,4	196	Zonage SEMALY étude transports collectifs (SEMALY, 2000)
	≈1,4	25	Zonage modèle stratégique (Lichère, Raux, 1997a, b)

Tableau 10.1 : Exemples de zonages pour différentes études
(Source : Ortuzar, Willumsen, 1994, sauf données relatives à Lyon)

Le temps et le budget disponibles pour une étude conditionnent bien souvent la précision pouvant être développée. Le temps nécessaire à la production des données et au calage des modèles relatifs à chacune des étapes augmente plus que proportionnellement à mesure que le nombre de zones s'accroît. Il en est de même du temps de calcul des ordinateurs, encore que cette contrainte se réduise à mesure que leur puissance s'accroît.

Le nombre de zones résulte donc d'un compromis entre des objectifs contradictoires. A titre d'illustration, quelques exemples de zonages sont présentés dans le tableau 10.1 (Ortuzar, Willumsen, 1994, sauf données relatives à Lyon).

3. CODIFICATION DES RESEAUX

La codification des réseaux permet de représenter l'offre de transport qu'elle soit viaire ou de transports collectifs. Elle vise à représenter la performance des différents modes de transport du point de vue des usagers de ces réseaux. Cette représentation peut s'opérer à différents niveaux de précision (section 3.1). La codification s'appuie généralement sur la théorie des graphes. Les réseaux sont définis à partir de nœuds et de liens qui possèdent certains attributs (section 3.2). Pour les transports collectifs, il est nécessaire d'y adjoindre les lignes de transports collectifs avec leurs caractéristiques d'offre (section 3.3).

3.1. Niveau d'agrégation

Comme pour le zonage, l'analyste doit spécifier le niveau de précision nécessaire dans la description des réseaux. Les critères avancés pour le choix du zonage peuvent être assez facilement transposés pour la description des réseaux. Il y a d'ailleurs un parallèle assez fort entre les niveaux de précision du zonage et de la description des réseaux.

En théorie plus la précision est grande, plus la pertinence l'est aussi. Mais comme pour le zonage, une très grande précision de la description des réseaux peut conduire à une pertinence illusoire. Elle n'est de plus généralement pas nécessaire et génère des coûts de production de données et d'étude excessifs. Prenons un exemple pour illustrer le propos. Considérons une étude où l'on souhaite décrire la totalité du réseau viaire, y compris les rues les plus petites permettant de circuler. Pour être efficiente, cette précision doit être accompagnée d'une description réaliste de tous les mouvements aux carrefours. Chaque mouvement tournant doit être doté de caractéristiques traduisant la plus ou moins grande difficulté à l'effectuer. Si ce réseau est congestionné en heure de pointe, il faudra être en mesure de simuler les effets de remontée de queues de bouchons sur les infrastructures ou carrefours amonts... Le niveau de précision du réseau doit donc être cohérent avec la précision des attributs le définissant. Par ailleurs, le zonage conduit à définir des centroïdes de zones concentrant la totalité des informations de la zone et des origines et des destinations des déplacements. De ce fait, le trafic interne à une zone ne peut être représenté puisqu'il a son origine au centroïde et sa destination au même centroïde. Donc avec ce mode de représentation, il n'y a aucun mouvement interne. Ces déplacements occupent pourtant la voirie de la zone sans possibilité de le modéliser finement (il existe toutefois des possibilités d'en tenir compte, mais elles ne peuvent s'appuyer sur la description des réseaux). Une description très fine des réseaux combinée à un zonage assez grossier serait donc totalement inutile. Enfin, si l'objet de l'étude est de simuler le fonctionnement du réseau à un horizon de long terme, il est totalement illusoire de chercher à représenter aussi finement le réseau, car on sera dans l'impossibilité de le définir avec le même niveau de précision à l'horizon d'étude.

L'analyste doit donc définir le niveau de précision recherché. Celui-ci doit être cohérent avec d'une part la précision des attributs de définition des réseaux et d'autre part le zonage choisi. Il doit également être cohérent avec les objectifs recherchés. Dans une étude qui visera à simuler le fonctionnement d'un carrefour ou d'un échangeur autoroutier ou encore le fonctionnement d'un quartier fortement congestionné, une description extrêmement fine du réseau pourra être nécessaire. A l'inverse dans une étude visant à simuler différentes stratégies de politiques de déplacements urbains à moyen ou long terme, le mode de représentation pourra être beaucoup plus simple, voire comme dans le cas du modèle stratégique lyonnais s'appuyer sur une représentation fictive du réseau. Dans ce dernier modèle, chaque lien correspond en fait à une agrégation d'infrastructures (Lichère, Raux, 1997a, b). Le niveau de précision dépend enfin également du budget disponible, la construction d'un réseau pouvant s'avérer très coûteuse compte tenu des informations nécessaires pour la représentation

des réseaux. Deux exemples de codification sont présentés dans les schémas 10.3 et 10.4 pour illustrer le propos.

3.2. Caractéristiques des liens et des nœuds

La description d'un réseau pour les besoins de modélisation nécessite la description de sa structure (définition de ses composants, par exemple des sections de voies, des intersections...), des propriétés de ses composants (longueur, vitesse...) et des relations entre ces éléments et les flux de déplacements (Lamb, Havers, 1970).

La description des réseaux s'appuie sur la théorie des graphes. Les réseaux sont ainsi composés de liens directionnels et de nœuds. Ces liens et ces nœuds sont de plusieurs sortes (section 3.2.1) et possèdent un certain nombre d'attributs (section 3.2.2).

3.2.1. Définition des éléments composant un réseau

Un réseau comporte des nœuds et des liens. On distingue principalement deux types de nœuds : le centroïde de zone et le nœud de réseau. Il y a également principalement deux types de liens : les connecteurs de centroïde et les liens de réseau. A ces éléments se superpose la description des lignes de transports collectifs (section 3.3, p. 316).

Le centroïde de zone est un nœud fictif qui concentre tous les attributs de la zone dont il constitue en quelque sorte le centre de gravité. Il s'agit d'un point fictif introduit à des fins de représentation (donc de simplification de la réalité). Il n'a pas réellement de position géographique. Souvent, on le situe au centre de gravité de la zone, mais il ne s'agit que d'une commodité de représentation graphique qui n'a pas de sens en termes de distance métrique. Toutes les localisations de la zone (ménages, emplois, origines et destinations des déplacements...) sont en fait concentrées en ce point. A l'intérieur de la zone, toutes les distances sont donc considérées comme nulles. Il s'agit bien évidemment d'une approximation. Et c'est pour réduire la portée de cette approximation, que l'on introduit le connecteur de centroïde (que nous décrivons ci-dessous). Le centroïde de zone est doté d'un certain nombre d'attributs traduisant les caractéristiques de la zone. Par définition, il ne peut y avoir qu'un seul centroïde par zone.

Le nœud de réseau correspond à des intersections entre des liens de réseau. Il permet d'assurer la connexité du réseau. Il correspond d'une certaine manière au carrefour du réseau viaire, même si dans un réseau simplifié le nœud de réseau n'a pas forcément une traduction physique directe (schéma 10.3c).

Lorsque l'on veut prendre en compte le trafic d'échange ou de transit, il est fréquent que l'on définisse une troisième catégorie de nœuds. Ces nœuds sont appelés nœuds d'injection, car ils servent à injecter la demande provenant des zones externes au périmètre d'étude. Ces nœuds jouent en quelque sorte un rôle similaire à celui des centroïdes de zones. En principe, ils sont connectés au réseau par des liens de réseau (parfois par des liens fictifs ayant des caractéristiques différentes des liens de réseau, notamment pour avoir un coût généralisé nul ou constant sur ce lien fictif, appelé aussi lien d'injection). Ce type de nœud peut être également utilisé pour isoler un générateur ponctuel (par exemple un centre commercial) sans pour autant le définir comme une zone. Dans ce cas, la demande générée par ce nœud est en général estimée en dehors du modèle ou bien de manière spécifique (sinon, il suffit d'isoler le générateur et de définir une zone ne comprenant que ce générateur). Ce type de nœud est caractérisé principalement par la demande qu'il génère ou attire.

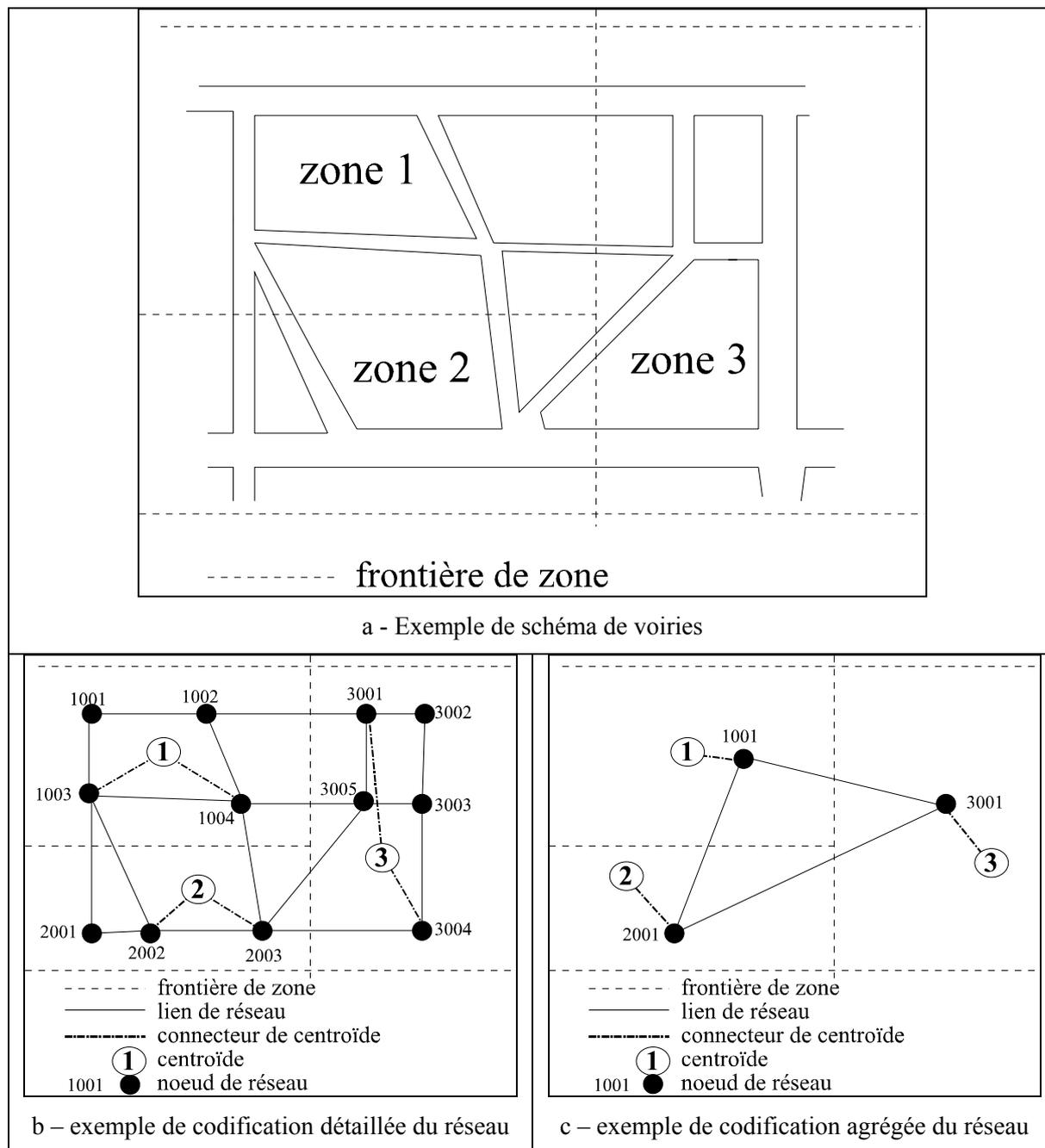


Schéma 10.3 : Exemples de codification de réseau

Le connecteur de centroïde est un lien fictif qui assure la liaison entre le centroïde de zone et le reste du réseau. Sa fonction est d'assurer la connexité du graphe de sorte que tous les centroïdes soient reliés au réseau par au moins un connecteur. Il est doté d'attributs permettant de déterminer le temps nécessaire pour relier le centroïde au réseau. Un centroïde peut être connecté au réseau par plusieurs connecteurs de centroïde.

Le lien de réseau correspond aux infrastructures de transports. La traduction peut être très proche de la réalité du réseau dans le cas d'une description fine du réseau (schéma 10.3b) ou s'en éloigner pour traduire davantage les fonctionnalités que les caractéristiques physiques dans le cas d'un réseau plus agrégé (schéma 10.3c). Le lien est doté d'un certain nombre d'attributs permettant de

traduire la fonctionnalité de la partie de réseau qu'il représente. En particulier, ces attributs doivent permettre de calculer le coût, ou temps généralisé, nécessaire pour parcourir le lien.

3.2.2. Attributs des liens

Les liens de réseau sont dotés d'attributs permettant de calculer le coût, ou temps généralisé, nécessaire pour parcourir les infrastructures que le lien représente. Généralement, on trouve les attributs suivants :

- la longueur du lien ;
- la vitesse à vide du lien (c'est-à-dire la vitesse en l'absence de toute circulation sur le lien). Cette information est parfois remplacée par le temps de parcours à vide, puisque la vitesse sert en fait à calculer le temps de parcours à partir de la longueur du lien ;
- la capacité du lien, cette capacité est généralement fournie en nombre d'unités de voitures particulières par heure (l'Unité de Voiture Particulière ou UVP est un indicateur permettant d'agréger différents types de véhicules pour tenir compte de leur occupation différenciée de l'espace de voirie. Par exemple, un poids lourds équivaut souvent à 2 à 3 UVP selon sa taille). Il s'agit de la capacité maximale pouvant être absorbée par le lien ;
- la tarification du lien qui permet de prendre en compte le péage. Le coût d'usage est, en revanche, généralement kilométrique et n'est donc pas défini au niveau du lien, lorsque la longueur du lien est disponible.

Selon les besoins et les objectifs de la modélisation, les attributs suivants peuvent être également définis :

- le type de voies : autoroutes, voies rapides, voies d'intérêt local, voies de desserte fine... Cette définition permet d'alléger le processus de description des attributs en donnant des attributs communs à l'ensemble des liens incluent dans un même type. La courbe débit-vitesse fait souvent partie de ces attributs. Elle est généralement identique pour l'ensemble des liens inclus dans un même type, mais différenciée selon le type de voies ;
- les mouvements possibles entre les liens connectés à chaque extrémité du lien. Cette information permet par exemple de représenter des interdictions de tourner à droite ou à gauche ;
- la nature du carrefour à chaque extrémité du lien : carrefour à feux, priorité à droite, stop... Dans ce cas, l'information est généralement complétée par un certain nombre d'attributs permettant de calculer le temps nécessaire pour traverser le carrefour, par exemple durée de cycle du feux et temps de vert disponible, lois d'écoulement pour un stop... Lorsque l'on souhaite avoir une représentation très détaillée du carrefour, on peut être conduit à retenir une représentation ne comportant pas un seul nœud, mais un ensemble de liens et de nœuds traduisant chacun des mouvements possibles (schéma 10.4) ;
- une loi de stockage de queue dans le cas d'une prise en compte assez fine des effets de la congestion sur l'écoulement des flux ;
- les véhicules autorisés à circuler. Cette information permet de spécialiser certaines infrastructures (voies réservées au bus, pistes cyclables...) ou d'interdire l'accès à certains types de véhicules (poids lourds, deux roues...).

Dans certains logiciels, les attributs relatifs au carrefour ne sont pas traités au niveau du lien, mais au niveau du nœud. Le principe de codification reste toutefois identique.

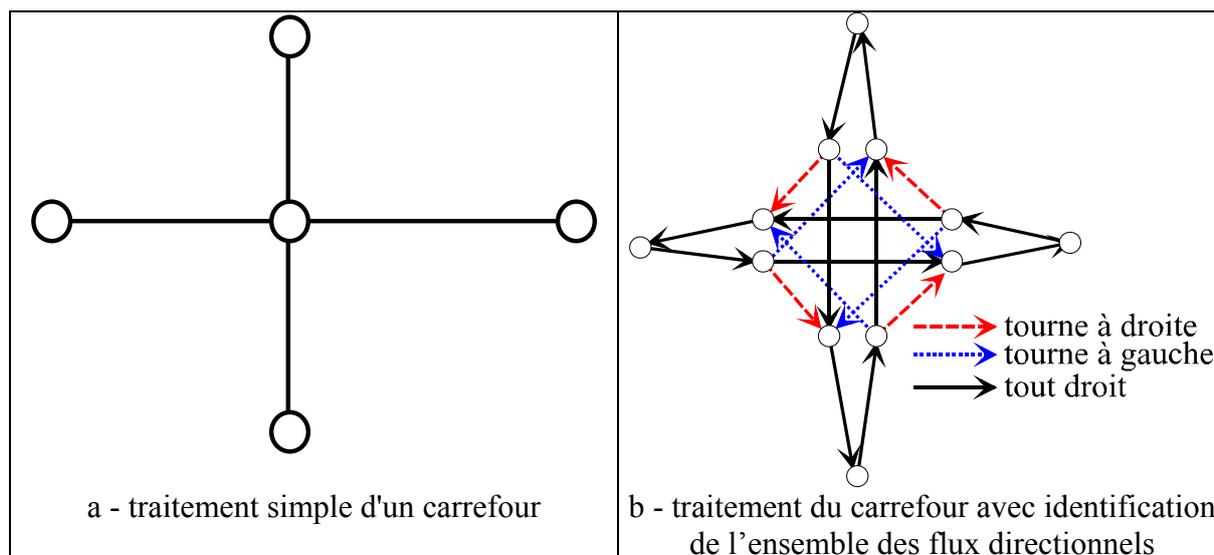


Schéma 10.4 : Exemples de codification d'un carrefour

Lorsqu'un lien représente plusieurs voies (comme dans le schéma 10.3c), la définition des caractéristiques du lien peut être problématique. Généralement, on additionne les capacités, mais pour la vitesse ou le temps de parcours l'opération n'est plus aussi simple. Si l'on fait une moyenne simple, on risque de distordre la fonctionnalité effective des infrastructures, notamment lorsque les deux (ou plus) voies ne supportent pas les mêmes trafics. On peut éventuellement faire une moyenne pondérée par les trafics supportés par chaque voie, mais encore faut-il disposer de l'information. De plus, en prévision rien ne garanti le maintien des poids respectifs. D'une manière générale, il est recommandé de privilégier la voie principale, qui traduit en général la fonctionnalité offerte entre les deux zones (comme dans le schéma 10.3, le lien entre les zones 1 et 3 privilégiera les caractéristiques du lien situé le plus haut sur le dessin 10.3a qui semble nettement plus important que le second situé dessous). Toutefois, il est recommandé de ne pas trop agréger des liens ayant des caractéristiques trop différentes.

Les connecteurs de centroïde possèdent un nombre d'attributs plus limités. Leur fonction est essentiellement d'assurer la connexité du réseau de telle sorte que tous les connecteurs de centroïde soient reliés au réseau. De ce fait, le temps généralisé de parcours du connecteur de centroïde est généralement supposé indépendant du trafic qu'il supporte. C'est pour cela qu'on les dote d'une capacité « infinie » ou plus exactement suffisamment élevée pour que le trafic supporté soit toujours inférieur à la capacité. Les connecteurs sont donc généralement définis par :

- le temps de parcours ou bien la distance et la vitesse qui permet de calculer le temps de parcours. Il n'est pas nécessaire de préciser qu'il s'agit d'un temps à vide car on présume une insensibilité à la charge de trafic sur ce type de lien ;
- la capacité définie pour être toujours supérieure au trafic.

Les nœuds de réseau ne possèdent généralement pas d'attributs. Leur fonction est d'assurer la connexité du réseau. Enfin, les centroïdes sont dotés d'attributs représentant les caractéristiques de la zone (population, emplois, motorisation...).

3.3. Codification des lignes de transports collectifs

La description du réseau de transports collectifs est de nature différente de celle du réseau viaire. Un réseau de transports collectifs est composé de lignes et le passage d'une ligne à une autre génère un coût ou temps généralisé supplémentaire qui n'a pas d'équivalent pour la voiture. Il est donc nécessaire de superposer au graphe une description des lignes de transports collectifs. Pour chaque

ligne, cette description consiste en une succession de nœuds de réseau correspondant à des possibilités de monter ou descendre de la ligne. Selon la finesse de représentation nécessaire, on peut isoler chacun des arrêts d'une ligne ou bien agréger plusieurs arrêts dans un même nœud (schéma 10.5).

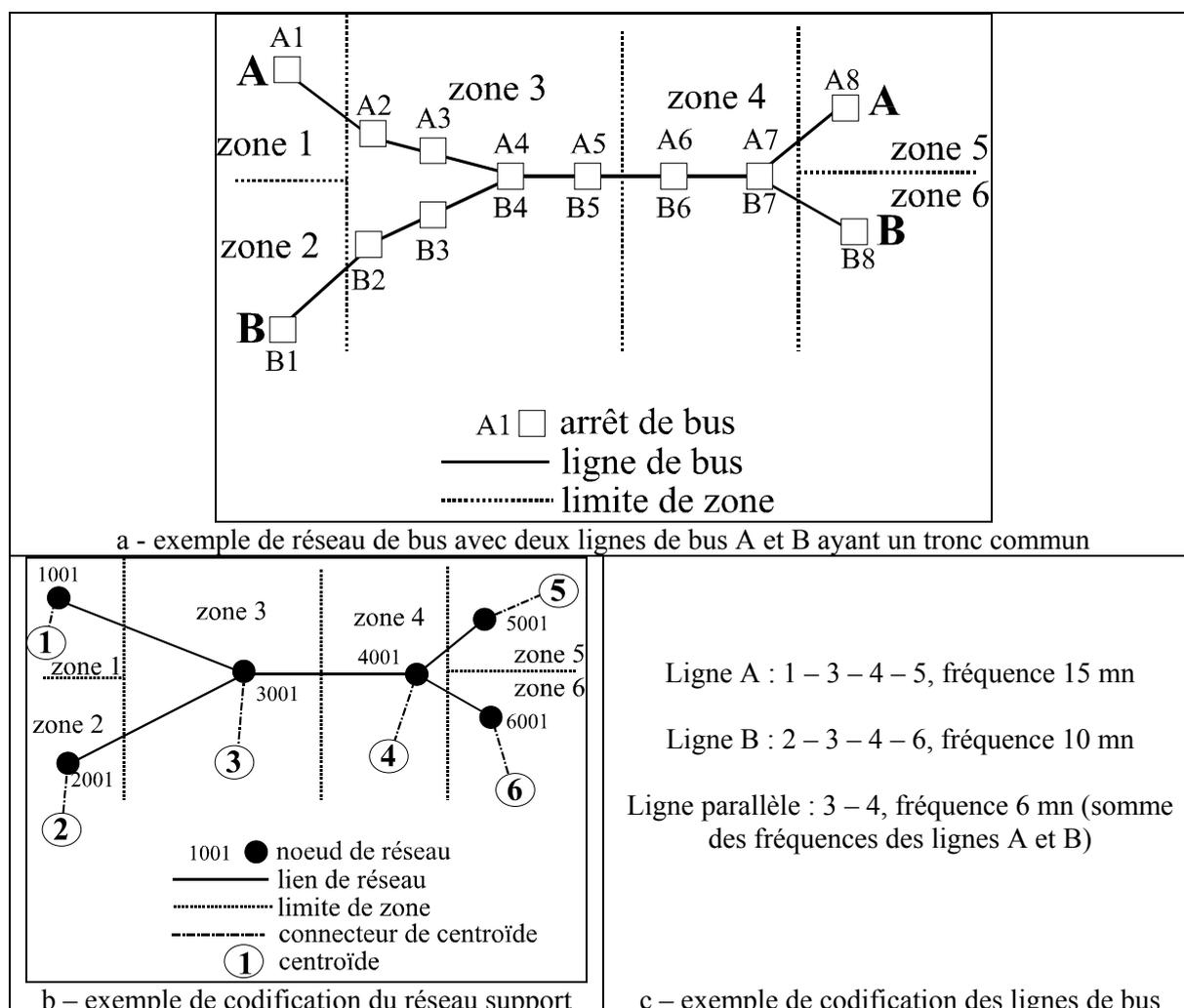


Schéma 10.5 : Exemples de codification de lignes de bus

Le graphe support de ce réseau peut être le graphe du réseau viaire. Toutefois, la description de lignes en site propre nécessite en général l'ajout de nœuds et de liens supplémentaires, car ces liens ne peuvent pas être utilisés par les autres modes de transport et sont bien souvent situés, au moins partiellement, en dehors du réseau viaire.

Une ligne de transports collectifs peut être dotée des attributs suivants :

- la fréquence ou le nombre de services par heure qui permet de calculer le temps d'attente. Il est parfois possible de spécifier plusieurs formules pour calculer le temps d'attente selon que l'on considère que les arrivées sont aléatoires (intervalles faibles entre deux véhicules) ou dépendantes des horaires de passage (intervalles élevés entre deux véhicules). Certains modèles (SEMALY, 2000) proposent d'intégrer un coefficient de régularité pour la ligne qui viendra moduler l'attente selon que la ligne est régulière ou non. La principale limite à ces améliorations du calcul du temps d'attente provient de la production de données nécessaire en situation de référence et de la capacité à simuler à l'horizon d'étude l'évolution de ces

- paramètres. Enfin, certains logiciels permettent de saisir directement les tables d'horaire des lignes ;
- la capacité de la ligne qui correspond en général à la capacité du véhicule. Cette information n'est toutefois pas toujours nécessaire si l'on ne prend pas en compte d'éventuel retard face à une capacité insuffisante de la ligne ou le temps nécessaire à l'arrêt pour assurer la montée des clients ;
 - la vitesse de la ligne est parfois spécifiée directement au niveau de la ligne et non des liens. Cet attribut n'est généralement défini que pour des lignes en site propre intégral pour lesquelles la vitesse est uniforme sur toute la ligne et peut être garantie indépendamment des conditions de circulation.

Généralement, les connecteurs de centroïde doivent être spécifiés pour chacun des modes de transport. Le temps d'accès au réseau n'est pas de même nature pour la voiture et pour les transports collectifs. L'accès aux transports collectifs nécessite en général un temps de marche qui est nettement supérieur au temps d'accès de la voiture. Le connecteur de centroïde peut dans certains cas être spécifique pour chacune des lignes ou groupes de lignes traversant une zone si le temps de marche pour accéder aux arrêts est différent pour chacune de ces lignes ou groupe de lignes. Le temps de marche étant généralement affecté d'un coefficient pondérateur au moins égal à deux pour traduire la pénibilité relative de la marche par rapport au temps en véhicule, il est important de définir assez précisément ce temps. Une mauvaise détermination des temps d'accès peut avoir des conséquences assez importantes sur les choix d'itinéraires tout particulièrement dans des réseaux fortement maillés. La détermination de cette distance de marche doit rester compatible avec la connaissance que l'on peut avoir des zones de chalandise des arrêts (de l'ordre de 300 mètres pour une ligne de bus et de 400 à 500 mètres pour du site propre). Ainsi, même si la zone est très étendue et que la distance moyenne de la population à l'arrêt est supérieure à 500 mètres, il est nécessaire de conserver une distance beaucoup plus courte pour le connecteur de centroïde. En général, on retient une distance assez proche de la distance moyenne effectivement parcourue par les utilisateurs des transports collectifs (et non pas une distance moyenne de la population aux arrêts). Toutefois, quel que soit le principe de codification retenu, le plus important est de maintenir une cohérence forte de ce principe pour l'ensemble de la codification du réseau. Cette cohérence devra également être maintenue lors de la phase de prévision qui suppose généralement une modification du réseau.

L'agrégation de plusieurs voiries routières dans un même lien (schéma 10.3c) s'effectue sans trop de difficulté surtout si les liens sont de même nature. En revanche, le problème des lignes de transports collectifs partageant un tronc commun est beaucoup plus complexe à traiter. Le traitement passe, en général, par la définition d'une ligne fictive sur l'ensemble de l'itinéraire commun à deux ou plus lignes de transports collectifs (schéma 10.5). La difficulté réside dans la définition des attributs de cette ligne fictive. Le choix d'une fréquence s'effectue souvent en sommant le nombre de passages de chacune des lignes ayant un tronc commun. Toutefois, la coordination des lignes n'existant pas forcément ou n'étant pas parfaite, cette règle conduit le plus souvent à minimiser le temps d'attente. De plus, certaines lignes peuvent être plus attractives que d'autres conduisant certains usagers à laisser passer un bus pour prendre le suivant jugé plus confortable. Cette construction des lignes fictives est généralement traitée par le logiciel.

La combinaison d'un service express avec un service omnibus est encore plus complexe puisque les lignes ne partagent en commun que certains arrêts. Dans ce cas, une représentation très précise du réseau peut s'avérer extrêmement complexe et lourde à mettre en œuvre. Comme toujours le niveau de précision dépendra des objectifs et des moyens disponibles pour l'étude. Mais dans tous les cas, il faut rester conscient de l'approximation et des conséquences que cela peut avoir, par exemple sur la répartition du trafic entre des lignes parallèles.

L'exemple proposé dans le schéma 10.5 permet de poser une partie des problèmes de la définition d'un réseau de transports collectifs. Tous les arrêts de la zone 4 ont été représentés par un

seul nœud 4001. Il ne sera donc pas possible de déterminer la charge de chacun des deux arrêts. Le problème est encore accru dans le cas de la zone 3, puisque les arrêts ne sont pas tous communs aux deux lignes. Un déplacement entre l'arrêt B8 et l'arrêt B2 ne pourra être distingué d'un déplacement entre l'arrêt B8 et l'arrêt A2. Pourtant, dans un cas, le déplacement est effectué en totalité sur la ligne B et dans l'autre, il y a une correspondance entre les lignes B et A. Compte tenu de la méthode d'affectation généralement utilisée pour les transports collectifs (méthode du plus court chemin), les deux déplacements seront affectés sur la ligne B. De même, les déplacements entre chacun des arrêts de la zone 3 et chacun des arrêts de la zone 4 seront affectés en totalité sur la ligne fictive (ils sont ensuite répartis entre les lignes composant la ligne fictive, généralement au prorata de leur fréquence de passage). Il ne sera donc pas possible de représenter la correspondance parfois nécessaire entre les lignes A et B, ni le fait que la ligne A sera privilégiée pour les montées ou descentes aux arrêts A2 et A3 et respectivement la ligne B aux arrêts B2 et B3. On remarquera également que chacune des zones n'a qu'un seul connecteur de centroïde, ce qui permet pas de distinguer le temps d'accès pour chacune des deux lignes au sein d'une même zone.

Le plus souvent ces approximations ne sont pas préjudiciables au regard des objectifs d'étude. Dans le cas contraire, la seule solution possible passe par une segmentation des zones 3 et 4 en autant de zones qu'il y a d'arrêts (alors que pour la voiture une définition plus précise du réseau est parfois possible sans pour autant redéfinir le zonage). Il est clair que cette méthode peut difficilement être généralisée à l'ensemble d'une agglomération, car elle conduirait à un nombre de zones extrêmement élevé et donc à un coût de production de données et d'analyse très lourd. L'objectif principal de cet ouvrage n'est pas d'approfondir la codification des réseaux, mais cet exemple illustre les difficultés d'une codification très fine des réseaux de transports collectifs, en général beaucoup plus complexe que pour la voiture particulière.

4. CONCEPTS DE BASE ET ETAPES PRELIMINAIRES

Plusieurs méthodes d'affectation sont disponibles comme nous l'illustrons plus loin (section 5, p. 328). Elles ont toutes en commun une première étape permettant de calculer le coût d'un ou plusieurs itinéraires entre chacune des origines i et chacune des destinations j (section 4.1). Pour ce faire, il est généralement nécessaire de définir une courbe débit-vitesse ou débit-coût, tout au moins pour l'affectation routière (section 4.2). Le coût ou temps d'un itinéraire servant à traduire le niveau de service de chacun des itinéraires, on privilégie généralement un temps ou coût généralisé (section 4.3).

Enfin, bien que ces étapes soient souvent un peu négligées, l'affectation ne peut généralement prendre place directement après la répartition modale. Les matrices origines-destinations par modes sont des matrices de déplacements et non des matrices de véhicules. Or, le modèle d'affectation traite de l'affectation de véhicules en ce qui concerne l'affectation de la demande automobile. Nous abordons rapidement l'étape de conversion véhiculaire qui permet de passer d'une demande de déplacements individuelles (en automobile) à une demande de véhicules automobiles dans la section 4.4. Les matrices de demande sont fréquemment des matrices journalières. Les niveaux de service étant très différents entre l'heure (ou la période) de pointe et le reste de la journée, on procède fréquemment à une segmentation de la matrice journalière. Elle permet d'isoler la matrice relative à l'heure (ou la période) de pointe (section 4.4). Ce passage à l'heure de pointe prend parfois place avant le choix du mode (sensible lui aussi au niveau de service respectif de la voiture et des transports collectifs).

4.1. Coût d'un itinéraire et notations

Le coût d'un déplacement pour aller d'une origine i à une destination j , dépend du choix de l'itinéraire. Pour un itinéraire donné, ce coût est constitué du coût nécessaire pour parcourir chacun des liens de l'itinéraire permettant d'aller du centroïde origine i au centroïde destination j . Le coût de chacun des liens peut être déterminé à partir des attributs du lien et d'une loi reliant le coût à ces

attributs. Cette loi peut être commune à l'ensemble des liens ou être définie par classe de liens (généralement en fonction de leur niveau de service : autoroute, voie rapide...). Deux éléments dans ce processus de calcul du coût d'un itinéraire méritent d'être précisés. Il s'agit de la courbe débit-vitesse ou débit-coût (section 4.2) et de la fonction de coût généralisé (section 4.3).

Avant de les aborder, il est nécessaire de préciser quelques notations que nous utilisons tout au long de chapitre :

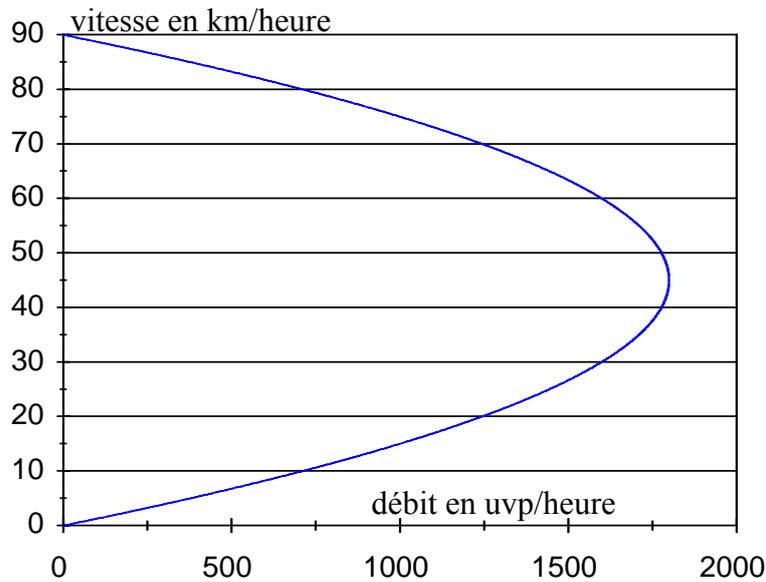
- T_{hdr} correspond au trafic entre une origine h et une destination d via une route (ou un itinéraire ou un chemin) r . Ce trafic correspond à un nombre de véhicules pour une affectation automobile ou un nombre de déplacements d'individus pour une affectation en transports collectifs ;
- V_a est le volume de trafic sur le lien a ;
- $c(V_a)$ est le coût du lien a pour un certain niveau de trafic V_a sur le lien a (on le notera parfois également $c_a(V_a)$) ;
- $C_a(V)$ est la fonction reliant le coût au volume pour le lien a , c'est-à-dire la courbe débit-coût ;
- c_{hdr} est le coût de l'itinéraire r (ou route r) pour aller de h à d ;
- δ_{hdr}^a est une variable qui prend la valeur 1 si le lien a est choisi pour aller de h à d par l'itinéraire r et 0 sinon.

Le coût total sur un lien est obtenu par $V_a * c(V_a)$ et $\sum_a V_a * c(V_a)$ représente le coût total des déplacements sur le réseau.

Le coût d'un lien ne dépend pas toujours du volume de trafic sur ce lien. C'est bien souvent le cas des transports collectifs. Dans ce cas, on pourra trouver la notation $c(a)$ pour exprimer cette absence de relation. Lorsque le trafic V_a est nul, le coût du lien est appelé coût à vide par opposition au coût $c(V_a)$ appelé en charge. Le coût à vide figure parmi les attributs du lien, ainsi que la fonction permettant de calculer le coût en fonction du volume de trafic.

Dans certains modèles, l'estimation du coût d'un lien ne dépend pas que du volume sur le lien a . Dans ce cas, on utilise la notation suivante $C_a = C_a(\{V\})$ pour exprimer la relation entre le coût du lien a et l'ensemble des volumes de trafic. Le plus souvent cette relation est limitée aux liens connectés au lien a . Il est évident qu'une formulation très générale conduit à des méthodes d'affectation et à des algorithmes beaucoup plus complexes à mettre en œuvre. Cette approche est toutefois plus réaliste dans la mesure où elle permet de rendre compte de manière explicite des phénomènes de remontée de queues de bouchons ou de saturation de carrefours dans des réseaux fortement congestionnés.

4.2. Courbe débit-vitesse ou débit-coût



Graphique 10.1 : Exemple de courbe débit-vitesse de forme parabolique (équation 10.1, pour une capacité de 1800 uvp/heure et une vitesse à vide de 90km/heure)

La vitesse d'un véhicule sur un lien dépend du trafic sur ce lien ou plus exactement du rapport entre le trafic sur ce lien et sa capacité. A mesure que ce rapport s'accroît, la vitesse diminue, faiblement dans un premier temps puis de plus en plus rapidement à mesure que l'on s'approche de la capacité maximale de la voie. On arrive ensuite dans une nouvelle phase où le débit et la vitesse diminuent conjointement. Cette forme de courbe est illustrée dans le graphique 10.1. Elle est appelée diagramme fondamental en ingénierie du trafic, et est classiquement présentée sous une forme parabolique. Une forme relativement simple est proposée par White (1963, cité par Thomas, 1991). Elle a pour équation :

$$s^2 = as - bV \quad (10.1)$$

avec s : vitesse du lien en kilomètre/heure ;

V : le volume sur le lien en UVP/heure ;

a : s_f , qui est la vitesse à vide sur le lien (*free flow speed*) = 90 km/heure sur le graphique 10.1 ;

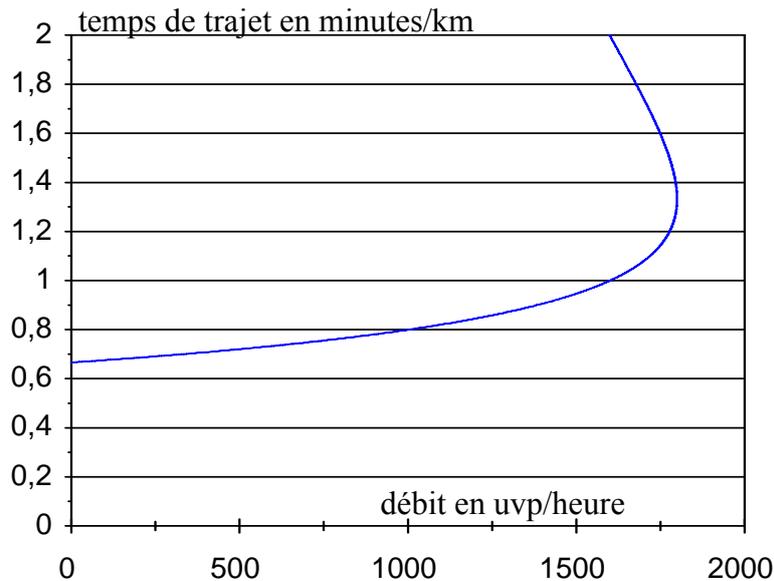
b : $= \frac{s_f^2}{4V_c}$, avec V_c la capacité du lien = 1800 uvp/heure sur le graphique 10.1.

Les données empiriques s'écartent toutefois un peu de cette forme de courbe. D'autres équations ont ainsi été proposées (Cohen, 1990). De plus, l'observation empirique montre une dispersion assez forte des débits pour une vitesse donnée surtout pour la zone de congestion correspondant aux vitesses inférieures à la vitesse obtenue pour le débit maximum. Ce type de courbe est employé pour le dimensionnement d'infrastructures ou en ingénierie du trafic (FHWA, 1994 ; Buisson, 1996 ; Cohen, 1991). On utilise une forme un peu différente pour les modèles d'affectation routière. La relation qui nous intéresse relie en fait le débit au temps de parcours ou au coût de transport. Cette relation conduit au graphique 10.2 (en retenant l'équation 10.1 avec les mêmes valeurs numériques).

Cette forme de courbe ne peut convenir pour une utilisation dans un modèle d'affectation, car elle ne permet pas d'estimer un temps de parcours pour des volumes supérieurs à la capacité. De plus, si la confrontation empirique permet de valider la courbe pour des vitesses supérieures à la vitesse correspondant à la capacité maximale (45 km/heure sur le graphique 10.1), la dispersion est très forte

pour des valeurs plus faibles (Cohen, 1991). Enfin, pour un débit donné, deux valeurs de temps de parcours sont possibles. On leur préfère donc d'autres formes de courbes, dont les propriétés recherchées sont les suivantes (Ortuzar, Willumsen, 1994) :

- réalisme, le temps de trajet modélisé doit être réaliste ;
- la fonction doit être non décroissante, monotone ;
- la fonction doit être continue et différentiable ;
- la fonction doit être définie pour des valeurs de trafic supérieures à la capacité (de telle sorte que le temps ne soit pas infini pour un trafic supérieur à la capacité) ;
- la fonction doit être facilement transférable d'un contexte à un autre. Il faut donc éviter des paramètres spécifiques à un terrain donné.



Graphique 10.2 : Exemple de courbe débit-temps (équation 10.1, pour une capacité de 1800 uvp/heure)

Nous donnons quelques exemples de courbes fréquemment utilisées.

La courbe proposée par le *Bureau of Public Roads* (1964) aux Etats-Unis :

$$t = t_0 \left(1 + \alpha \left(\frac{V}{V_c} \right)^\beta \right) \quad (10.2)$$

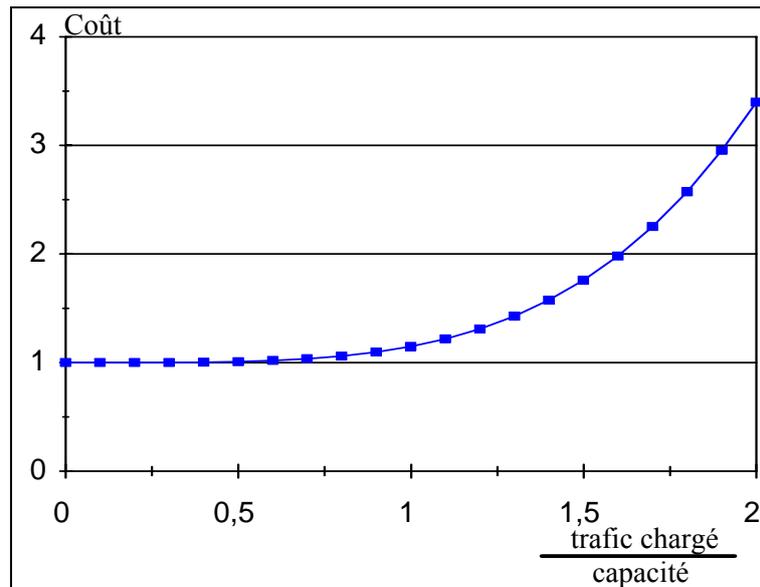
où t est le temps de parcours sur un lien ;

t_0 est le temps de parcours à vide ;

V le volume sur le lien ;

V_c la capacité maximale du lien ;

α et β sont deux paramètres de calage. Les valeurs par défaut proposées par Horowitz (1993) sont $\alpha = 0,15$ et $\beta = 4$ (graphique 10.3 pour un coût à vide = 1).



Graphique 10.3 : Courbe débit-coût du *Bureau of Public Roads* des Etats-Unis (1964, pour un coût à vide = 1)

La formulation la plus courante, en France, est celle utilisée par le logiciel DAVISUM, issue du logiciel DAVIS :

$$C = C_0 * \frac{1,1 - A * x_i}{1,1 - x_i} \text{ pour } x_i = \frac{\text{trafic chargé sur le lien}}{\text{capacité du lien}} \leq 0,9 \text{ (non saturé)}$$

$$\text{tangente de la courbe pour } x_i > 0,9 \text{ (arc saturé)} \quad (10.3)$$

avec : C : coût du lien ;

C_0 : temps de parcours à vide ;

A : coefficient de la courbe débit/vitesse selon la catégorie de l'arc. Les valeurs suivantes sont souvent proposées par défaut :

- 0 pour la voirie tertiaire ou voies locales ;
- 0,3 pour la voirie secondaire ou voies de desserte ;
- 0,5 pour la voirie primaire ou grandes artères urbaines ;
- 0,666 pour les voies rapides ;
- 0,92 pour les autoroutes.

Les courbes correspondantes sont présentées dans le graphique 10.4 (avec $C_0 = 1$) pour différentes valeurs de A.

Enfin, le logiciel CartOPERA (CETE Méditerranée, CERTU, 2000a,b) propose une courbe proche de celle du logiciel DAVIS :

$$C = (C_0 * ((1,1 - A * X) / (1,1 - X))) + K \text{ si } X < 1.$$

$$C = (C_0 * (((1,1 - A) / 0,1) * X^2)) + K \text{ si } X \geq 1. \quad (10.4)$$

avec : C : coût du lien ;

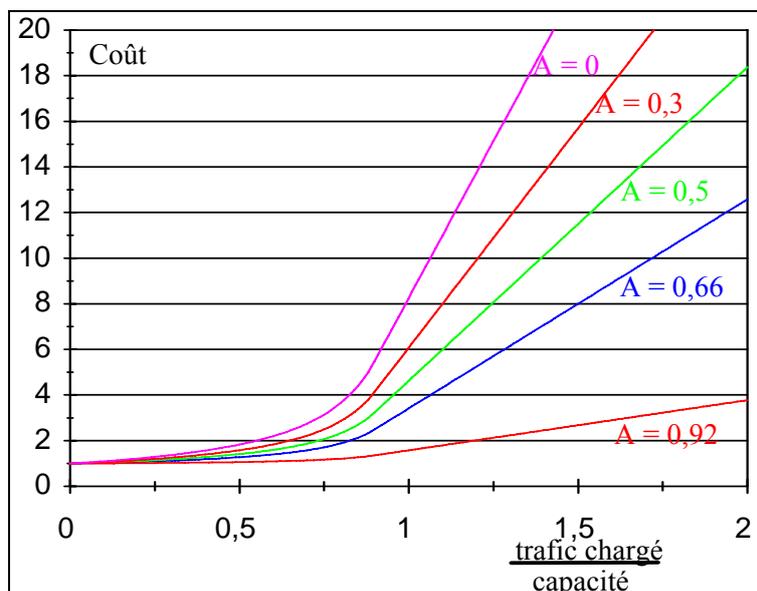
K : (distance * coût kilométrique) + coût supplémentaire (péages...) ;

C_0 : temps de parcours à vide [* valeur du temps] ;

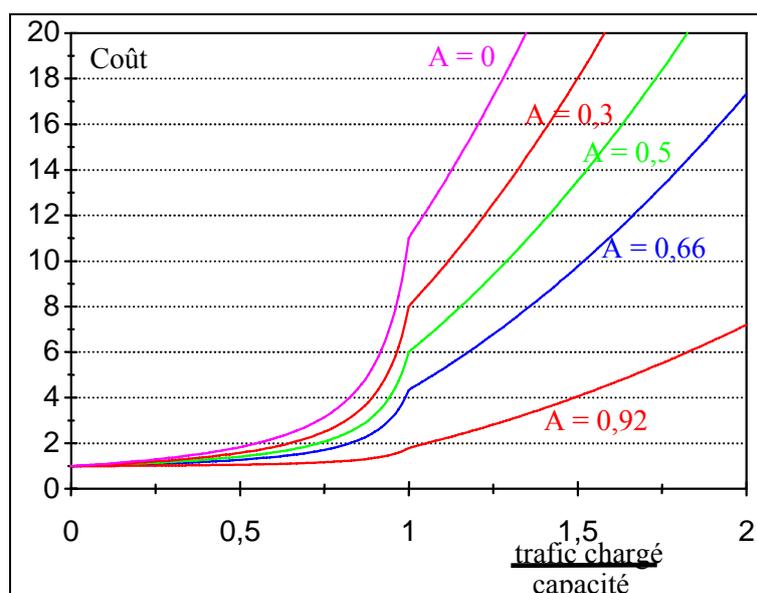
X : charge de trafic, soit volume (trafic chargé sur le lien) / capacité ;

A : coefficient de la courbe débit/vitesse selon la catégorie de l'arc. Le logiciel autorise 10 types de courbes différenciées par la valeur donnée au paramètre A avec les valeurs par défaut proposées ci-dessus.

Les courbes correspondantes sont présentées dans le graphique 10.5 (avec $K = 0$ et $C_0 = 1$).



Graphique 10.4 : Courbe débit-coût du logiciel DAVISUM (pour un coût à vide = 1)



Graphique 10.5 : Courbe débit-coût du logiciel CartOPERA (pour un coût à vide = 1)
(source : CETE Méditerranée, CERTU, 2000a,b)

L'avantage de l'ensemble de ces courbes par rapport à celle du graphique 10.2 est de fournir des valeurs de coûts ou de temps sur le lien pour des valeurs de trafic supérieures à la capacité du lien. Ortuzar et Willumsen (1994) proposent d'autres courbes notamment celles produites par le Ministère des transports britanniques (voir également Lindsey, Verhoef, 2000).

Lorsque l'on dispose de suffisamment de données sur les temps de parcours en fonction de la charge des infrastructures, il est possible de caler sa propre courbe débit-coût sur les données observées. Toutefois, cette information est rarement disponible ou d'une qualité insuffisante. Les

principales sources de données sont issues des boucles de comptage qui peuvent fournir vitesse et débit. Leur utilisation est toutefois limitée aux sections autoroutières ou aux voies rapides pour lesquelles la vitesse peut être relativement uniforme sur des sections suffisamment longues. En revanche, cette information est difficilement utilisable (lorsqu'elle est disponible) sur des sites réellement urbains. Des mesures de temps de parcours sont parfois réalisées avec des véhicules circulant sur certaines origines-destinations en conditions réelles de circulation. Mais ce type de mesure ne peut être généralisé en raison de son coût très élevé. La mise en œuvre est également assez délicate et nécessite des protocoles extrêmement rigoureux, notamment sur le style de conduite et sur le plan d'échantillonnage.

De ce fait, les courbes par défaut (comme celles décrites ci-dessus) sont fréquemment utilisées. Même si les données manquent pour estimer des courbes spécifiques sur le terrain d'étude, il reste souhaitable de valider le choix des courbes et des coefficients de calage. Cette validation s'opère avec le type de données mentionné au paragraphe précédent. Cette information étant fréquemment manquante, les données de temps de parcours déclarés dans les enquêtes peuvent également servir de base. Il faut toutefois être vigilant quant à la fiabilité de ces données qui correspondent à des temps déclarés donc « déformés » par la perception des individus enquêtés. Enfin, des données issues des questionnaires de flottes de véhicules (taxi, livraison...) sont très utiles lorsqu'il est possible de se les procurer.

4.3. Coût généralisé

Nous souhaitons revenir rapidement sur la fonction de coût généralisé, bien que nous ayons déjà abordé la question (notamment chapitre 5, p. 152). La formule générique du coût généralisé est donnée par l'équation 5.8 (p. 152) :

$$C_g = P + \left(\sum_i \alpha_i T_i \right) * V_t$$

où C_g est le coût généralisé ;

P est le prix du déplacement ;

T_i est une partie élémentaire du temps nécessaire pour effectuer le déplacement ;

V_t est la valeur que l'individu attribue à son temps ;

α_i est un coefficient pondérateur du temps élémentaire T_i .

Il paraît utile de rappeler que l'estimation des différentes composantes de ce coût généralisé pose des problèmes spécifiques.

Le prix est généralement défini de manière moyenne, en fonction de la distance pour la voiture (plus les éventuels péages) et selon le mode de tarification pour les transports collectifs. Toutefois, ce prix moyen recouvre des situations qui peuvent être fortement contrastées. Pour les transports collectifs, on utilise généralement la recette moyenne fournie par l'exploitant, mais entre le titre unitaire et les différentes formules tarifaires d'abonnement et les réductions tarifaires à caractère commercial ou social, la disparité peut être très élevée sans que l'information soit généralement disponible. De même, pour la voiture, le coût d'usage dépend du véhicule utilisé, de son ancienneté, du style de conduite... Il faut enfin prendre en compte le stationnement bien que l'information sur le coût du stationnement soit rarement disponible dans les enquêtes déplacements. Dans un modèle désagrégé, il est rarement possible de prendre en compte l'ensemble de ces éléments de manière désagrégée pour déterminer le prix. De plus, si dans certains cas certains éléments sont pris en compte, il faut encore être capable de simuler leur évolution à l'horizon d'étude.

Le temps de parcours est généralement décomposé en temps élémentaires. Pour la voiture, on peut considérer :

- le temps en véhicule obtenu à partir des courbes débit-coût ;
- le temps d'accès à l'origine et à la destination (parfois en isolant le temps de recherche d'une place de stationnement). Ce dernier temps est généralement affecté d'un coefficient pondérateur proche de deux pour traduire sa perception négative par les individus ;
- un bonus pour les autoroutes ou voies rapides qui permet de prendre en compte un facteur de confort, d'attractivité de ce type de voie, en plus des facteurs liés à la vitesse.

Pour les transports collectifs, on distingue fréquemment :

- le temps en véhicule ;
- le temps de marche pour accéder à l'arrêt à l'origine et à la destination ;
- le temps d'attente du véhicule. Il est généralement souhaitable de tenir compte de l'irrégularité des services. En supposant une arrivée aléatoire, la SEMALY dans son logiciel TERESE propose, pour des intervalles suffisamment faibles, un calcul du temps d'attente = αI , où I est l'intervalle moyen entre deux véhicules et α un coefficient de régularité qui varie entre 0,5 (régularité parfaite, dans la pratique cette valeur n'est jamais choisie) et 0,8 (pour des services peu réguliers) (SEMALY, 2000). Ortuzar et Willumsen (1994) proposent la formule $\frac{(I^2 + \sigma^2)}{2I}$, où I est l'intervalle moyen entre deux services et σ son écart-type (un écart-type nul conduit à $I/2$; le choix $\alpha = 0,5 + \frac{\sigma^2}{2I^2}$ conduit à une identité de la formulation du logiciel TERESE avec cette formulation). Toutefois, pour des intervalles entre véhicules plus élevés, l'arrivée aléatoire n'est plus pertinente. Dans ce cas, la plupart des usagers connaissent les horaires et le temps d'attente est généralement borné à un maximum de 5 à 10 minutes (ce qui est déjà très pénalisant compte tenu du coefficient pondérateur) ;
- le temps de correspondance défini indépendamment de l'attente du véhicule suivant, qui peut être forfaitaire (en général plusieurs minutes pour traduire la pénibilité de la correspondance indépendamment du temps réel supplémentaire) ou ajusté sur les lignes en correspondance en fonction du temps réellement nécessaire pour la correspondance + une constante qui traduit la pénibilité de la correspondance indépendamment du temps réel supplémentaire (cette constante peut valoir de 2 à 10 minutes selon les réseaux).

Les temps de marche et d'attente sont affectés de coefficients le plus souvent supérieurs à deux (en principe compris entre 1,5 et 3, leur estimation résulte généralement de calage sur des données de préférences révélées ou parfois de préférences déclarées). La localisation précise de l'origine et de la destination du déplacement par rapport aux différents arrêts de transports collectifs n'étant pas connue, il est généralement nécessaire de retenir une valeur moyenne. Cette valeur fait partie des attributs du connecteur de centroïde.

La valeur du temps peut être moyenne en l'absence de données. Toutefois, une valeur individuelle relative au déplacement est préférable pour tenir compte de la forte dispersion des valeurs individuelles du temps. Il est en effet important en modélisation de privilégier une valeur perçue par rapport à une valeur tutélaire ou moyenne pour traduire correctement la perception de l'utilité de chacun des itinéraires pour une origine-destination donnée. L'information individuelle n'étant pas toujours disponible, on utilise parfois une loi de distribution des valeurs du temps dans la population des déplacements. La forme de ces distributions est généralement log-normale (chapitre 2, p. 54). Le plus souvent, la formalisation du modèle ne permet pas de retenir une distribution des valeurs du temps. Dans ce cas, il est souvent préférable d'utiliser une valeur du temps médiane plutôt qu'une valeur moyenne plus sensible aux très hauts revenus.

Compte tenu de ces différents éléments, il est rarement (pour ne pas dire jamais) possible d'estimer un coût généralisé au niveau du déplacement. Cette remarque s'applique tout autant au

modèle agrégé qu'au modèle désagrégé. Cet aspect est important lorsque l'on doit considérer le choix entre une modélisation agrégée ou désagrégée.

Au-delà des éléments pris en compte dans le coût généralisé, il est important de garder en mémoire que beaucoup d'autres composantes du déplacement ou de l'individu influent sur la perception des caractéristiques du déplacement. La perception du confort, l'attrait esthétique ou fonctionnel de l'itinéraire sont par exemple des dimensions évoquées par des individus dans leur choix d'itinéraire. L'habitude et la connaissance préalable des itinéraires contribuent souvent à une perception moindre des coûts sur ces itinéraires relativement à des itinéraires inconnus. L'absence d'information relativement à certains itinéraires conduit de fait à leur attribuer un coût quasi-infini.... Tous ces éléments contribuent à une perception subjective du coût des itinéraires par les individus.

Enfin, il est important de souligner les limites de cette formalisation mathématique de la fonction de coût généralisé. L'additivité de l'ensemble des éléments entrant dans la fonction de coût n'est pas évidente, car il n'est pas sur que les individus disposent ainsi d'un référent unique permettant d'assurer la conversion de l'ensemble des composantes du coût généralisé.

4.4. Conversion véhiculaire et passage à la pointe

Les matrices origines-destinations produites à l'issue de l'étape de choix modal sont généralement des matrices de déplacements de personnes. Pour l'affectation routière, il est nécessaire de transformer les matrices pour avoir des déplacements de véhicules (section 4.4.1). Les conditions d'offre de transport étant souvent fortement différenciées selon les périodes de la journée, on procède généralement à une segmentation des matrices jour en différentes périodes temporelles (section 4.4.2).

4.4.1. Conversion véhiculaire

L'étape de choix du mode produit des matrices origines-destinations de déplacements individuels. Pourtant, le modèle d'affectation routière requiert en entrée des matrices de déplacements de véhicules, notamment pour l'estimation des temps de parcours à partir de la courbe débit-coût. A moins d'identifier explicitement les modes voiture conducteur et voiture passager parmi les options du choix modal, il est donc nécessaire de convertir les matrices de déplacements d'individus en matrices de déplacements de véhicules. C'est cette conversion qui est appelée « conversion véhiculaire ».

La conversion s'effectue par le biais du taux d'occupation des véhicules. Ce taux peut être défini pour l'ensemble de l'agglomération, mais il est préférable de le définir par segment. La segmentation s'appuie habituellement sur le type de liaison (en termes d'origines-destinations), le motif de déplacement et certaines variables socio-économiques. La précision dépend bien évidemment des données disponibles pour estimer la valeur du taux d'occupation pour chacun des segments ainsi définis.

En prévision, il est nécessaire d'estimer la valeur de ce taux. En l'absence de modèle spécifique pour l'estimation de ce taux, généralement, on prolonge les tendances antérieures lorsque l'on peut disposer de données relatives à différentes périodes temporelles, sinon, on conserve une valeur constante. Ces choix sont toutefois problématiques lorsque des mesures de covoiturage ou d'accroissement des taux d'occupation sont incluses dans la politique de déplacements devant être simulée. Dans ce cas, il est recommandé, soit de construire un modèle permettant de simuler ce taux, soit lors de l'étape de choix modal de définir explicitement des modes de transport prenant en compte l'occupation du véhicule (en séparant les modes voiture conducteur et voiture passager, en identifiant explicitement le covoiturage...).

4.4.2. Passage à la période de pointe

Les étapes de génération et de distribution portent généralement sur une période de 24 heures correspondant le plus souvent à une journée de semaine. L'offre de transport est pourtant fortement différenciée selon la période de la journée. Pour la voiture, la performance est accrue dans les périodes creuses par rapport à celle de la pointe du fait d'une moindre charge des infrastructures routières. En revanche, pour les transports collectifs, l'offre est généralement plus abondante (bien que parfois de qualité inférieure pour les réseaux très chargés) en période de pointe. De ce fait, on distingue généralement l'affectation relative à la période de pointe du matin et du soir de celle relative à la ou les périodes creuses. La distinction de ces périodes est également assez fréquente pour l'étape de choix du mode de transport.

Les méthodes employées sont identiques selon que le calcul est effectué avant ou après l'étape de choix modal. La méthode la plus simple est similaire à celle que nous venons de voir pour la conversion véhiculaire. Elle consiste à estimer le pourcentage des déplacements quotidiens prenant place au cours de la période de pointe. Ce pourcentage n'est pas identique pour l'ensemble des déplacements. Généralement, on retrouve les variables de segmentation évoquées pour la conversion véhiculaire : le type de liaison (en termes d'origines-destinations), le motif de déplacement et certaines variables socio-économiques. On peut difficilement faire l'économie du motif de déplacement, les rythmes quotidiens étant fortement différenciés selon le motif. De plus, si la pointe du soir est distinguée de la pointe du matin, le type de liaison, notamment pour les échanges centre-périphérie, doit être impérativement pris en compte.

De nouveau en prévision, il est nécessaire d'estimer la valeur du pourcentage. En l'absence de modèle spécifique, on prolonge les tendances antérieures lorsque l'on peut disposer de données relatives à différentes périodes temporelles, sinon, on conserve une valeur constante. Ce dernier choix est toutefois problématique dans la mesure où l'on observe un étalement des périodes de pointe dans la plupart des agglomérations sous le double effet de la congestion et de l'évolution des rythmes quotidiens (aménagement du temps de travail, 35 heures...).

Le pouvoir prédictif de ces approches étant fortement limité, de nombreux travaux se sont développés ces dernières années pour modéliser le choix de l'heure de départ, notamment pour les déplacements domicile travail. On trouve une bonne synthèse de ces travaux dans (Mahmassani, 2000). Nous pouvons également citer les travaux de De Palma et Fontan (2001) sur le choix de l'heure d'arrivée appliqués à l'Ile de France qui s'appuient sur une production de données spécifiques (De Palma, Fontan, 2000).

5. PRINCIPALES FAMILLES DE MODELES D'AFFECTION

Les fondements théoriques des modèles d'affectation sont similaires à ceux des étapes précédentes : l'individu choisit l'alternative qui maximise son utilité. Compte tenu de la nature du choix relatif à cette étape, choix de l'itinéraire compte tenu d'une origine-destination et d'un mode donné, l'utilité (ou plus exactement la désutilité) est mesurée à travers le coût généralisé. Le principe de base de tous les modèles d'affectation est donc de considérer que l'individu cherche à minimiser le coût généralisé de son déplacement pour une origine-destination et un mode donné. Les principes généraux de calcul du coût généralisé ont été exposés dans les sections précédentes.

5.1. Classification des modèles d'affectation

Si les modèles ont un principe de base commun, ils sont généralement classés selon deux dimensions (Thomas, 1991 ; Ortuzar, Willumsen, 1994) :

- prise en compte des contraintes de capacité. La capacité (généralement en UVP) est un des attributs des liens. Certains modèles ne prennent pas en compte la contrainte de capacité du

lien. Dans ce cas, le coût du lien ne dépend pas de la charge du lien. Le risque est alors important d'avoir un chargement qui dépasse largement la capacité du lien. Une application de la courbe débit-coût montrerait alors que le coût du lien est nettement plus élevé que celui qui a été pris en compte dans le processus de choix de l'itinéraire, ce qui remettrait en cause les choix des individus. Ce type de modèle n'est donc applicable que sur des réseaux non ou très faiblement congestionnés (ce qui est rarement le cas en urbain) ou pour des applications en transports collectifs où l'incidence de la charge sur les temps de parcours est rarement prise en compte ;

- utilité déterministe ou probabiliste. Dans le premier cas, on considère que tous les individus perçoivent d'une manière identique le coût d'un lien ou d'un itinéraire. L'expérience démontre facilement le contraire pour deux raisons principalement. Tout d'abord, la perception des différents éléments constituant le coût généralisé peut être différente selon les individus. Pour tel individu, la marche sera considérée comme très pénalisante par exemple du fait d'une faible motricité, alors que pour un autre elle pourra être considérée comme faiblement pénalisante car l'individu aime marcher ou que la marche s'effectue dans une zone commerçante ou un parc... De plus, les individus ne disposent le plus souvent que d'une information partielle sur l'ensemble des itinéraires potentiels. La seconde concerne l'ensemble des facteurs du choix non pris en compte dans la fonction de coût généralisé : l'agrément du site traversé, le climat... Par ailleurs, l'estimation des grandeurs entrant dans le coût généralisé est elle-même entachée d'incertitudes comme nous l'avons illustré section 4.3 (p. 325).

La combinaison de ces deux dimensions conduit à la classification présentée dans le tableau 10.2.

		Procédure de choix de l'itinéraire	
		Déterministe	Probabiliste
Contrainte de capacité	Non	Méthode tout-ou-rien ou plus court chemin	Procédure stochastique pure (multichemin Dial, Burrell...), prix-temps
	Oui	Méthode d'équilibre (équilibre de Wardrop, affectation par tranche)	Méthode d'équilibre avec utilité stochastique

Tableau 10.2 : Classification des modèles d'affectation

Nous développons dans les sections qui suivent chacune de ces méthodes :

- affectation tout-ou-rien ou plus court chemin : (section 5.2) ;
- affectation stochastique pure ou multichemin : (section 5.3) ;
- affectation à l'équilibre : (section 5.4) ;
- affectation à l'équilibre avec utilité stochastique : (section 5.5).

Nous n'aborderons pas ici l'affectation dynamique, bien qu'elle bénéficie d'un grand nombre de travaux de recherche. Son utilisation n'est pas très fréquente pour la planification à moyen terme qui est le terme temporel principalement couvert par les modèles à quatre étapes. Nous renvoyons le lecteur vers la littérature correspondante (Friesz, Bernstein, 2000 ; Henn, 2001 ; De Palma et al., 1997).

5.2. Affectation tout-ou-rien ou plus court chemin

Il s'agit de la méthode la plus simple. Le principe de la méthode est contenu dans le nom qui lui est donné. Il s'agit d'une procédure qui affecte la totalité du trafic sur une origine-destination donnée

sur le chemin qui est le plus court (en termes de coût généralisé) et rien sur tous les autres chemins possibles.

Deux hypothèses sont nécessaires pour cela :

- tous les individus perçoivent le coût d'un lien de manière identique ;
- le coût d'un lien est fixe et donc ne dépend pas de sa charge. Le coût d'un lien, dans cette méthode est toujours calculé à partir du coût à vide du lien. Si l'hypothèse est acceptable pour les transports collectifs, pour la voiture son utilisation doit être limitée à des réseaux sans congestion.

La première hypothèse conduit à un plus court chemin unique. La seconde conduit à l'affectation de la totalité de la demande sur ce plus court chemin puisque la contrainte de capacité n'est pas prise en compte et que le temps de parcours est indépendant de la charge sur les différents liens de l'itinéraire.

La réalisation de l'affectation au plus court chemin nécessite la construction de l'arbre de plus court chemin à partir de chacun des centroïdes. Il s'agit de l'arbre définissant le chemin le plus court à partir d'une origine donnée vers l'ensemble des autres destinations (schéma 10.6).

Il existe plusieurs algorithmes de plus courts chemins (Thomas, 1991). Les plus connus sont ceux de Dijkstra (1959) et de Moore (1957). On trouve parfois également celui de d'Esopo (Van Vliet, 1978 cité par Thomas, 1991). Ils se distinguent par leur rapidité de convergence qui tient principalement à l'ordre de sélection des nœuds à considérer pour la détermination du nœud suivant (lors de l'étape 2 de l'algorithme de plus court chemin). Bien que ceux de Dijkstra et de d'Esopo aient les meilleures performances (Ortuzar, Willumsen, 1994), c'est le premier qui est le plus répandu.

Exemple d'algorithme de plus court chemin

Le principe de détermination de l'arbre de plus court chemin est décrit ci-dessous.

Notation :

c_i^* est le coût du plus court chemin de l'origine à i ;

c_{ij} est le coût du lien (i,j) ;

R est la table comprenant l'ensemble des nœuds marqués et \bar{R} la table contenant l'ensemble des nœuds non marqués. Quand un nœud i est marqué, il est entré dans R avec la valeur (p_i, c_i^*) ou p_i est le nœud précédent dans l'arbre de plus court chemin entre l'origine et i ;

h le nœud origine pour lequel on détermine l'arbre de plus court chemin.

Etape 1 : initialisation

Mettre tous les nœuds (sauf l'origine) dans la table \bar{R} , mettre l'origine h dans la table R avec la valeur $(0, 0)$.

Etape 2 : détermination du nœud suivant

Choisir le nœud j suivant appartenant à \bar{R} tel que :

$$c_j^* = \min_{x \in R} (c_x^* + \min_{y \in \bar{R}} (c_{xy})) = c_i^* + c_{ij}, \text{ pour tout } i \in R \text{ et pour tout } j \in \bar{R}. \quad (10.5)$$

$$p_j = i.$$

Etape 3 : mettre ce nœud dans la table R

Mettre le nœud j dans la table R avec la valeur (p_j, c_j^*) , enlever j de la table \bar{R} .

Etape 4 : test de fin de l'algorithme

Si \bar{R} est non vide aller à l'étape 2, sinon fin de l'algorithme.

Fin de l'algorithme

Cet algorithme permet de déterminer l'arbre de plus court chemin (schéma 10.6), ainsi que le coût généralisé correspondant à ce plus court chemin entre chacune des origines-destinations. Lors de l'utilisation de l'algorithme, les connecteurs de centroïde ne doivent pas être inclus dans les plus courts chemins (sauf à l'origine et la destination de l'itinéraire). De ce fait, les centroïdes connectés au réseau par un connecteur de centroïde ne sont pas inclus dans la table \bar{R} lors de la première étape d'initialisation. A l'issue de l'algorithme, il est, en revanche, nécessaire de les relier à l'arbre ainsi construit en choisissant pour chaque centroïde le connecteur offrant le chemin le plus court.

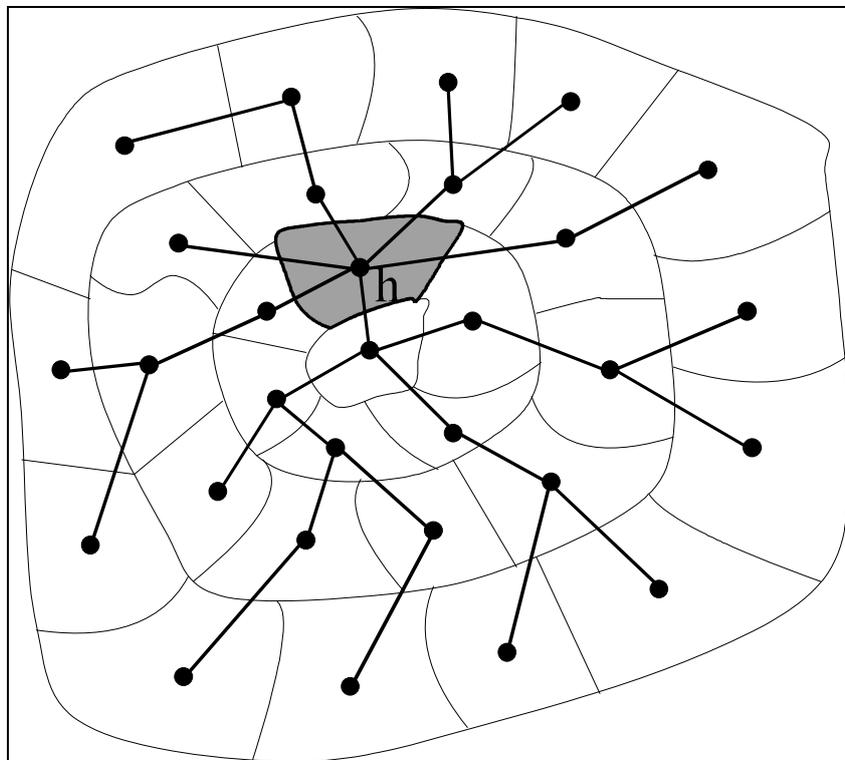


Schéma 10.6 : Arbre de plus court chemin depuis l'origine h

La charge est ensuite affectée sur le plus court chemin pour chacune des origines-destinations. La principale limite de cette méthode est de conduire à des charges pouvant excéder fortement la capacité. De plus, la charge est fortement dépendante de la définition du réseau. Si deux itinéraires ont des coûts très proches sur une origine-destination, la modification du coût sur un lien peut faire basculer le plus court chemin sur l'autre itinéraire et donc toute la charge. Sur les réseaux importants et fortement maillés, la charge sur les différents liens est donc très dépendante de la définition du réseau. Sachant que la définition du coût des différents liens est une opération délicate et fortement chronophage, l'incertitude est souvent assez forte quant aux paramètres des liens. Un lien peut facilement basculer d'une situation fortement congestionnée à une situation de trafic faible voire nul (ou inversement) suite à une modification mineure du réseau. De ce fait, cette méthode est utilisée principalement pour l'affectation en transports collectifs. Il est de plus conseillé de limiter son

utilisation à des réseaux dont le maillage n'est pas trop important. Son utilisation pour la voiture est fortement déconseillée d'autant que le temps de calcul ordinateur est de moins en moins un frein à l'utilisation d'autres méthodes.

Le principal intérêt de cette méthode est d'indiquer l'itinéraire désiré dans le cas d'un réseau non congestionné. Par ailleurs, elle sert de base à l'implantation des autres méthodes d'affectation.

5.3. Affectation stochastique ou multichemin

Nous présentons tout d'abord les fondements théoriques de l'affectation stochastique (section 5.3.1). Nous développons ensuite les méthodes les plus couramment utilisées : l'affectation multichemin de Dial (section 5.3.2) et de Burrell (section 5.3.3) et la méthode d'Abraham appelée également méthode Ariane (section 5.3.4) utilisée en France pour l'interurbain. Nous présentons enfin l'affectation prix-temps bien que l'approche ne soit pas stochastique. En revanche, elle est effectivement multichemin (section 5.3.5). Nous concluons avec une présentation des principales limites de l'affectation multichemin (section 5.3.6).

5.3.1. Fondements théoriques de l'affectation stochastique

L'affectation stochastique permet d'abandonner la première hypothèse du modèle d'affectation selon le plus court chemin. Elle considère que tous les individus ne perçoivent pas le coût d'un chemin de manière identique. Comme nous l'avons déjà mentionné (section 5.1, p. 328) plusieurs dimensions justifient cette approche :

- la perception des différents éléments constituant le coût généralisé d'un lien peut être différente selon les individus ;
- certains facteurs explicatifs du choix d'un itinéraire ne sont pas pris en compte dans la fonction de coût généralisé du lien : l'agrément du site traversé, le climat... ;
- l'estimation des grandeurs entrant dans le coût généralisé est entachée d'incertitude (section 4.3, p. 325).

En revanche, la seconde hypothèse émise pour l'affectation selon le plus court chemin subsiste :

- le coût d'un lien est fixe et donc ne dépend pas de sa charge. Le coût d'un lien est toujours calculé à partir du coût à vide du lien.

Pour tenir compte des différences de perception, on définit une loi de distribution autour d'une valeur moyenne qui est celle calculée par le modèle à partir de la description du réseau. On peut également interpréter cette distribution comme la distribution du terme d'erreur dans une formulation aléatoire de l'utilité.

L'affectation tout-ou-rien correspond ainsi à une formulation déterministe de l'utilité associé à un itinéraire, tandis que l'affectation multichemin retient une formulation aléatoire de l'utilité (chapitre 5, p. 157). Les fondements théoriques de ces approches s'appuient donc sur la théorie des choix discrets avec une approche de l'utilité aléatoire pour l'affectation multichemin et déterministe pour l'affectation tout-ou-rien. L'affectation tout-ou-rien s'appuie toujours sur une définition agrégée de l'utilité comme nous l'avons présentée dans la section précédente (section 5.2, p. 329) (il arrive toutefois que la population des déplacements soit segmentée, principalement sur les caractéristiques des véhicules plutôt que sur les caractéristiques des individus réalisant le déplacement). L'affectation stochastique s'appuie sur une définition agrégée de l'utilité, mais parfois également sur une définition désagrégée prenant en compte les caractéristiques des individus ou des véhicules. L'estimation de tels modèles est évidemment beaucoup plus complexe.

Les différentes méthodes d'affectation multichemin se distinguent principalement par quatre dimensions :

- la loi de distribution du terme d'erreur. Une distribution uniforme a été proposée par Burrell. Une distribution normale conduira à un modèle probit (chapitre 9, p. 251) qui pose des problèmes d'estimation comme nous l'avons déjà mentionné. Enfin, la distribution la plus courante est celle de Gumbel qui conduit à la formulation logit multinomial. Une structure hiérarchique est parfois utilisée pour la formulation logit ;
 - le niveau auquel s'applique le terme d'erreur. Dans la méthode multichemin de Dial, la différence de perception s'applique au choix de la route, tandis que dans l'affectation multichemin de Burrell, elle s'applique au niveau du lien lui-même, ce qui interdit pratiquement l'énumération des chemins possibles et donc conduit à des méthodes s'appuyant sur des techniques de simulation ;
 - la procédure de sélection des chemins déterminant l'univers de choix (pour les méthodes, comme celle de Dial, nécessitant une énumération). Elle peut être sélective ou bien viser à une énumération de l'ensemble des chemins acceptables. La définition des chemins acceptables est un problème complexe. En général, on lui donne la définition suivante : un chemin est considéré comme acceptable s'il ne comporte pas de lien conduisant à un « retour en arrière ». Cette condition est nécessaire pour avoir un nombre fini et pas trop élevé de chemins possibles. Deux définitions similaires du « retour en arrière » (mais ne conduisant pas forcément au même univers de choix) sont utilisées :
 - « *The forward pass* » (Ortuzar, Willumsen, 1994) ou l'ordre ascendant dans la définition des chemins possibles. Un chemin entre une origine h et une destination d sera considéré comme acceptable s'il ne comporte pas de lien a dont la destination j soit plus proche que l'origine i de l'origine h du chemin ;
 - « *The backward pass* » (Ortuzar, Willumsen, 1994) ou l'ordre descendant dans la définition des chemins possibles. Un chemin entre une origine h et une destination d sera considéré comme acceptable s'il ne comporte pas de lien a dont l'origine i soit plus proche que la destination j de la destination d du chemin.
- Le terme « plus proche » est entendu au sens du plus court chemin en temps généralisé ;
- les points de choix. On distingue principalement les modèles qui considèrent un unique moment de choix de ceux qui redéfinissent le choix à chaque nœud du réseau. Dans les premiers modèles, le choix porte sur l'ensemble de l'itinéraire entre l'origine h et la destination d du déplacement. En revanche, dans les seconds modèles, la procédure de choix est renouvelée à chaque nœud du réseau et dépend donc des chemins possibles entre le nœud de réseau considéré i et la destination finale d du déplacement. Ces derniers modèles sont généralement utilisés pour la modélisation dynamique, notamment en présence d'information routière (Henn, 2001).

Nous focalisons notre attention sur les deux modèles d'affectation multichemin statiques les plus répandus : ceux de Dial et Burrell (pour une présentation plus large, voir notamment Thomas, 1991).

5.3.2. Affectation multichemin de Dial

Le modèle proposé par Dial (1971) dérive du modèle logit multinomial dont la formulation générale est la suivante :

$$P(h, d, r) = \frac{e^{-\theta c_{hdr}}}{\sum_{r'} e^{-\theta c_{hdr'}}} \quad (10.6)$$

où $P(h, d, r)$ est la probabilité de choix de la route r sur l'origine-destination h-d parmi l'ensemble des routes acceptables r' sur l'origine-destination h-d ;

c_{hdr} est le coût du chemin entre h et d par la route r ;

θ est un paramètre à estimer qui influe sur la dispersion du trafic sur l'origine-destination h-d entre les différentes routes r'.

Ce modèle postule donc l'ensemble des hypothèses nécessaires à l'utilisation d'un logit multinomial (chapitre 9, p. 253). Il postule donc l'indépendance entre les termes d'erreur relatifs aux différentes routes, ce qui est évidemment problématique lorsque des routes partagent en commun une partie importante de leur itinéraire.

Dial propose de présenter l'équation 10.6 sous une forme un peu différente en divisant le numérateur et le dénominateur par $e^{-\theta c_{hd}^*}$ où c_{hd}^* est le plus court chemin pour aller de h à d. L'équation 10.6 devient :

$$P(h, d, r) = \frac{e^{-\theta \delta c_{hdr}}}{\sum_{r'} e^{-\theta \delta c_{hdr'}}} \quad (10.7)$$

où δc_{hdr} est la différence entre le coût de la route r sur h/d et le coût du plus court chemin sur h/d : $\delta c_{hdr} = c_{hdr} - c_{hd}^*$

Exemple d'algorithme pour l'affectation multichemin de Dial

Nous empruntons à Thomas (1991) la présentation de l'algorithme proposé par Fisk (1977) pour la résolution de la méthode multichemin de Dial.

Etape 1 : initialisation

Calcul du plus court chemin.

Etape 2 : calcul du poids de chaque arc et de chaque nœud parmi l'ensemble des chemins acceptables

Etape 2-a : initialisation des poids W de chaque nœud et w de chaque lien

Soit h l'origine de l'arbre de plus court chemin.

$W_h = 1$. $W_i = 0$ pour tout i différent de h.

$w_{ij} = 0$ pour tout i, j.

Etape 2-b : calcul des poids w des liens et W des nœuds

Considérer chaque nœud j du réseau par ordre croissant des plus courts chemins c_j^* depuis h.

Etape 2-b1 : pour chaque nœud j, considérer l'ensemble des nœuds i connectés à j par un lien (i-j) :

Calculer la pertinence e_{ij} du lien (i-j) dans le chemin entre h et j :

$$e_{ij} = \begin{cases} e^{-\theta \Delta_{ij}} & \text{si } c_i^* < c_j^* \\ 0 & \text{si } c_i^* \geq c_j^* \end{cases} \quad (10.8)$$

Cette condition traduit l'hypothèse nécessaire à la détermination des chemins acceptables. Elle traduit l'interdiction d'un « retour en arrière ».

Avec : $\Delta_{ij} = (c_i^* + c_{ij}) - c_j^*$ qui représente le supplément de coût engendré par le passage par le lien (i-j) par rapport au plus court chemin.

Calculer ensuite le poids du lien (i-j) :

$$w_{ij} = W_i e_{ij} \quad (10.9)$$

Etape 2-b2 : déterminer le poids du nœud j :

$$W_j = \sum_i w_{ij} \quad (10.10)$$

Etape 3 : calcul des volumes de trafic

Etape 3-a : initialisation des volumes

Initialiser pour tout nœud j $V_j = 0$.

Initialiser pour chaque nœud d de destination de déplacements $V_d = T_{hd}$

avec T_{hd} le trafic entre le nœud origine de l'arbre de plus court chemin et le nœud d de destination des déplacements.

Etape 3-b : calcul des volumes

Considérer chaque nœud j du réseau par ordre décroissant des plus courts chemins c_j^* .

Pour chaque nœud j, considérer l'ensemble des nœuds i connectés à j par un lien (i-j) :

- calculer le volume V_{ij} sur le lien (i-j) :

$$V_{ij} = V_j * \frac{w_{ij}}{W_j} ; \quad (10.11)$$

- actualiser ensuite V_i :

$$V_i = V_i + V_{ij}. \quad (10.12)$$

Fin de l'algorithme

Il reste à démontrer que cet algorithme conduit effectivement à l'équation 10.7. En reportant les équations 10.9 et 10.10 dans l'équation 10.11, on obtient :

$$V_{ij} = V_j * \frac{W_i e^{\theta \Delta_{ij}}}{\sum_i W_i e^{\theta \Delta_{ij}}} \quad (\text{en considérant que } e^{\theta \Delta_{ij}} = 0 \text{ si } e_{ij} = 0) \quad (10.13)$$

Si l'on reprend l'équation 10.7, on peut écrire :

$$V_{ij} = \sum_{d,r} \left(V_d \frac{e^{-\theta \delta c_{hdr}}}{\sum_{r'} e^{-\theta \delta c_{hdr'}}} \right), \quad (10.14)$$

en posant que $e^{-\theta \delta c_{hdr}} = 0$ si le lien (i-j) n'est pas inclus dans la route r pour aller de h (origine de l'arbre de plus court chemin) à d (destination du déplacement). Il s'agit ainsi de considérer l'ensemble des routes r incluant le lien (i-j) pour aller de h à d et ceci pour chacune des destinations de déplacements d.

On peut alors montrer, pour une origine-destination donnée, en remplaçant successivement d'une part V_j par sa valeur en remontant dans l'ordre croissant des plus courts chemins et d'autre part W_i par sa valeur dans l'ordre décroissant des plus courts chemins, que l'on énumère bien l'ensemble

des routes r incluant le lien $(i-j)$ pour aller de h à d et ceci pour chacune des destinations de déplacements d (ce qui correspond à la somme sur d et r dans l'équation 10.14). Le terme V_j de l'équation 10.13 devient alors V_d ; le terme W_i devient $W_h = 1$, tandis que le terme $e^{\theta\Delta_{ij}}$ devient

$$\prod_{(i-j)} e^{\theta\Delta_{ij}} = e^{\theta \sum_{(i-j)} \Delta_{ij}}$$

ou l'ensemble $\{(i-j)\}$ représente l'ensemble des liens $(i-j)$ d'une route r pour

aller de h à d . Compte tenu de la définition de Δ_{ij} , ce produit devient $e^{-\theta\delta c_{hdr}}$. On retrouve de même le dénominateur, ce qui démontre l'égalité des équations 10.13 et 10.14. L'algorithme de Fisk répond donc bien à la méthode proposée par Dial (on trouvera une démonstration plus formalisée dans Thomas, 1991).

D'autres modèles similaires à celui de Dial, s'appuyant sur une énumération complète des chemins acceptables sont proposés dans Thomas (1991).

5.3.3. Affectation multichemin de Burrell

Les modèles s'apparentant à la méthode proposée par Burrell (1968, nous nous appuyons sur la présentation de Thomas, 1991) stipulent une distribution aléatoire du terme d'erreur (qui traduit notamment les différences de perception selon les individus) au niveau de chacun des liens. De ce fait, la procédure d'énumération des chemins acceptables n'est pratiquement plus possible. Ces modèles ont donc en commun d'utiliser des techniques de simulation pour résoudre le problème.

Pour cela, Burrell (et d'une manière générale l'ensemble des méthodes s'appuyant sur des techniques de simulation) pose les hypothèses suivantes :

- la moyenne de la distribution des coûts perçus pour un lien donné correspond à l'utilité mesurée par le modèle à partir des attributs du lien (ce qui peut aussi s'exprimer la moyenne du terme d'erreur est nulle) ;
- pour chaque individu, la perception des coûts d'un lien est indépendante de la perception des coûts sur tous les autres liens.

Le principe de ces méthodes est le suivant.

Exemple d'algorithme pour l'affectation multichemin de Burrell

Etape 1 : détermination du trafic à chaque itération

Pour chaque itération (3 en Grande Bretagne selon Thomas (1991), 3 à 5 selon Ortuzar, Willumsen (1994)), et pour chaque nœud h origine.

Etape 1a : génération du coût de chaque lien

Pour chaque lien a , générer aléatoirement une valeur pour z_a (la procédure de génération aléatoire doit respecter la distribution du terme d'erreur) et calculer le coût du lien a :

$c_a = \bar{c}_a + z_a$ avec \bar{c}_a le coût moyen (donc estimé par le modèle à partir des attributs du lien) du lien a .

Cette méthode peut générer des coûts négatifs. Pour y faire face, soit on réduit l'écart-type de la distribution, soit on rejette la valeur de z_a lorsque le coût est négatif.

Etape 1b : calcul des plus courts chemins

A l'aide des coûts ainsi estimés pour chacun des liens, calculer l'arbre de plus court chemin.

Etape 1c : affectation du trafic sur ce plus court chemin

Affecter la matrice origines-destinations sur les plus courts chemins pour déterminer le trafic V_a de chacun des liens.

Etape 2 : détermination du trafic final sur chaque lien a

Pour chaque lien, faire la moyenne des trafics V_a estimée pour le lien a à chacune des itérations.

Fin de l'algorithme

Dans son modèle pour simplifier la procédure de calcul, Burrell a retenu une distribution uniforme du terme d'erreur. Cette simplification n'est toutefois pas retenue par l'ensemble des auteurs (Thomas, 1991).

Dans cette méthode, deux éléments conditionnent l'étendue de la dispersion des itinéraires pour une même origine-destination :

- le choix de la distribution et surtout de son écart-type. Un écart-type élevé générera potentiellement davantage de valeur de z_a élevée. De ce fait, un plus grand nombre de chemins pourront être potentiellement considérés comme le plus court chemin à l'étape 1.b ;
- le nombre d'itérations. Plus il y a d'itérations, plus il y aura potentiellement de chemins différents (à condition que l'écart-type de la distribution soit suffisant pour le permettre).

Si l'étendue de la dispersion des itinéraires est plus grande, il y a un risque que la méthode conduise à sélectionner des itinéraires nettement plus coûteux que le plus court chemin (c'est potentiellement le cas lorsque l'écart-type est élevé). Dans ce cas, il est souhaitable d'accroître le nombre d'itérations, le poids de cet itinéraire nettement plus coûteux parmi l'ensemble des itinéraires étant ainsi réduit.

5.3.4. Méthode d'Abraham ou méthode Ariane : logit logarithmique

La méthode d'Abraham, qui est également communément appelée méthode Ariane, car elle est utilisée dans le modèle ARIANE de la Direction des Routes (SETRA), s'appuie sur la loi d'Abraham du nom de son inventeur (Morellet et al., 1980 ; Quinet, 1998). Cette loi permet de répartir le trafic entre deux itinéraires pour une même origine-destination à partir de l'équation suivante :

$$\frac{T_1}{T_2} = \left(\frac{c_1}{c_2} \right)^{-\alpha} \quad (10.15)$$

avec T_1 et T_2 les trafics sur les chemins 1 et 2 ;

c_1 et c_2 les coûts généralisés de ces chemins ;

α un coefficient positif auquel l'analyse statistique donne des valeurs comprises en général entre 8 et 10.

Cette formulation est utilisée en interurbain car l'incidence d'une différence donnée de coût généralisé entre deux itinéraires n'est probablement pas identique selon que le coût généralisé est faible ou à l'inverse très élevée. En revanche, en urbain l'incidence est probablement plus réduite.

En apparence, s'il s'agit bien d'une affectation multichemin, il ne semble pas qu'il s'agisse d'une affectation stochastique. Il semble donc surprenant de classer cette méthode d'affectation dans cette section. En fait, l'interprétation stochastique est facile à montrer (Leurent, 1995a). Considérons la fonction d'utilité U_i suivante :

$$U_i = \theta \ln(c_i) + \varepsilon \quad (10.16)$$

où c_i est le coût généralisé du chemin i ;

θ un paramètre à estimer ;

ε le terme résiduel d'erreur.

Si l'on postule que les résidus de chacun des itinéraires sont indépendants et identiquement distribués selon une loi de Gumbel, la part de marché du chemin i pour une origine-destination donnée, s'écrit :

$$P(i) = \frac{e^{\theta \ln(c_i)}}{\sum_j e^{\theta \ln(c_j)}} = \frac{c_i^\theta}{\sum_j c_j^\theta} \quad (10.17)$$

La propriété IIA d'indépendance vis-à-vis des autres alternatives (chapitre 9, p. 258) permet d'écrire :

$$\frac{T_i}{T_j} = \frac{P(i)}{P(j)} = \left(\frac{c_i}{c_j} \right)^\theta$$

θ est un paramètre qui a forcément un signe négatif dans la fonction d'utilité. De ce fait, on peut poser $\theta = -\alpha$ et l'on retrouve l'équation 10.16.

Cette démonstration conduit Leurent (1995a) a dénommé également cette méthode le logit logarithmique.

Si l'on remonte aux sources de la Loi d'Abraham (Abraham, Coquand, 1961), les hypothèses sur le terme d'erreur sont beaucoup flexibles. Les auteurs proposent la formulation suivante pour l'utilité déterministe :

$$V = l * L + t * T \quad (10.18)$$

où L est la longueur de l'itinéraire ;

T est le temps de parcours probable ou estimé ;

l est « le coût kilométrique unitaire où peuvent intervenir :

- des consommations du véhicule ;

- une appréciation de la sécurité ;

- une prise en compte de l'agrément de conduite ;

t la valeur unitaire du temps pour l'usager considéré, et pour l'itinéraire en cause. » (Abraham, Coquand, 1961)

Les individus n'ayant pas la même évaluation de chacun de ces paramètres, les auteurs définissent les fonctions suivantes pour l'individu n : k_n relative aux kilomètres parcourus et s_n relative au temps de parcours :

$$k_n = l_0 * L_0 + \varepsilon_n$$

$$s_n = t_0 * T_0 + \eta_n \quad (10.19)$$

ε_n et η_n étant des variables aléatoire de moyenne nulle ;

l'indice 0 traduit la valeur moyenne de chacun des paramètres et variables, pour un itinéraire donné.

Les auteurs proposent ensuite une distribution normale pour les termes résiduels ce qui conduit à la formulation probit. Sans vouloir faire de procès d'antériorité, il semble bien que l'on puisse situer Abraham et Coquand parmi les pionniers de l'approche désagrégée et assurément du développement du modèle probit (Leurent, 1995a exprime également ce point de vue). Dans cet article, les auteurs confrontent les résultats du modèles probit avec les différentes formulations alors en vigueur aux

Etats-Unis (méthode californienne (Moskowitz), courbes de Détroit (Campbell, McCargar), courbe de l'AASHO (Barker, 1952), références citées par Abraham et Coquand (1961)) pour mettre en évidence la supériorité du modèle probit sur le plan de la reproduction des données empiriques.

5.3.5. Affectation prix-temps

Le modèle prix-temps (chapitre 2, p. 51) permet d'arbitrer entre une composante monétaire et une composante temporelle intervenant dans la fonction d'utilité dans un contexte de choix discret. L'application que nous avons présentée dans le cas du choix du mode entre l'avion et le train peut être facilement transposée dans le cas du choix entre itinéraires concurrents. Le principe est identique. L'utilisateur choisit l'itinéraire qui possède l'utilité la plus élevée, c'est-à-dire le coût généralisé le plus faible. La formulation de l'utilité pour un itinéraire r entre une origine-destination donnée devient :

$$U_{nr} = V_{nr} = -Cg_{nr} = -(P_r + h_{nr}T_r) \quad (10.20)$$

où U_{nr} et V_{nr} représentent l'utilité du chemin r pour l'individu n . L'utilité n 'est représentée que par sa partie déterministe, ce qui explique l'égalité $U_{nr} = V_{nr}$;

Cg_{nr} est le coût généralisé du chemin r pour l'individu n ;

P_r est le prix moyen du chemin r ;

T_r est le temps moyen de parcours du chemin r (il s'agit d'un temps fixe qui ne dépend pas de la charge comme pour les autres approches développées dans cette section 5.3) ;

h_{nr} est la valeur du temps pour l'individu n relative au déplacement effectué sur l'itinéraire r .

L'égalité $U_{nr} = V_{nr}$ exprime clairement que l'approche n 'est pas stochastique. Nous avons tout de même inclus ce modèle dans cette section, car il s'agit d'une méthode d'affectation multichemin.

La résolution du problème nécessite d'identifier l'ensemble des « chemins efficaces » (Leurent, 1995a), c'est-à-dire susceptible d'être utilisés. La définition est en fait assez simple (nous reprenons la définition proposée par Leurent (1995a) ainsi que le graphique qui permet de l'illustrer, graphique 10.6). « *Un chemin est efficace s'il n'y a pas d'autre chemin à la fois plus rapide et moins cher. On peut repérer les chemins efficaces à l'aide d'un simple diagramme, en portant pour chaque chemin un point de coordonnées : le prix et le temps (c'est la frontière de Pareto du système)* ». Tout autre chemin ne sera pas utilisé, car il existe un chemin dont le prix et le temps est plus faible donc ce chemin aura un coût généralisé plus faible quelle que soit la valeur du temps de l'individu. Dans la pratique, cela suppose tout d'abord d'identifier l'ensemble des prix (ou classes de prix) différents parmi l'ensemble complet des chemins sur l'origine-destination considérée. Ensuite pour chacun des prix ou classes de prix, il faut identifier le chemin de temps minimum.

L'arbitrage entre le prix et le temps est effectué par l'intermédiaire de la valeur du temps. Il est donc nécessaire de postuler une fonction de distribution des valeurs du temps que nous notons $f(h)$. Si l'on range les chemins efficaces dans l'ordre des prix croissants, la valeur du temps d'indifférence HC_k entre les chemins k et $k+1$ est définie par (équation 2.5, p. 51) :

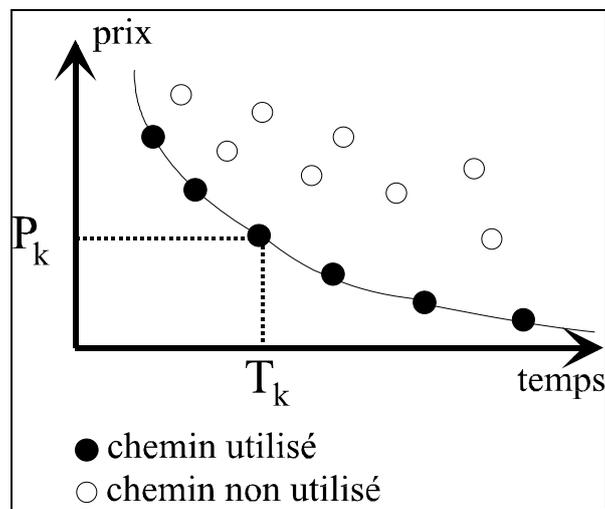
$$HC_k = \frac{P_{k+1} - P_k}{T_k - T_{k+1}} \quad (10.21)$$

avec P_k et P_{k+1} le prix du chemin appartenant à la classe de prix $k, k+1$;

T_k et T_{k+1} le temps du chemin appartenant à la classe de prix $k, k+1$.

L'utilisateur choisissant le chemin qui possède la plus forte utilité compte tenu de sa valeur du temps, le domaine de pertinence du chemin efficace k correspond à l'intervalle de valeur du temps $[HC_{k-1}; HC_k]$. La probabilité de choix du chemin k , $P(k)$ est alors obtenue par (équation 2.7, p. 52) :

$$P(k) = \int_{HC_{k-1}}^{HC_k} f(h)dh \quad (10.22)$$



Graphique 10.6 : Sélection des chemins efficaces dans le modèle prix-temps
(Source : Leurent, 1995a)

La distribution des valeurs du temps étant généralement assimilée à une distribution log-normale, en reprenant l'équation 2.13 (p. 54), la probabilité devient :

$$P(k) = \Phi\left(\frac{\ln(HC_k) - m}{\sigma}\right) - \Phi\left(\frac{\ln(HC_{k-1}) - m}{\sigma}\right) \quad (10.23)$$

où Φ est la distribution cumulative de la loi normale standard ;

σ est l'écart-type de la distribution log-normale des valeurs du temps associées aux déplacements effectués sur une origine-destination donnée ;

m est la moyenne de la distribution log-normale des valeurs du temps associées aux déplacements effectués sur une origine-destination donnée.

En général, pour simplifier le problème, on postule une distribution des valeurs du temps identique pour chacune des origines-destinations.

5.3.6. Principales limites de l'affectation stochastique ou multichemin

La principale limite tient au maintien de l'hypothèse rappelée dans la section 5.3.1 (p. 332) :

- le coût d'un lien est fixe et ne dépend donc pas de sa charge. Le coût d'un lien est toujours calculé à partir du coût à vide du lien.

De ce fait, le coût d'un lien ne tient pas compte de la congestion. Si l'affectation multichemin limite les risques d'une charge excédant fortement la capacité relativement à l'affectation tout-ou-rien, ces risques ne sont pas éliminés. De ce fait, dans des réseaux fortement congestionnés, il n'est pas rare d'avoir des liens supportant une charge supérieure à la capacité, même en utilisant des méthodes favorisant une forte dispersion des itinéraires.

Cette méthode n'est donc pas recommandée pour des réseaux routiers en situation de congestion. En revanche, pour les réseaux routiers non ou peu congestionnés (période creuse par exemple en urbain ou trafic interurbain) et pour les transports collectifs, les méthodes multichemin sont plus pertinentes que l'affectation au plus court chemin. L'intérêt de ces méthodes réside également dans leur capacité à répartir le trafic entre plusieurs chemins tout en conservant une

attractivité accrue pour les chemins les plus courts (définis à partir d'une formulation déterministe de l'utilité). Enfin, ces méthodes ne nécessitent pas la définition de courbes débit-coût pour chacun des liens (puisque le coût ne tient pas compte de la charge). Cet avantage n'est pas mineur dans la mesure où la définition et la validation de telles courbes est une opération qui n'est pas toujours simple.

Au-delà du problème de capacité, certaines limites sont plus spécifiques à certaines méthodes.

Pour l'affectation de Dial (et les méthodes similaires), on considère généralement les limites suivantes :

- les hypothèses relatives au choix de la distribution du terme d'erreur. En particulier, l'utilisation d'un logit multinomial postule l'indépendance entre les termes d'erreur relatifs aux différentes routes, ce qui est évidemment problématique lorsque des routes partagent en commun une partie importante de leur itinéraire ;
- la sensibilité à la définition du réseau. Considérons un réseau simplifié comprenant une rocade et un itinéraire central qui se décompose en deux itinéraires secondaires qui se distinguent par une petite portion de route (schéma 10.7). L'ensemble des chemins conduit à un coût similaire. Dans ce cas, le trafic sera réparti à parts égales entre les trois itinéraires ainsi définis (routes 1 et 2 par le centre et rocade). On pourrait pourtant penser que les routes 1 et 2 par le centre représentent en fait une seule alternative et donc que le trafic devrait être partagé par moitié entre la rocade et les deux routes du centre. D'ailleurs, une définition un peu plus simplifiée du réseau ne faisant pas apparaître ces deux routes dans le centre conduirait à cette solution. Ce problème découle en fait de l'hypothèse précédente qui conduit à la propriété IIA d'indépendance vis-à-vis des autres alternatives (chapitres 8, p. 233 et 9, p. 258). Il est évident que dans cet exemple, cette propriété n'est pas respectée, car les termes d'erreur des alternatives routes 1 et 2 dans le centre ne sont pas indépendants. De ce fait, l'affectation de Dial a tendance à affecter trop de trafic dans les parties du réseau définies de manière très fine (généralement dans le centre) et pas assez dans les parties définies moins finement. Pour y faire face, certains auteurs proposent une structure de choix hiérarchique. Toutefois, dans un réseau comportant un grand nombre de liens, cette procédure est extrêmement complexe à implanter et n'est donc pas très réaliste ;
- la procédure de sélection des itinéraires acceptables élimine des itinéraires que l'on observe pourtant sur le terrain. Face à une situation congestionnée, il n'est pas rare qu'un conducteur choisisse de changer l'itinéraire initialement prévu et de ce fait, fasse ponctuellement un « retour en arrière » (au sens du plus court chemin).

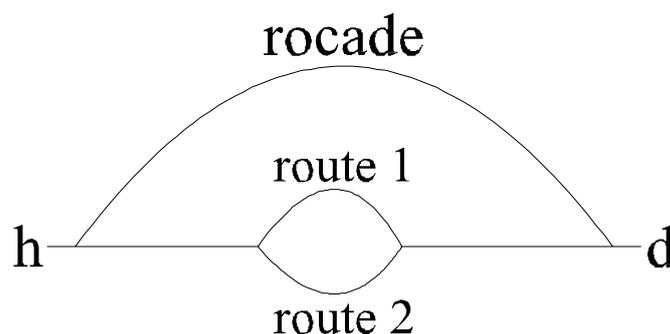


Schéma 10.7 : Illustration de la sensibilité à la définition du réseau pour l'affectation de Dial

Pour l'affectation de Burrell, les limites suivantes sont généralement avancées :

- la principale limite provient de l'hypothèse rappelée section 5.3.3 (p. 336) : pour chaque individu, la perception des coûts d'un lien est indépendante de la perception des coûts sur tous les autres liens. Il est pourtant évident que cette perception n'est pas indépendante et tient compte des préférences de chaque individu : la préférence pour l'utilisation des grands axes

pour éviter de se perdre, ou au contraire l'utilisation fréquente des voies secondaires pour éviter la congestion, une plus ou moins grande propension à éviter les carrefours à feux ou ceux comportant des stops... Pour y faire face, il est nécessaire soit de segmenter la population en fonction des caractéristiques pouvant être communes à un groupe de conducteurs ou d'adopter une approche désagrégée. L'information est toutefois rarement disponible concernant ces préférences relatives au choix d'itinéraire ;

- comme pour toutes les procédures de simulation, les résultats de l'affectation dépendent de la séquence des valeurs de z_a obtenues aléatoirement. Pour y faire face, il est nécessaire d'accroître le nombre d'itérations.

5.4. Affectation à l'équilibre ou avec contrainte de capacité

Les deux familles de modèles précédentes ne prennent pas en compte les phénomènes de congestion. De ce fait, elles sont peu efficaces pour traiter de l'affectation des déplacements pour des réseaux routiers fortement congestionnés. C'est pour répondre à ces problèmes que les modèles que l'on range dans la famille des modèles à l'équilibre ont été développés. Le principe général de ces modèles est de prendre en compte la congestion à travers une définition du coût du lien en fonction de sa charge. En revanche, ils considèrent une approche déterministe de l'utilité à travers un coût unique pour le lien défini à partir des attributs du lien.

L'affectation sur plusieurs chemins n'est plus assurée par une formalisation probabiliste des coûts comme dans l'affectation stochastique, mais par une prise en compte de la capacité des infrastructures. Lorsqu'un lien est trop chargé son coût devient excessif. De ce fait, d'autres itinéraires deviennent plus intéressants (au sens du coût généralisé). Le principe des méthodes à l'équilibre est de faire en sorte que plusieurs itinéraires soient choisis pour une même origine-destination lorsque l'on rencontre des phénomènes de congestion, de telle sorte que ces itinéraires soient d'un coût comparable.

Plusieurs modèles sont inclus dans cette famille. Toutefois, seule l'affectation à l'équilibre de Wardrop répond pleinement à l'objectif visé (section 5.4.1). Les autres modèles plus simples à mettre en œuvre permettent seulement d'approcher la solution obtenue par l'équilibre de Wardrop. Nous nous limiterons à l'affectation par tranche (aussi appelée affectation incrémentale, section 5.4.2) dont l'utilisation est assez fréquente. L'ensemble de ces méthodes d'affectation est parfois également appelé affectation avec contrainte de capacité du fait de la relation entre le coût d'un lien et sa charge.

On trouvera la description d'autres méthodes dans Thomas (1991) ou encore dans Ortuzar et Willumsen (1994). Ces auteurs présentent notamment une méthode appelée « *Method of Successive Averages* » (méthode des moyennes successives) qui est similaire dans son principe à la méthode d'équilibre et que nous décrivons succinctement dans la section 5.5 (p. 354).

5.4.1. Affectation à l'équilibre de Wardrop

Nous présentons tout d'abord les principes de l'affectation selon le premier principe de Wardrop qui respecte la théorie de la maximisation de l'utilité individuelle (section 5.4.1.1). Nous décrivons ensuite l'algorithme de Frank-Wolfe qui permet de résoudre une affectation selon le premier principe de Wardrop (section 5.4.1.2), puis l'algorithme fondé sur la procédure d'égalisation par transvasement (section 5.4.1.3). Nous présentons ensuite rapidement le principe du modèle prix-temps selon l'équilibre de Wardrop (section 5.4.1.4). Wardrop a également explicité un second principe d'affectation qui conduit à une affectation à l'optimum collectif et non plus individuel (section 5.4.1.5). Les résultats distincts de ces deux affectations ont conduit à souligner plusieurs paradoxes dont les plus connus sont les paradoxes de Braess et de Villé (section 5.4.1.6).

5.4.1.1. Principes généraux de l'affectation selon le premier principe de Wardrop

La formulation de la condition d'équilibre est due à Wardrop (1952) : Une affectation est à l'équilibre lorsque dans un réseau congestionné, toutes les routes utilisées sur une même origine-destination ont un coût équivalent et minimum tandis que toutes les autres routes possibles non utilisées ont un coût supérieur. Si cette condition n'est pas remplie, il existe au moins un conducteur qui a intérêt à changer d'itinéraire car il pourra trouver un itinéraire plus court. Dans une affectation à l'équilibre aucun conducteur n'a intérêt à changer d'itinéraire, car en changeant d'itinéraire il aurait un coût égal ou supérieur.

Cette condition est appelée le principe de Wardrop ou l'équilibre de Wardrop. On l'appelle aussi le 1^{er} principe de Wardrop ou le principe d'optimisation individuelle de Wardrop. Wardrop a proposé un second principe appelé principe d'optimisation collective ou sociale de Wardrop (section 5.4.1.5).

En général, plusieurs hypothèses complémentaires sont posées dans les affectations selon l'équilibre de Wardrop :

- le coût sur un lien dépend de la charge sur le lien, mais ne dépend pas de la charge sur les autres liens. Il ne s'agit donc que d'une prise en compte partielle de la congestion. Les phénomènes de remontée de queue ou de diffusion de la congestion par les carrefours ne peuvent donc pas être pris en compte (cette condition n'est toutefois pas impérative, Thomas (1991) présente ainsi certains modèles traitant ces questions. On trouvera également des développements dans (Fernandez, Friesz, 1983 ou Friesz, Bernstein, 2000) en ce qui concerne l'affectation dynamique) ;
- le principe de Wardrop suppose une information parfaite sur les conditions de trafic et donc sur le coût des différents liens. Seule une approche probabiliste des coûts généralisés permet de contourner le problème (section 5.5, p. 354) ;
- enfin les conducteurs choisissent leur itinéraire indépendamment du choix des autres conducteurs de manière à maximiser leur utilité individuelle (on retrouve ici une hypothèse récurrente commune à l'ensemble des modèles d'affectation à la base de la théorie des choix discrets).

A la différence de la plupart des méthodes évoquées jusqu'ici, l'équilibre de Wardrop ne peut être déterminé algébriquement de manière exacte. Il est donc nécessaire de recourir à des algorithmes de résolution. L'algorithme le plus courant est de Frank-Wolfe (1956, pour sa présentation, nous nous appuyons sur l'ouvrage de Thomas (1991) qui lui-même fait référence aux travaux de Van Vliet et Dow (1979). Il existe d'autres algorithmes permettant une affectation à l'équilibre de Wardrop, notamment un algorithme appelé « Procédure d'égalisation par transvasement » décrit par Leurent (1995b) (nous présentons succinctement cet algorithme car il nécessite une énumération de l'ensemble des itinéraires à la différence de l'algorithme de Frank-Wolfe, section 5.4.1.3).

Nous conservons les notations présentées à la section 4.1 (p. 319), ainsi que certaines notations de la section 4.3 (p. 325) :

h est un indice correspondant à l'origine d'un déplacement et d à sa destination ;

c_{hd}^* est le coût du plus court chemin sur l'origine-destination h/d . A la différence de l'affectation stochastique, il s'agit du coût en charge (donc recalculé après avoir affecté le trafic).

Une solution qui respecte le principe de Wardrop doit respecter le système d'équations suivant :

$$\begin{cases} c_{hdr} = c_{hd}^* , \text{ pour tout chemin } hdr \text{ tel que } T_{hdr} > 0 \\ c_{hdr} \geq c_{hd}^* , \text{ pour tout chemin } hdr \text{ tel que } T_{hdr} = 0 \end{cases} \quad (10.24)$$

avec T_{hdr} trafic entre l'origine h et la destination d , via le chemin r .

La résolution mathématique de ce principe est due à Beckman et al. (1956). Ils ont montré que la détermination du trafic sur chacun des liens, de telle sorte que ces trafics respectent le premier principe de Wardrop, est équivalent à minimiser la fonction Z :

$$Z = \sum_a \int_0^{V_a} C_a(v) dv \quad (10.25)$$

en respectant les contraintes :

$$\sum_r T_{hdr} = T_{hd} \quad (10.26)$$

$$\text{et } T_{hdr} \geq 0 \text{ pour tout } h, d \text{ et } r \quad (10.27)$$

avec V_a volume de trafic sur le lien a ;

$C_a(v)$ la fonction reliant le coût au volume pour le lien a .

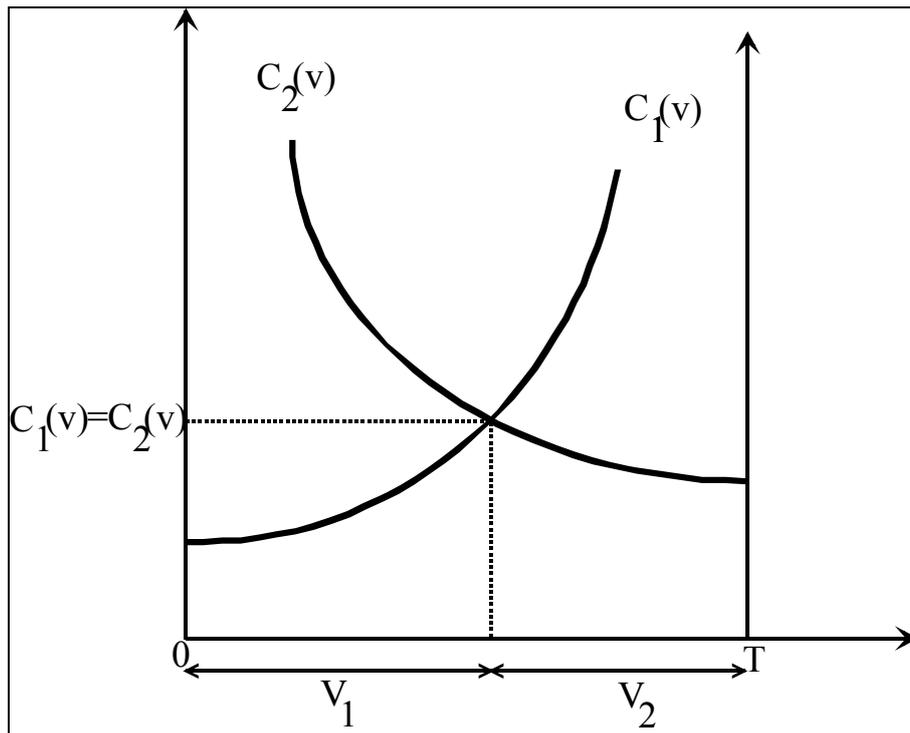
Lorsque les fonctions $C(V)$ sont strictement croissantes, la solution est unique (la démonstration est assez simple. Elle consiste à démontrer que la fonction Z est convexe. Si les fonctions sont seulement croissantes (c'est-à-dire non décroissantes, ce qui correspond effectivement à la réalité), la solution $\{V_a\}$ peut ne pas être unique, même si la solution des $\{c_a\}$ (c_a : coût du lien a) est unique.

L'égalité entre la résolution de l'équilibre de Wardrop et la minimisation de la fonction Z est intuitivement facile à montrer. Si une solution ne respecte pas le principe de Wardrop, c'est qu'il y a un conducteur qui peut changer d'itinéraire en ayant un coût plus faible. Dans ce cas, la fonction Z aura une valeur plus faible. Inversement si une solution respecte l'équilibre de Wardrop, si un conducteur change d'itinéraire le coût de son trajet sera accru. Dans ce cas, si la fonction de coût est strictement croissante, la fonction Z aura une valeur plus grande (si la fonction de coût n'est pas strictement croissante, dans ce cas le coût peut rester identique et donc Z conservera la même valeur, ce qui démontre au passage la non unicité de la solution $\{V_a\}$ dans ce cas). On trouvera une démonstration formalisée dans Ortuzar et Willumsen (1994, elle est assez simple, il suffit de poser le Lagrangien des équations 10.25 à 10.27 et d'annuler les dérivées premières par rapport aux T_{hdr}).

Comme pour tout algorithme, il est nécessaire de définir une règle de fin. La règle la plus simple est de déterminer a priori un nombre d'itérations. Il ne s'agit toutefois pas de la règle la plus pertinente (surtout pour des algorithmes dont la convergence n'est pas strictement croissante avec le nombre d'itérations). Ortuzar et Willumsen (1994) proposent la règle suivante, où le calcul itératif ne s'arrête que lorsque δ est inférieur à une certaine valeur définie a priori :

$$\delta = \frac{\sum_{hdr} T_{hdr} (c_{hdr} - c_{hd}^*)}{\sum_{hd} T_{hdr} c_{hd}^*} \quad (10.28)$$

cet indicateur δ exprime de manière relative le surcoût de la solution à l'itération n par rapport à une solution où tous les conducteurs pourraient se déplacer au coût minimum de chacune des origines-destinations obtenu à l'itération n . En effet $(c_{hdr} - c_{hd}^*)$ représente le surcoût de la route r sur l'origine-destination $h-d$ par rapport au plus court chemin de l'origine-destination $h-d$ obtenu à l'issue de l'itération n . Dans le calcul effectué à l'équation 10.28, l'ensemble des coûts sont recalculés après avoir effectué l'affectation du trafic à l'itération n . Il est important de noter que le plus court chemin calculé à l'itération n n'est pas la solution respectant l'équilibre de Wardrop. Toutefois, lorsque δ est suffisamment petit, cela signifie qu'aucun individu ne peut espérer un gain significatif en changeant d'itinéraire, donc que l'on est proche de la solution.



Graphique 10.7 : Illustration de l'équilibre selon le premier principe de Wardrop

De nombreux autres indicateurs d'arrêts sont proposés par Thomas (1991).

Le graphique 10.7 permet d'illustrer le principe sur un exemple simplifié avec deux itinéraires c_1 et c_2 ayant pour fonction de coûts $C_1(v)$ et $C_2(v)$. Soit T la demande sur l'origine-destination. Les chemins c_1 et c_2 ne supportent que le trafic de cette origine-destination. Soit V_1 et V_2 , les volumes sur les chemins c_1 et c_2 :

$$T = V_1 + V_2 .$$

Les fonctions de coûts des chemins c_1 et c_2 sont représentées sur le graphique 10.7. La fonction $C_1(v)$ a l'origine de son abscisse en 0, tandis que la fonction $C_2(v)$ a son origine en T , le sens de l'abscisse étant inversée pour cette dernière fonction de coût. La solution du premier principe de Wardrop se trouve à l'intersection des deux courbes où $C_1(V_1) = C_2(V_2)$. L'illustration graphique montre que c'est l'unique solution qui minimise la fonction Z (équation 10.25).

5.4.1.2. Un algorithme de résolution de l'équilibre selon le premier principe de Wardrop : l'algorithme de Frank-Wolfe

Etape 1 : initialisation

Soit n (le compteur sur les itérations) = 0.

Pour chaque lien a pour l'itération 0, définir $c_a^{(0)}$, en général, on retient le coût du lien à vide.

A l'aide de l'ensemble $\{c_a^{(0)}\}$ des coûts à vide de chacun des liens, faire une affectation tout-ou-rien, puis affecter la matrice origines-destinations des déplacements $\{T_{hd}\}$.

Soit $F_a^{(0)}$ appelé volume auxiliaire : le volume affecté sur le lien a à l'issue de l'initialisation (itération 0).

$V_a^{(0)} = F_a^{(0)}$: volume affecté sur le lien a à l'issue de l'initialisation (itération 0).

Etape 2 : calcul itératif

$n = n + 1$.

Etape 2a : actualisation du coût des liens en fonction du volume

$$c_a^{(n)} = c_a(V_a^{(n-1)}).$$

Etape 2b : détermination du volume auxiliaire

Avec les nouveaux $\{c_a^{(n)}\}$, faire une affectation tout-ou-rien, puis affecter la matrice origines-destinations des déplacements $\{T_{hd}\}$.

$F_a^{(n)}$ est le volume affecté sur le lien a à l'issue de l'itération n .

Etape 2c : calcul des nouveaux volumes sur les liens a

$$V_a^{(n)} = V_a^{(n-1)} + \lambda^{(n)}(F_a^{(n)} - V_a^{(n-1)}) \quad (10.29)$$

où $\lambda^{(n)}$ est choisi de telle sorte qu'il minimise :

$$Z^{(n)} = Z(\lambda^{(n)}) = \sum_a \int_0^{V_a^{(n)}} c_a(v) dv, \text{ avec } 0 \leq \lambda^{(n)} \leq 1 \quad (10.30)$$

$Z^{(n)}$ est la fonction à minimiser, qui lorsque le minimum est atteint permet de satisfaire l'équilibre de Wardrop (équation 10.26).

Etape 2d : test de la fin de l'algorithme

Si la règle d'arrêt est respectée, fin de l'algorithme, sinon retour à l'étape 2.

Fin de l'algorithme

Le choix de $\lambda^{(n)}$ est important pour assurer la rapidité de la convergence de l'algorithme. Thomas (1991) propose l'une des deux méthodes exposées ci-dessous. En fait, sachant que Z est convexe n'importe quel algorithme de minimisation d'une fonction convexe définie seulement un segment ($[0 \leq \lambda^{(n)} \leq 1]$) peut s'appliquer.

Le principe de la première méthode est très simple, mais ne converge pas forcément très rapidement. Il consiste à couper l'intervalle initial en deux parties égales et à regarder dans quel sous intervalle est incluse la valeur de $\lambda^{(n)}$ qui minimise Z (nous notons Z par simplification au lieu de $Z(\lambda^{(n)})$). L'opération est renouvelée tant que l'intervalle n'est pas suffisamment petit, donc que la précision de $\lambda^{(n)}$ n'est pas suffisante.

Le principe de la seconde méthode est similaire, mais permet de définir des segments convergeant plus rapidement (le principe est explicité à l'étape 2 du deuxième algorithme ci-dessous). Toutefois, dans certains cas, il ne permet pas d'obtenir de solution.

Les deux algorithmes nécessitent de calculer la dérivée de Z :

$$\text{Posons } I_a(V_a^{(n)}) = \int_0^{V_a^{(n)}} c_a(v) dv$$

$$Z(\lambda^{(n)}) = \sum_a I_a(V_a^{(n)}) = \sum_a I_a \left\{ V_a^{(n-1)} + \lambda^{(n)} (F_a^{(n)} - V_a^{(n-1)}) \right\} \quad (10.31)$$

Donc :

$$Z'(\lambda^{(n)}) = \sum_a (F_a^{(n)} - V_a^{(n-1)}) * c_a \left\{ V_a^{(n-1)} + \lambda^{(n)} (F_a^{(n)} - V_a^{(n-1)}) \right\} \quad (10.32)$$

$F_a^{(n)}$ et $V_a^{(n-1)}$ étant connus (Cf. algorithme précédent $F_a^{(n)}$ est calculé à l'étape 2b de l'itération en cours et $V_a^{(n-1)}$ est le résultat de l'itération précédente), Z' peut être calculé facilement (ce qui n'est pas le cas de Z) pour toutes valeurs $\lambda^{(n)}$ (noté λ dans la suite pour simplifier les notations).

Etape 1 : initialisation

$\lambda = 0$; calcul de $Z'(\lambda)$; si $Z'(0) > 0$, alors $\lambda = 0$ car Z' est une fonction croissante ; fin de l'algorithme.

$\lambda = 1$; calcul de $Z'(\lambda)$; si $Z'(1) < 0$, alors $\lambda = 1$ car Z' est une fonction croissante ; fin de l'algorithme.

On définit $\text{inf} = 0$ (pour borne inférieure de l'intervalle) et $\text{sup} = 1$ (pour borne supérieure de l'intervalle).

Etape 2 : détermination de l'intervalle contenant λ

$m = (\text{inf} + \text{sup})/2$.

Calcul de $Z'(m)$.

Si $Z'(m) = 0$; $\lambda = m$; fin de l'algorithme.

Si $Z'(m) < 0$; $\text{inf} = m$.

Si $Z'(m) > 0$; $\text{sup} = m$.

Si $(\text{sup} - \text{inf}) < \text{au test d'arrêt}$ (c'est-à-dire un intervalle suffisamment petit pour que λ soit connu avec la précision voulue) ; fin de l'algorithme ; sinon retour au début de l'étape 2.

Fin de l'algorithme

Le second algorithme est similaire dans son principe :

Etape 1 : initialisation

$\lambda_0 = 0$; calcul de $Z'(\lambda_0)$; si $Z'(0) > 0$, alors $\lambda = 0$ car Z' est une fonction croissante ; fin de l'algorithme.

$\lambda_1 = 1$; calcul de $Z'(\lambda_1)$; si $Z'(1) < 0$, alors $\lambda = 1$ car Z' est une fonction croissante ; fin de l'algorithme.

$r = 1$.

Etape 2 : détermination de λ

$r = r + 1$.

$$\text{Calcul de } \lambda_{(r)} = \lambda_{(r-1)} - Z'(\lambda_{(r-1)}) * \left\{ \frac{(\lambda_{(r-1)} - \lambda_{(r-2)})}{Z'(\lambda_{(r-1)}) - Z'(\lambda_{(r-2)})} \right\} \quad (10.33)$$

Il s'agit en fait de trouver une valeur entre $\lambda_{(r-1)}$ et $\lambda_{(r-2)}$ qui permette une convergence plus rapide que le choix d'une valeur moyenne. Le principe est assez simple si l'on veut le représenter graphiquement. Il s'apparente à la détermination du point où les droites représentées par les dérivées de Z aux points $\lambda_{(r-1)}$ et $\lambda_{(r-2)}$ se coupent et de choisir ce point (en fait, il ne s'agit pas rigoureusement de ce point qui nécessiterait de calculer la valeur de Z aux points $\lambda_{(r-1)}$ et $\lambda_{(r-2)}$, mais une valeur proche obtenue en égalisant $Z(\lambda_{(r-1)})$ et $Z(\lambda_{(r-2)})$).

Si $\lambda_{(r)} < 0$ ou $\lambda_{(r)} > 1$, l'algorithme n'est pas convergent ; fin de l'algorithme et utilisation de l'algorithme précédent.

Si $|\lambda_{(r)} - \lambda_{(r-1)}| <$ au test d'arrêt (c'est-à-dire un intervalle suffisamment petit pour que λ soit connu avec la précision voulue) ; fin de l'algorithme ; sinon retour au début de l'étape 2.

Fin de l'algorithme

5.4.1.3. Principe de l'algorithme fondé sur la procédure d'égalisation par transvasement

Le principe est très simple, même si sa mise en œuvre algorithmique ne l'est pas autant. La première étape comme pour l'algorithme de Frank-Wolfe décrit consiste à réaliser une affectation au plus court chemin.

Aux itérations suivantes, on réalise de nouveau un plus court chemin en tenant compte de la charge sur les liens obtenue à l'itération précédente pour estimer le coût de chacun des liens. On identifie alors l'ensemble des « chemins possibles » pour chacune des origines-destinations. Pour une origine-destination donnée, ces « chemins possibles » correspondent à l'ensemble des chemins identifiés comme plus court chemin au cours de l'une des itérations.

Ensuite pour chacune des origines-destinations, la procédure de transvasement est mise en place parmi les « chemins possibles ». Si tous les chemins utilisés ont un coût identique et que les chemins non utilisés ont un coût supérieur, le principe de Wardrop est respecté sur l'origine-destination. Sinon, une partie du trafic sur le chemin présentant le coût le plus élevé est transféré sur le chemin ayant le coût le plus faible. L'opération est renouvelée jusqu'au respect du principe de Wardrop sur l'origine-destination traitée. Si le test d'arrêt n'est pas respecté, on passe alors à l'itération suivante.

On trouvera une description de l'algorithme proposé par Schittenhelm (1990) dans Leurent (1995b). Leurent (1995b) montre également, sur un exemple, sa supériorité par rapport à l'algorithme de Frank-Wolfe et à l'algorithme de la méthode des moyennes successives que nous présentons section 5.5 (p. 354). Leurent (1995c) démontre également que cet algorithme permet de traiter d'autres méthodes d'affectation telles que les affectations stochastiques multichemins, les affectations prix-temps, les affectations à l'équilibre avec utilité aléatoire ou encore les méthodes d'affectation avec demande élastique (la demande pour une origine-destination dépend du coût de l'origine-destination).

5.4.1.4. Affectation prix-temps selon le premier principe de Wardrop

Le principe de l'affectation prix-temps selon le premier principe de Wardrop combine d'une part la méthode exposée pour l'affectation prix-temps (section 5.3.5, p. 339) dans le cas où le coût des liens ne dépend pas de la charge avec le principe de l'affectation à l'équilibre de Wardrop défini plus haut faisant au contraire intervenir la charge dans la détermination du coût du lien.

Le premier principe de Wardrop ne se présente toutefois plus tout à fait d'une manière identique puisque tous les individus n'ont pas la même valeur du temps. En revanche, pour une valeur du temps donnée, le principe peut s'énoncer de manière identique. Une affectation est à l'équilibre lorsque dans un réseau congestionné, toutes les routes utilisées, *par les individus ayant la même valeur du temps*, sur une même origine-destination ont un coût équivalent et minimum tandis que toutes les autres routes possibles non utilisées ont un coût supérieur. Le problème est que l'on ne peut pas résoudre le problème pour chacune des valeurs du temps. Il s'agit donc de transposer la méthode que nous avons exposée pour l'affectation prix-temps, dans le cas où le coût des liens ne dépend pas de la charge, pour y appliquer le principe de Wardrop.

Leurent (1993) propose d'utiliser l'algorithme de la méthode des moyennes successives que nous présentons plus loin (section 5.5, p. 354) et pour chacune des itérations de réaliser une affectation prix-temps. Pour cela, il convient de définir à chaque itération :

- l'ensemble des chemins efficaces. On procède comme nous l'avons exposée section 5.3.5 (p. 339) : tout d'abord identification de l'ensemble des prix (ou classes de prix) différents parmi l'ensemble complet des chemins sur l'origine-destination considérée ; ensuite pour chacun des prix ou classes de prix, identification du chemin de temps minimum ; enfin classement de l'ensemble de ces chemins par prix croissants ;
- la valeur du temps d'indifférence HC_k entre les chemins des classes successives de prix k et

$$k+1 \text{ est définie par (équation 10.21) : } HC_k = \frac{P_{k+1} - P_k}{T_k - T_{k+1}}$$

avec P_k et P_{k+1} le prix du chemin appartenant à la classe de prix $k, k+1$;

T_k et T_{k+1} le temps du chemin appartenant à la classe de prix $k, k+1$.

- la demande affectée sur l'intervalle de valeur du temps $[HC_{k-1}; HC_k]$. La probabilité de choix du chemin k , $P(k)$ est alors obtenue par (équation 10.22) :

$$P(k) = \int_{HC_{k-1}}^{HC_k} f(h)dh$$

où $f(h)$ est la fonction de densité des valeurs du temps. On retient généralement une distribution log-normale, ce qui conduit à l'expression suivante pour la probabilité (équation 10.23) :

$$P(k) = \Phi\left(\frac{\ln(HC_k) - m}{\sigma}\right) - \Phi\left(\frac{\ln(HC_{k-1}) - m}{\sigma}\right)$$

Leurent (1993) démontre que cet algorithme est convergent et conduit vers une solution qui respecte le premier principe de Wardrop. Il peut donc maintenant être défini de la manière suivante : il est possible de définir une succession d'intervalles de valeur de temps au sein desquels on peut appliquer le premier principe de Wardrop. Leurent (1995c) montre que l'algorithme fondé sur la procédure d'égalisation par transvasement que nous avons présenté dans la section précédente peut également s'appliquer. Barbier de Saint Hilaire (2001) propose un algorithme proche de ce dernier pour l'affectation TRIBUT (affectation prix-temps à l'équilibre de Wardrop) dans le logiciel DAVISUM).

5.4.1.5. Second principe de Wardrop

Le second principe de Wardrop également appelée principe d'affectation à l'équilibre selon l'optimum collectif stipule que tous les conducteurs choisissent leur itinéraire de telle sorte que le coût total des déplacements sur le réseau est minimum. On suppose ici qu'au lieu de rechercher une optimisation individuelle, les conducteurs coopèrent entre eux pour obtenir un optimum collectif de telle sorte que le coût moyen soit minimum.

En général, l'affectation selon le second principe de Wardrop conduit à une solution différente du premier principe. Cette solution n'est évidemment pas très réaliste. Elle contredit de plus le principe de maximisation de l'utilité individuelle. Toutefois, si cette affectation n'est pas très réaliste, il est intéressant de comparer les affectations selon les deux principes de Wardrop et de calculer le coût total relatif aux deux affectations. La différence entre le coût total selon le premier principe et le coût total selon le second principe est une quantification du surcoût supporté par les conducteurs par rapport à un réseau qui comporterait des mesures de régulation conduisant les usagers à se conformer au résultat de l'affectation selon le second principe. Cette analyse peut indiquer des pistes pour une meilleure régulation.

La traduction mathématique du second principe est très simple. Il s'agit de minimiser la fonction S :

$$S(T_{\text{hdr}}) = \sum_a V_a C_a(v) \quad (10.34)$$

sous les conditions des équations 10.26 et 10.27, rappelées ci-dessous :

$$\sum_r T_{\text{hdr}} = T_{\text{hd}}$$

et $T_{\text{hdr}} \geq 0$ pour tout h, d et r.

L'équation 10.34 peut également s'écrire de la façon suivante (Thomas, 1991 ; Ortuzar, Willumsen, 1994) :

$$S = \sum_a \int_0^{V_a} C_{m_a}(v) dv \quad (10.35)$$

où C_{m_a} est le coût marginal du lien a.

La résolution de ce système d'équations peut s'effectuer avec l'algorithme de Frank-Wolfe en remplaçant la fonction Z de l'équation 10.25 par la fonction S de l'équation 10.35. Il faut toutefois que les fonctions $C_a'(v)$ soient croissantes (et même strictement croissantes pour garantir l'unicité de la solution). Cette condition est souvent respectée, mais pas systématiquement, comme l'indique l'exemple des courbes proposées dans le logiciel CartOPERA (graphique 10.5) pour une charge dépassant la capacité. Dans le cas de la courbe proposée dans le logiciel DAVISUM (graphique 10.4), la dérivée est croissante, mais est constante au-delà d'un rapport volume sur capacité de 0,9 ne garantissant pas l'unicité de la solution.

En poursuivant le parallèle établi ci-dessus avec la fonction Z, la solution est obtenue lorsque le coût marginal de chacun des itinéraires pour une origine-destination donnée est identique et minimum. Il ne s'agit donc plus de minimiser le coût moyen de chacun des itinéraires comme dans le premier principe de Wardrop conduisant à l'optimisation individuelle, mais de minimiser le coût marginal pour obtenir l'optimum collectif. On retrouve là un principe de tarification bien connu des économistes. Il serait donc plus efficace, d'un point de vue collectif, d'agir sur les coûts des liens (par exemple par du péage) pour garantir cette égalité des coûts marginaux.

L'analyse des différences entre les affectations à l'équilibre selon les deux principes de Wardrop a donné naissance à des paradoxes dont les plus connus sont les paradoxes de Braess et de Villé.

5.4.1.6. Paradoxes de Braess et de Villé

Le paradoxe de Braess met en évidence l'inefficacité de la construction de certaines infrastructures dans un réseau sous certaines configurations de trafic. Pour illustrer ce paradoxe, nous considérons le réseau illustré par le schéma 10.8. Ce réseau comporte 4 nœuds (1, 2, 3 et 4) et 4 liens (1-2, 1-3, 2-4, 3-4) le lien 2-3 n'existant pas dans un premier temps. Ce réseau supporte uniquement un volume de trafic de 18 unités allant de 1 à 4.

On vérifiera aisément que la seule solution respectant le premier principe de Wardrop est obtenue pour un partage égal entre les chemins 1-2-4 et 1-3-4. Dans ce cas, le coût de chacun des chemins est identique :

$$\text{chemin 1 : 1-2-4, 9 unités, coût du chemin 1} = (3 + 3*9) + (36 + 9) = 75 ;$$

$$\text{chemin 2 : 1-3-4, 9 unités, coût du chemin 2} = (36 + 9) + (3 + 3*9) = 75.$$

Le coût moyen d'un chemin est donc de 75 et le coût total $75*18 = 1350$.

Le déplacement d'une unité d'un chemin sur l'autre conduit en effet à des coûts supérieurs pour les usagers du chemin le plus chargé qui n'ont donc aucun intérêt à ce changement. De plus, la perte totale des usagers du chemin le plus chargé est supérieure au gain total des usagers du chemin moins chargé, ce qui conduit à un coût total sur le réseau $= ((3 + 3*10) + (36 + 10))*10 + ((36+8) + (3 + 3*8))*8 = 1358$. La solution du premier principe de Wardrop satisfait donc également le second principe de Wardrop.

Supposons que l'on cherche à renforcer la capacité du réseau en construisant le lien supplémentaire 2-3 qui a une fonction de coût $= 6 + V$. On constatera facilement que la solution du second principe de Wardrop reste inchangée. En revanche, le nouveau chemin 1-2-3-4 est d'un coût plus faible que les deux autres chemins 1-2-4 et 1-3-4. Dans une affectation selon le premier principe de Wardrop, des usagers de ces chemins ont donc intérêt à changer d'itinéraire. On constate facilement qu'une solution selon le premier principe de Wardrop conduit à la répartition suivante entre les trois chemins :

$$\text{chemin 1 (1-2-4), 8 unités, coût du chemin 1} = (3 + 3*10) + (36 + 8) = 77 ;$$

$$\text{chemin 2 (1-3-4), 8 unités, coût du chemin 2} = (36 + 8) + (3 + 3*10) = 77 ;$$

$$\text{chemin 3 (1-2-3-4), 2 unités, coût du chemin 3} = (3 + 3*10) + (9 + 2) + (3 + 3*10) = 77.$$

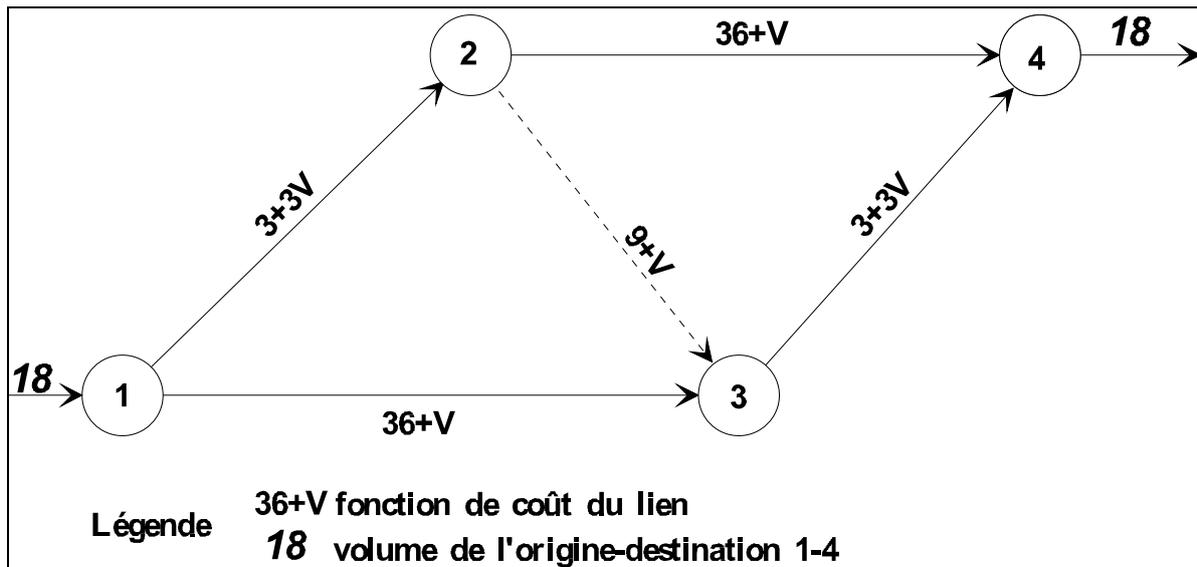


Schéma 10.8 : Illustration du paradoxe de Braess
 (Source : note de cours du Professeur Chapleau, Ecole Polytechnique de Montréal)

Les trois chemins ont un coût identique et tout changement conduit à la détérioration de la situation d'au moins un usager. Le coût total est de 1386. La situation du réseau s'est donc détériorée par rapport à la situation antérieure. La création d'un lien supplémentaire sur le réseau sans le but d'améliorer la situation du réseau a conduit en fait à une détérioration du réseau.

Quinet (1998) mentionne le paradoxe de Villé (1970). « Deux itinéraires assurent le même déplacement entre une origine et une destination urbaines. Le premier est constitué par exemple par une rocade longue et peu encombrée, et le coût généralisé y est donc indépendant du trafic. Le deuxième est un itinéraire urbain court et encombré. Le partage des usagers se fait entre les deux de telle manière que les coûts généralisés sur chacun d'eux soient égaux. Une analyse « myope » de l'encombrement du deuxième itinéraire peut inciter à l'élargir pour réduire la congestion. Après l'élargissement, la répartition du trafic va changer, et les utilisateurs de l'itinéraire urbain vont être plus nombreux ; mais leur coût de transport ne va pas changer puisqu'il est toujours égal au coût de l'itinéraire long, qui est constant. L'élargissement n'aura fait que déplacer des usagers d'un itinéraire sur l'autre, sans gain de temps. L'exemple est encore plus probant si l'itinéraire alternatif de coût stable est en fait constitué par une ligne de transport public. Alors la baisse du trafic sur cette ligne conduit logiquement l'opérateur à baisser ses fréquences, et le coût généralisé sur le transport public augmente, donc le coût sur l'itinéraire routier urbain aussi, et finalement la situation s'est globalement détériorée. » (Quinet, 1998).

5.4.2. Affectation par tranche ou incrémentale

Le principe de l'affectation par tranche est très simple. Il consiste à segmenter la matrice de demande en plusieurs tranches et d'affecter chacune des tranches à l'aide d'une affectation au plus court chemin tenant compte du trafic déjà chargé aux itérations précédentes. Pour la première tranche, on retient le coût à vide. Ce principe peut être implémenté à l'aide de l'algorithme suivant.

Exemple d'algorithme pour l'affectation par tranche ou incrémental

Etape 1 : initialisation

Pour chaque lien a , définir $c_a^{(0)}$, en général, on retient le coût du lien à vide.

Initialiser $V_a^{(0)} = 0$ pour tous les liens a.

Définir le nombre N d'itérations.

Définir les tranches $p^{(n)}$, de telle sorte que $\sum_n p^{(n)} = 1$.

$n = 0$.

Etape 2 : calcul du volume affecté à chaque itération

$n = n + 1$.

A l'aide de l'ensemble $\{c_a^{(n-1)}\}$ des coûts de chacun des liens, faire une affectation tout-ou-rien, puis affecter la tranche de la matrice origines-destinations des déplacements $p^{(n)}\{T_{hd}\}$.

Soit $F_a^{(n)}$ appelé volume auxiliaire : le volume affecté sur le lien a à l'issue de l'itération n.

$$V_a^{(n)} = V_a^{(n-1)} + F_a^{(n)}. \quad (10.36)$$

Calculer pour chacun des liens $c_a^{(n)}$ avec le nouveau trafic $V_a^{(n)}$.

Si $n < N$, reprendre au début de l'étape 2, sinon fin de l'algorithme.

Fin de l'algorithme

Cet algorithme permet d'affecter le trafic sur plusieurs itinéraires pour chacune des origines-destinations (le nombre d'itinéraires est au plus égal au nombre de tranches). Lorsque le réseau est fortement congestionné, il est souhaitable d'accroître le nombre de tranches. Il est recommandé de définir les tranches avec un poids dégressif, par exemple : 40 %, 25 %, 15 %, 10 %, 5 %, 5 %. Si le nombre de tranches et les pourcentages sont adaptés à la congestion du réseau, l'affectation par tranche permet de limiter les risques de charges dépassant fortement la capacité de certains liens comme dans l'affectation au plus court chemin et parfois les affectations stochastiques.

Lorsque le nombre de tranches s'accroît, cette affectation ne converge pas forcément vers une affectation à l'équilibre (Ortuzar, Willumsen, 1994). Ce résultat découle du fait que lorsqu'un trafic est affecté, il ne peut être réaffecté ensuite si l'itinéraire sur lequel le trafic a été affecté s'avère d'un coût supérieur à d'autres itinéraires une fois l'ensemble des itérations terminées. Ce problème est assez fréquent pour les premières itérations. C'est pourquoi certains auteurs préconisent une fois les N itérations terminées, d'ôter le trafic des premières itérations et de renouveler l'opération avec des tranches plus fines (car les première tranches sont en général plus importantes) pour ce trafic qui a été enlevé.

Une autre critique porte sur l'instabilité de la méthode d'une part en fonction du nombre de tranches choisies et du pourcentage de chacune de ces tranches et d'autre part en fonction de la définition du réseau. Cette instabilité provient du fait qu'à chaque itération un seul itinéraire est choisi et qu'une fois affecté, le choix de l'itinéraire n'est plus remis en cause comme dans le cas de l'affectation à l'équilibre. De ce fait, deux affectations comportant un même nombre de tranches, mais ne différant que marginalement au niveau des pourcentages de chacune des tranches ou au niveau de la définition du réseau peuvent conduire à des résultats dont les différences ne sont plus marginales (Thomas, 1991).

Malgré ces limites, cette méthode est couramment utilisée, car elle est très facile à mettre à en œuvre.

5.5. Affectation à l'équilibre avec utilité stochastique

Pour cette présentation, nous empruntons largement à Ortuzar et Willumsen (1994). Le lecteur souhaitant approfondir cette question pourra consulter Thomas (1991) ou Sheffi (1985)

L'affectation stochastique multichemin permet de tenir compte des différences de perception entre les individus et d'autres sources d'erreurs potentielles relatives à l'estimation de l'utilité, mais ne tient pas compte des contraintes de capacité, ni de la charge sur les liens dans l'estimation de l'utilité. A l'inverse, l'affectation à l'équilibre de Wardrop tient compte de la relation entre les coûts des liens et leur charge, mais adopte une approche déterministe de l'utilité. C'est pour répondre à ces deux objectifs que sont développées les méthodes d'affectation à l'équilibre avec une approche probabiliste de l'utilité. Ces méthodes sont baptisées « *Stochastic User Equilibrium Assignment* » ou SUE en anglais.

En reprenant le premier principe de Wardrop, une affectation à l'équilibre avec utilité stochastique repose sur les principes suivants : chaque usager choisit le chemin qui lui permet de minimiser le coût perçu. De ce fait, pour chaque usager le chemin choisi est celui qui présente le coût perçu le plus faible et donc aucun usager ne peut trouver un chemin lui offrant un coût perçu plus faible.

On se trouve donc effectivement dans une situation d'équilibre puisqu'aucun conducteur n'a intérêt à changer d'itinéraire tout en respectant le principe de maximisation de l'utilité individuelle.

En théorie, les modèles répondant à ce principe sont particulièrement attractifs, puisqu'ils permettent d'abandonner les deux conditions restrictives (perception déterministe des coûts et non prise en compte de la charge dans la détermination des coûts, tableau 10.2, p. 329) dont nous devons tenir compte dans les sections précédentes. Leur mise en œuvre est toutefois complexe et doit faire face à des problèmes pratiques. Le principal d'entre eux est le problème de convergence des algorithmes devant être mis en œuvre. Sheffi (1985) a en effet montré que certaines conditions étaient en général nécessaires à cette convergence.

L'algorithme suivant exposé par Ortuzar et Willumsen (1994) peut permettre cette convergence. La convergence n'est toutefois pas très rapide (au minimum une cinquantaine d'itérations selon Sheffi (1985)). Cet algorithme est appelé « méthode des moyennes successives » (« *Method of Successive Averages* » (MSA)). Son principe est similaire à l'algorithme de Frank-Wolfe, mais le coefficient λ de l'étape 2c n'est pas optimisé, ce qui assure la (lente !) convergence.

Exemple d'algorithme pour la méthode des moyennes successives

Etape 1 : initialisation

Soit n (le compteur sur les itérations) = 0.

Pour chaque lien a pour l'itération 0, définir $c_a^{(0)}$, en général, on retient le coût du lien à vide et $V_a^{(0)} = 0$.

Etape 2 : affectation

$n = n + 1$.

Faire une affectation multichemin (de Dial ou de Burrell par exemple) en affectant la totalité de la matrice de demande (Thomas (1991) ou Sheffi (1985) pour avoir des précisions sur les meilleures stratégies à mettre en œuvre pour les affectations multichemins).

Soit $F_a^{(n)}$ le volume ainsi affecté, appelé volume auxiliaire comme dans les autres algorithmes.

Etape 3 : calcul des nouveaux volumes

Calcul des nouveaux volumes à partir de l'équation suivante :

$$V_a^{(n)} = (1 - \lambda)V_a^{(n-1)} + \lambda F_a^{(n)} \quad (10.37)$$

Etape 4 : calcul des nouveaux coûts des liens

Calcul des nouveaux coûts des liens $c_a^{(n)}$ tenant compte du trafic $V_a^{(n)}$ sur ces liens.

Si les coûts $c_a^{(n)}$ n'ont pas trop changé par rapport à l'itération précédente, fin de l'algorithme ; sinon retour à l'étape 2. On trouvera une discussion des règles d'arrêt dans Thomas (1991).

Fin de l'algorithme

En général, on choisit $\lambda = 1/n$, ce qui permet de donner un poids égal à chacune des itérations, d'où le nom de la méthode. L'expérience montre toutefois que cette méthode ne converge pas forcément continûment vers la solution. De ce fait, il faut éviter d'avoir un test d'arrêt portant sur le nombre d'itérations, car on peut avoir une solution moins pertinente que celle d'une des précédentes itérations. La convergence est d'autant moins rapide que le réseau est fortement congestionné. Mais Sheffi (1985) a montré que pour des réseaux fortement congestionnés, l'affectation à l'équilibre de Wardrop (avec utilité déterministe) conduit à des résultats très proches de l'affectation à l'équilibre avec utilité stochastique. De ce fait, l'utilisation de l'affectation à l'équilibre avec utilité stochastique n'est vraiment avantageuse que pour des réseaux qui ne sont pas trop congestionnés.

6. PRINCIPALES LIMITES DE L'ETAPE D'AFECTATION

Pour cette dernière section, nous conservons le plan habituel qui a maintenant fait ses preuves pour structurer nos remarques. Nous organisons donc notre analyse de l'étape d'affectation à travers les trois conditions de l'opérationnalité des modèles (chapitre 2, p. 45) :

- pertinence des modèles. Nous revenons tout d'abord sur les hypothèses nécessaires aux différentes familles de modèles et notamment sur les deux principales : perception des coûts et prise en compte des contraintes de capacité (section 6.4). L'équilibre de Wardrop est aujourd'hui la principale méthode utilisée pour l'affectation routière. Mais la notion même d'équilibre est remise en cause par certains auteurs (section 6.5) ;
- cohérence des modèles. Nous nous limitons ici à la cohérence interne des modèles. Les méthodes d'affectation nécessitent le plus souvent des procédures itératives ne conduisant pas forcément à une solution unique. Les questions de convergence et des conditions nécessaires à cette convergence méritent donc une attention particulière (section 6.6). La prévision des données d'entrée ou des coefficients de calage renvoie évidemment à des questions de pertinence comme nous l'avons déjà évoquée dans les précédents chapitres. Mais en ce qui concerne notamment la définition des réseaux, elle pose également des problèmes redoutables de cohérence (section 6.7) ;
- mesurabilité des modèles. C'est probablement dans cette étape que les problèmes de mesurabilité se posent avec le plus d'acuité à travers : le zonage (section 6.2) et la définition des réseaux et de leurs attributs (section 6.3). Nous abordons les questions de mesurabilité avant les autres dimensions de l'opérationnalité des modèles pour suivre le plan du chapitre.

Mais avant nous souhaitons aborder la question du calage et de la validation de l'étape d'affectation. A la différence des chapitres précédents, nous n'avons pas abordé cette dimension lors de la présentation des méthodes elles-mêmes. Ce choix tient au fait que les procédures de calage et de validation ne dépendent pas directement des méthodes d'affectation elles-mêmes, car il ne s'agit pas seulement de caler les coefficients de ces modèles. De plus, ce calage ne peut s'appuyer facilement sur des indicateurs statistiques (section 6.1).

6.1. Calage et validation

Le principe du calage d'un modèle est de parvenir à reproduire un jeu de données en estimant un certain nombre de paramètres du modèle. La validation consiste à s'assurer que les résultats produits par le modèle sont conformes à ce que l'on connaît lorsque l'on fait varier certaines des données d'entrée. La difficulté principale du calage de l'étape d'affectation est dû au fait qu'il ne s'agit pas seulement d'estimer les paramètres de certaines équations. C'est en fait l'ensemble de la définition du réseau qui doit être calé puis validé. Lors des étapes précédentes, nous disposions de données pour assurer ce calage et de tests statistiques permettant d'apprécier la qualité de la reproduction. Ici, les données sont partiellement manquantes et les indicateurs statistiques que l'on peut construire sont généralement indirects.

Lors de cette étape, il est nécessaire de caler l'ensemble des attributs des fonctions servant à codifier les réseaux et à définir les coûts des liens, tout en estimant parfois certains des paramètres des modèles d'affectation (certains modèles n'ont toutefois aucun paramètre de calage, comme l'affectation au plus court chemin par exemple). Mais les données pour caler ces attributs ne sont pas toujours directement disponibles. Par exemple, pour le calage de relations débit-coût, on ne dispose pas toujours de données sur les coûts ou les temps en fonction de la charge de la voie. La relation ne sera pas calée directement, mais indirectement. De ce fait, il n'y a pas de procédure standard de calage de l'étape d'affectation (Ortuzar, Willumsen, 1994) comme nous avons pu en présenter dans les autres chapitres.

Avant de présenter les indicateurs habituellement utilisés, nous rappelons les données disponibles (ou plus exactement qui sont parfois disponibles), tant pour définir que pour caler les réseaux ou le modèle d'affectation :

- les matrices de déplacements, qu'elles soient directement issues d'une production de données ou estimées à partir d'un modèle. Nous avons déjà eu l'occasion de souligner l'incertitude souvent assez forte entourant cette information ;
- les données de comptages qui indiquent le nombre de véhicules circulant en un point donné d'une infrastructure. Selon les boucles de comptage, il est parfois possible de connaître également la composition de ce trafic (part de poids lourds), ainsi que la vitesse des véhicules. Au carrefour, des boucles judicieusement placées permettent de connaître les mouvements tournants. En France, des comptages permanents sont effectués par les DDE et par certaines collectivités. Les systèmes de régulation des feux peuvent également fournir de l'information. Cette information n'est toutefois pas toujours facile à interpréter dans la mesure où une partie du trafic qui circule sur les infrastructures ne rentre pas (ou plus exactement pas toujours) dans le champ de la modélisation (trafic de poids lourds, livraison, déplacements intrazone, déplacements des non résidents effectués à l'intérieur du périmètre d'enquête... (chapitre 4, p. 119) ;
- les longueurs et caractéristiques géométriques des voies. De plus en plus de bases de données sont disponibles à un niveau national ou local ;
- les données d'exploitation des exploitants de réseau qui permettent de connaître les longueurs, capacité, fréquence, vitesse... des différents services de transports collectifs. En général, il s'agit d'une information très fiable ... dans la mesure où l'exploitant accepte de la fournir ! ;
- les enquêtes de temps de parcours. Elles consistent généralement à faire circuler des véhicules dans les conditions réelles de circulation et à calculer les temps de parcours. Parfois, ce type d'informations peut être obtenu auprès de certaines sociétés faisant circuler des flottes importantes de véhicules (compagnies de taxi, de livraison...).

La qualité et la portée de l'information est évidemment très inégale selon la source d'information. La précision des matrices origines-destinations est souvent assez médiocre, tout au moins pour des zonages assez fin (à l'exception de certaines matrices produites par des exploitants qui

peuvent avoir une connaissance très fine de leur marché), alors que d'autres données peuvent être connues avec beaucoup plus de précision comme certains comptages (cette précision n'est cependant pas systématique et l'information est de portée beaucoup plus limitée). La qualité des données doit donc être prise en compte lors de l'étape de codification, puis de calage du réseau.

S'il n'y a pas de procédure standard de calage de l'étape d'affectation, cela ne veut pas dire qu'il n'y a pas de recommandations, ni d'indicateurs pouvant être utilisés. Nous évoquons les principaux aspects à prendre en compte lors du calage de l'affectation.

Le principal élément pour apprécier la qualité du calage de l'affectation consiste à s'assurer que les flux calculés par le modèle sont effectivement observés à travers les données collectées. Il est généralement recommandé de produire la statistique suivante :

$$\sum_a \left\{ \frac{(\text{flux observés}_a - \text{flux calculés}_a)^2}{\text{flux observés}_a} \right\} \quad (10.38)$$

Plus cette statistique est faible, meilleure est la reproduction des flux observés. L'analyse est toutefois délicate, car les données disponibles ne sont pas toujours très fiables, pas forcément relatives à la même année ou la même période de l'année, ou ne couvrent pas toute l'aire d'étude mais une partie seulement ou un certain type de réseau.... De plus, elles couvrent des champs de mobilité qui peuvent être différents si les flux observés correspondent à des comptages (les comptages prennent en compte tous les flux quels qu'ils soient, alors que la modélisation peut ne porter que sur les déplacements de personnes). L'amélioration de la statistique doit être analysée avec prudence car elle n'est pas automatiquement signe d'un gain équivalent de qualité du modèle. La statistique peut être estimée pour l'ensemble du réseau, mais il est souvent intéressant de l'estimer pour des segments de liens uniquement (les axes directement concernés par l'objet d'étude, les axes structurants du réseau, un secteur de l'agglomération, les infrastructures coupant un cordon, les axes avec carrefours à feux versus les axes sans feux de signalisation, par classe de courbe débit-coût...). La définition des segments pertinents doit permettre d'indiquer des pistes pour améliorer le calage (l'exemple axes avec carrefours à feux versus axes sans feux peut indiquer une mauvaise prise en compte des carrefours à feux ou bien de ceux sans feux...).

Au-delà de cet indicateur synthétique, l'analyse peut être poursuivie de manière plus désagrégée pour étudier la charge d'une infrastructure routière ou d'une ligne de transports collectifs. On peut également étudier les mouvements tournants ou les correspondances ou le serpent de charge pour une ligne de transports collectifs... La finesse de l'analyse résultera comme toujours d'un trade-off entre les objectifs de la modélisation et le budget disponible, réalisé sous la contrainte des données disponibles et de leur qualité.

La reproduction des charges n'est pas le seul indicateur disponible. Il est également possible de calculer les temps de trajet et les temps ou coûts généralisés. Les premiers sont parfois disponibles (temps déclarés dans les enquêtes, enquêtes avec véhicules dans la circulation, sources d'exploitants...). Mais de nouveau l'analyse doit être menée avec précaution. Un temps déclaré peut être biaisé. L'expérience indique que le temps déclaré en transports collectifs est souvent majoré par rapport au temps réel et inversement pour la voiture (sauf en cas de congestion). De même, les vitesses produites par le modèle sur certains axes peuvent être comparées avec les vitesses disponibles.

Ces indicateurs sont produits à l'issue de l'affectation de la matrice de demande et correspondent à la finalité principale du modèle à savoir reproduire puis prévoir les flux et les temps de parcours pour pouvoir évaluer différents scénarios de transports. Mais il y a une autre série de diagnostics qui doit être menée avant même de procéder à l'affectation de la matrice de déplacement. La première étape consiste à valider certaines fonctionnalités du réseau. Pour ce faire, il est

souhaitable d'analyser les arbres de plus courts chemins. Cette analyse permettra de détecter certaines erreurs « grossières » (et parfois d'autres moins importantes) sur la définition du réseau : un nœud qui ne peut être atteint car un lien a été oublié, des arbres de plus courts chemins non intuitifs qui peuvent désigner des sur/sous-coûts sur certains liens ou une correspondance impossible.... L'analyse des temps à vide sur les plus courts chemins est également très instructive avec par exemple la production de courbes isochrones. Ces courbes peuvent permettre d'isoler des zones anormalement « proches » ou « lointaines ». Il est recommandé de mener ces analyses sur la distance seule, sur le temps réel seul, sur le temps généralisé (sans prendre en compte le coût), sur le coût seul si nécessaire, et enfin sur le coût généralisé. Ces différentes analyses facilitent la détection d'erreurs sur certains paramètres dont l'effet pourrait être « noyé » sur le coût généralisé global, mais qui peut être « amplifié » sur une de ses composantes prise isolément.

Après avoir suggéré quelques pistes et indicateurs pour valider un réseau, nous revenons rapidement sur les attributs et fonctions devant être examinés. A priori tous ! ... sinon, c'est que l'attribut ou la fonction est inutile ! Toutefois, la masse d'informations devant être validée est rapidement considérable (pour un zonage comportant quelques centaines de zones, on peut facilement avoir entre 2 et 5 fois plus de nœuds et plusieurs milliers de liens, chacun de ces liens comportant plusieurs attributs) ce qui rend l'opération complexe. De ce fait, il est recommandé de commencer par les éléments les plus globaux du réseau pour continuer avec ceux qui concernent des groupes de liens puis des liens spécifiques. L'ordre suivant peut être utile :

- validation des fonctions de coûts généralisés ou de tout autre fonction définie au niveau de l'ensemble du réseau (par exemple si le calcul du temps d'accès résulte d'une fonction appliquée à l'ensemble des zones) et des paramètres de ces fonctions. Le choix des paramètres de ces fonctions influe sur l'ensemble des plus courts chemins et des flux affectés (par exemple vitesse de marche à pied, coefficient de pondération de l'attente, temps de recherche du stationnement...);
- validation de fonctions et de leur paramètres influant sur des segments de liens, par exemple paramètre d'une fonction débit-coût pour une classe de voies, ou paramètres appliqués à un groupe de lignes de transports collectifs ;
- validation par lignes de transports collectifs (sur la fréquence par exemple) ou par axes routiers structurants (sur le traitement des carrefours par exemple) ;
- validation plus désagrégée (paramètre d'un lien particulier, analyse des correspondances à un nœud donné...).

Cet ordre doit bien évidemment être adapté au contexte d'étude, au budget et aux données disponibles. Il doit de plus être combiné avec la séquence évoquée précédemment : validation des données à l'issue du calcul des plus courts chemins, puis validation à l'issue de l'affectation de la demande.

La « correction » du réseau doit toutefois être menée avec discernement. Il ne s'agit pas de modifier le réseau pour assurer une reproduction parfaite des données de trafic. Cette reproduction parfaite n'a aucun intérêt si elle est le résultat d'un ajustement « aveugle » des coefficients pour viser la nullité de l'indicateur de l'équation 10.38 par exemple. La « correction » des attributs ne doit se faire que s'il y a effectivement une « erreur » et être menée de manière cohérente sur l'ensemble du réseau. Sinon, le modèle sera capable de reproduire la réalité, mais risque d'être moins performant quant à la simulation pour laquelle il est développé.

Ceci nous amène à la dernière étape après avoir calé un modèle, il est nécessaire de le valider, c'est-à-dire de le faire tourner en modifiant certains paramètres pour analyser la plausibilité des résultats produits. Une démarche similaire à celle évoquée ci-dessus pourra être développée à cet effet en commençant par analyser les résultats les plus agrégés puis en affinant progressivement.

Le temps nécessaire à la validation d'un réseau et à la validation du modèle ne doit pas être mésestimé. Certains logiciels offrent des fonctionnalités permettant de faciliter le travail (par exemple, la vérification de la connexité du réseau ou certaines erreurs « grossières »), mais l'empirisme reste de rigueur au delà des recommandations que nous avons pu formuler.

6.2. Zonage

La question a déjà été largement évoquée. Toutefois, son importance nous oblige à rappeler les enjeux. La précision du zonage conditionne directement la précision de la définition des réseaux. Elle conditionne également la précision de l'estimation des temps d'accès. Même avec un modèle désagrégé, le calcul du temps d'accès reste tout autant agrégé que pour un modèle agrégé. Ainsi même avec un découpage fin, la reproduction des itinéraires est très mauvaise au niveau individuel, même si à un niveau agrégé les résultats peuvent être acceptables (chapitres 9, p. 299 et 4, p. 127). Cette définition peut donc influencer sur l'estimation des coûts généralisés, et donc sur la répartition modale.

Enfin, le zonage influe sur l'importance du trafic intrazone qui occupe une partie de l'espace du réseau modélisé. L'estimation étant assez délicate et dépendante du niveau de congestion, sa prise en compte est généralement assez difficile.

6.3. Définition des réseaux

Ce point a également été largement abordé à la section 3 (p. 312). Nous le rappelons à travers quelques exemples. Il est évident que la précision de la codification influera sur la capacité à reproduire finement les itinéraires utilisés. Si un carrefour n'est pas bien représenté au niveau de ses mouvements tournants, on risque de sous/sur-estimer des problèmes de congestion qui peuvent conduire à une sur/sous-estimation du trafic sur un lien, dont les conséquences peuvent être amplifiées par le choix de méthodes ne prenant pas en compte les contraintes de capacité.

Selon la taille des zones, il y a généralement une partie du réseau considéré comme étant de desserte fine au niveau de la zone qui n'est pas pris en compte dans la codification. Cette approximation est généralement acceptable pour un réseau non congestionné. En revanche, lorsque la congestion s'accroît, une partie des conducteurs a tendance à rechercher des voies d'évitement quitte à emprunter des voies locales. La non-prise en compte de ce réseau local risque de conduire à une congestion accrue par rapport à la réalité. Inversement, la prise en compte de ce réseau introduit probablement une capacité excessive en situation peu congestionnée, car les déplacements intrazonaux utilisant ces infrastructures ne sont pas pris en compte dans la modélisation.

De plus, la définition du réseau interagit avec le choix de la méthode d'affectation (par exemple l'affectation multichemin de Dial, schéma 10.7, p. 345).

Comme pour le zonage, il n'est pas possible de recommander « La » bonne solution. Il s'agit toujours d'un compromis entre les objectifs et les moyens disponibles sous contraintes de données. Ceci étant il faut être conscient des biais introduits.

6.4. Hypothèses des modèles d'affectation

Deux hypothèses principales ont été formulées. Elles conduisent à la classification des modèles d'affectation :

- prise en compte ou non des contraintes de capacité à travers une fonction reliant le coût d'un lien à sa charge ;
- perception déterministe ou stochastique des coûts et donc approche déterministe ou probabiliste de l'utilité.

Les deux hypothèses ont été longuement discutées tout au long de la section 5, nous ne faisons donc que rappeler leur importance. De plus, l'abandon de ces hypothèses conduit à en formuler d'autres, même si leur incidence est plus limitée. La prise en compte de la charge sur le coût du lien est généralement limitée à la charge du lien. Il est pourtant évident que la charge sur d'autres liens peut influencer sur le coût d'un lien, notamment par l'intermédiaire de la propagation de la congestion au niveau du carrefour (section 5.4.1.1, p. 343). La prise en compte de la perception des coûts nécessite de faire une hypothèse sur la distribution du terme d'erreur et généralement de postuler une indépendance entre les distributions relatives à chaque itinéraire au niveau individuel...

Au-delà de ces deux hypothèses, les différents modèles s'appuient sur une hypothèse de rationalité du choix du consommateur qui maximise son utilité. Les limites en sont connues. Nous ne ferons qu'en rappeler une : la connaissance de l'information. Pour maximiser son utilité, l'individu choisit l'itinéraire le plus court au sens du coût généralisé. Implicitement, il dispose de toute l'information lui permettant de calculer le coût généralisé de tous les itinéraires concurrents potentiels. L'évidence démontre le contraire. C'est notamment pour y faire face que l'utilité est souvent définie avec un terme d'erreur qui est censé intégrer notamment cet aspect.

6.5. L'équilibre dans l'affectation

La notion d'équilibre n'est pas aussi évidente qu'il y paraît (chapitre 5, p. 159). Elle suppose une situation stabilisée. L'expérience de la circulation quotidienne amène à en douter tant la variabilité est forte d'un jour à l'autre. On considère donc que les données recueillies ne sont pas représentatives de la situation qui prévaut chaque jour, mais plutôt d'une situation moyenne. Mais même à ce niveau l'équilibre n'est pas évident, car les comportements évoluent sans cesse en réagissant non pas continûment aux changements de l'état du sous-système des transports, mais avec un retard plus ou moins important. Goodwin (1998) montre ainsi que les réponses à des modifications de l'offre peuvent apparaître quelques années, voire quelques dizaines d'années plus tard.

Certes les réponses sont probablement plus rapides en matière de changement d'itinéraires qu'en matière de changement modal ou de localisation, mais elles ne sont pas immédiates et peuvent prendre un certain temps pour se stabiliser. Prenons un exemple. La plupart des individus sont habitués à une certaine régularité (habitude) dans leur comportement. Cette habitude pourra être rompue si la congestion devient trop pénible ou si l'individu connaît un incident l'obligeant à changer d'itinéraire. Si l'itinéraire concurrent qui a été utilisé à cette occasion est suffisamment attractif, l'automobiliste pourra être conduit à l'essayer de nouveau hors incident et peut être à changer de manière définitive après plusieurs tentatives confirmant à ses yeux la supériorité du nouvel itinéraire. Ce processus est plus ou moins long selon les individus et peut prendre plusieurs mois ou années avant que l'individu n'expérimente l'itinéraire. Ce que l'on mesure n'est donc peut être pas une situation en équilibre.

6.6. Convergence des modèles

La convergence des procédures itératives utilisées par les modèles d'équilibre (de Wardrop ou avec utilité stochastique) nécessite de poser certaines hypothèses. En particulier pour l'équilibre de Wardrop, il est nécessaire d'avoir une fonction coût-débit croissante. C'est généralement le cas de la plupart des fonctions utilisées. Mais l'unicité de la solution n'est garantie que dans la mesure où la courbe est strictement croissante (section 5.4.1.1, p. 343).

Dans ce cas, l'hypothèse n'est pas restrictive. En revanche, pour les modèles prenant en compte une relation entre le coût du lien et la charge non seulement du lien, mais aussi d'autres liens, les conditions pour que la convergence soit garantie sont plus restrictives (Thomas, 1991).

Dans le cas d'une affectation selon le second principe de Wardrop, une condition supplémentaire est posée sur la dérivée de la courbe de coût qui doit être positive (et même strictement positive pour l'unicité de la solution). La condition est ici plus restrictive que pour le premier principe de Wardrop, puisque deux des courbes exposées dans ce chapitre ne la respectent pas. L'incidence reste toutefois limitée, puisque l'affectation selon le second principe n'est pas réaliste sur le plan comportemental, même si elle peut être intéressante pour analyser le réseau (section 5.4.1.5, p. 349).

Enfin, il n'est pas inutile de rappeler que l'affectation incrémentale ou par tranche ne converge pas forcément vers une solution respectant le premier principe de Wardrop (section 5.4.2, p. 352).

6.7. Prévision

L'utilisation d'un modèle en prévision pose toujours le problème de la stabilité ou de la loi d'évolution des paramètres de calage en fonction du temps. Ce problème a été déjà été évoqué pour chacune des étapes. Mais comme pour le calage, le problème se pose d'une manière un peu spécifique pour l'affectation. La question de la stabilité ou de l'évolution ne se pose pas que pour les paramètres de calage des modèles, mais également pour la définition des réseaux.

Il est nécessaire de s'interroger sur la stabilité des paramètres de définition des réseaux. Certes, la longueur d'un lien a peu de chances d'évoluer (sauf si le lien est le support d'une nouvelle infrastructure de transport), en revanche, d'autres paramètres peuvent évoluer. Quelques exemples permettent de situer le problème. A mesure que la congestion s'accroît, il n'est pas exclu que les comportements des automobilistes évoluent. Cette évolution peut avoir des conséquences sur la définition de la capacité des infrastructures. Par exemple, il est connu qu'à Paris, la capacité d'une voie rapide est supérieure de l'ordre de 10 % à la capacité d'une voie similaire en province du fait d'une interdistance entre véhicules plus faible. Le développement de la congestion en province réduira-t-il les interdistances ? On peut également imaginer une incidence sur les courbes débit-coût.

Le développement des systèmes de régulation avec l'information dynamique embarquée ou non modifie les stratégies d'itinéraires des conducteurs. A un horizon de 10 ans, comment faut-il prendre en compte cette incidence ? à travers de nouvelles lois de comportement, à travers la modification de la description des réseaux, à travers la modification des hypothèses sur la perception des coûts... ?

La prévision concerne également la modification des réseaux pour simuler l'incidence de nouvelles infrastructures de transports ou de nouvelles politiques de déplacements. Il est important que les nouveaux réseaux soient définis avec le même niveau de précision que lors de la codification du réseau de référence. Par exemple, un soin tout particulier doit être apporté à la définition des temps d'accès si ceux-ci doivent être modifiés au niveau d'une zone. Des principes de codification qui seraient trop différents entre les situations de référence et de prévision pourraient par exemple avoir une incidence très forte sur la charge d'un nouvel axe de transports collectifs. L'exemple pourrait être repris avec le traitement des correspondances....

Compte tenu de l'ensemble de ces limites, nous serions tenté de reprendre à notre compte le propos de Thomas (1991) selon lequel la sophistication des méthodes n'est peut être pas toujours nécessaire compte tenu de l'incertitude concernant la définition des réseaux, les matrices origines-destinations et le comportement des usagers.

Annexe 1 - Données nécessaires au calage du modèle

« Prix-temps » et corrigé

Ces données correspondent à un cas d'école. Elles ont été fournies par la SNCF et Air France pour des besoins pédagogiques. Elles ne reflètent toutefois qu'imparfaitement la réalité pour des raisons de confidentialité de certaines données. Elles restent malgré tout globalement cohérentes, permettant la réalisation de l'exercice pédagogique.

1. TRAFICS EN 1996

entre Paris et... en milliers de passagers	SNCF 1996			AERIEN 1996		
	1ère classe	2ème classe	total	Air France entre Orly et...	Air France entre CDG et...	Total y compris autres compagnies
Bordeaux	413	1592	2005	860	274	1400
Brest	67	404	471	376	51	428
Clermont- Ferrand	76	353	429	227	17	243
Limoges	116	359	475	63	0	70
Lyon	2024	3345	5369	410	300	710
Marseille	253	1069	1322	1406	375	2700
Metz	156	384	540			
Metz/Nancy				0	0	75
Montpellier	144	584	728	707	165	1150
Nancy	252	454	706			
Nantes	583	1712	2295	168	125	350
Nice	105	351	456	1273	601	2900
Strasbourg	99	542	641	795	198	1250
Toulon	139	490	629	461	0	650
Toulouse	115	634	749	1272	358	2650

Tableau annexe 1.1 : Trafic ferroviaire et aérien en 1996

Les trafics sont exprimés en milliers de voyageurs aller + retour. CDG = Charles de Gaulle. Pour l'aérien le trafic total comprend les deux aéroports et toutes les compagnies assurant la liaison.

2. TEMPS DE PARCOURS EN MINUTES EN 1996

entre Paris et	S N C F		AIR FRANCE (ORLY et...)	
	trajet	temps terminal	trajet	temps terminal
Bordeaux	177	65	60	105
Brest	243	65	65	105
Clermont-Ferrand	201	65	50	90
Limoges	170	65	70	110
Lyon	120	65	55	115
Marseille	254	65	70	120
Metz/Nancy			50	90
Metz	163	65		
Montpellier	254	65	70	115
Nancy	157	65		
Nantes	119	65	50	125
Nice	386	65	80	130
Strasbourg	231	65	55	105
Toulon	303	65	75	110
Toulouse	300	65	70	120

Tableau annexe 1.2 : Temps de parcours ferroviaire et aérien en 1996

Le temps terminal = temps d'accès à l'aéroport (ou à la gare + temps d'accès à la destination finale à partir de l'aéroport ou de la gare).

3. TARIFS DE BASE AUTOMNE-HIVER 1996 (CONVERTI EN EQUIVALENT EUROS 96)

entre Paris et...	Tarif moyen (moyenne sur les résa)		Air France tarif moyen entre Orly et...
	1ère	2ème	
Bordeaux	62,20	44,71	82,61
Brest	72,57	48,07	98,60
Clermont-Ferrand	44,96	30,06	89,44
Limoges	43,15	28,77	98,02
Lyon	64,27	43,41	90,02
Marseille	72,90	50,77	89,10
Metz/Nancy	39,39	26,31	125,83
Montpellier	72,76	50,67	92,69
Nantes	57,53	38,36	90,37
Nice	84,68	60,18	101,96
Strasbourg	51,57	34,34	78,09
Toulon	77,66	54,15	105,78
Toulouse	80,21	56,76	87,36

Tableau annexe 1.3 : Tarifs ferroviaire et aérien en 1996

Pour le calcul du tarif moyen SNCF 1ère + 2ème classe, on prendra la moyenne des deux classes au prorata des trafics fournis dans le tableau annexe 1.1.

Pour le trafic aérien, compte tenu de la structure tarifaire, le tarif moyen effectivement payé par l'utilisateur est obtenu en appliquant un coefficient de 76 % du plein tarif. Pour la SNCF, ce pourcentage est de 85 % du plein tarif (obtenu en faisant la moyenne du plein tarif pour chaque niveau de

réservation pondérée par le poids du trafic de chaque niveau de réservation). C'est ce tarif qui figure dans le tableau ci-dessus (tableau annexe 1.3).

4. CORRIGE DE L'EXERCICE DE CALAGE DU MODELE PRIX-TEMPS - ENSEMBLE DES DONNEES

Calcul des tarifs moyens en 1996 en €96

	tarif moyen SNCF			tarif
	1e cl	2e cl	1/2e cl	aérien
Bordeaux	62,20	44,71	48,31	82,61
Brest	72,57	48,07	51,56	98,60
Clermont-Ferrand	44,96	30,06	32,70	89,44
Limoges	43,15	28,77	32,28	98,02
Lyon	64,27	43,41	51,27	90,02
Marseille	72,90	50,77	55,01	89,10
Metz/Nancy	39,39	26,31	30,59	125,83
Montpellier	72,76	50,67	55,04	92,69
Nantes	57,53	38,36	43,23	90,37
Nice	84,68	60,18	65,82	101,96
Strasbourg	51,57	34,34	37,00	78,09
Toulon	77,66	54,15	59,35	105,78
Toulouse	80,21	56,76	60,36	87,36

Valeur d'indifférence du temps en 1996 €96

SNCF	AIR	h_i	$\ln(h_i)$
en mn	en mn	heures	heures
242	165	26,73	3,29
308	170	20,45	3,02
266	140	27,02	3,30
235	180	71,72	4,27
185	170	155,00	5,04
319	190	15,86	2,76
225	140	67,22	4,21
319	185	16,86	2,82
184	175	314,29	5,75
451	210	9,00	2,20
296	160	18,13	2,90
368	185	15,22	2,72
365	190	9,26	2,23

Calcul des parts de marché

	trafic en milliers en 1996						u_i
	SNCF	SNCF	SNCF	Aérien	%fer	%air	transformée
	1ère cl	2ème cl	1+2e cl				air
Bordeaux	413	1592	2005	1400	0,589	0,411	-0,225
Brest	67	404	471	428	0,524	0,476	-0,060
Clermont-Ferrand	76	353	429	243	0,638	0,362	-0,354
Limoges	116	359	475	70	0,872	0,128	-1,134
Lyon	2024	3345	5369	710	0,883	0,117	-1,191
Marseille	253	1069	1322	2700	0,329	0,671	0,444
Metz/Nancy	408	838	1246	75	0,943	0,057	-1,582
Montpellier	144	584	728	1150	0,388	0,612	0,285
Nantes	583	1712	2295	350	0,868	0,132	-1,115
Nice	105	351	456	2900	0,136	0,864	1,099
Strasbourg	99	542	641	1250	0,339	0,661	0,415
Toulon	139	490	629	650	0,492	0,508	0,021
Toulouse	115	634	749	2650	0,220	0,780	0,771

Il faut donc réaliser la régression entre le vecteur u_i (transformée de Laplace-Gauss de la part de marché de l'avion) et le vecteur du logarithme népérien de la valeur d'indifférence du temps

Résultats de la régression

constante	2,132
pente	-0,682
écart-type log valeurs du temps	1,467
m moyenne log valeurs du temps	3,127
médiane valeur du temps	22,81€/h

Comparaison modèle théorique, données d'enquête (avec valeur du temps en heure)

	Modèle théorique		Données d'enquête		Ecart différence en %		
	u_i transformée air	%air	u_i transformée air	%air	absolue	relative /air	/fer
Bordeaux	-0,108063	45,7	-0,225	41,1	4,581	11,14	7,78
Brest	0,0744038	53	-0,060	47,6	5,357	11,25	10,23
Clermont-Ferrand	-0,11549	45,4	-0,354	36,2	9,242	25,56	14,48
Limoges	-0,781046	21,7	-1,134	12,8	8,895	69,25	10,21
Lyon	-1,306545	9,6	-1,191	11,7	2,111	18,08	2,39
Marseille	0,2479256	59,8	0,444	67,1	7,34	10,93	22,33
Metz/Nancy	-0,736956	23,1	-1,582	5,7	17,38	306,1	18,43
Montpellier	0,2061158	58,2	0,285	61,2	3,07	5,014	7,921
Nantes	-1,788541	3,7	-1,115	13,2	9,548	72,16	11,00
Nice	0,6343178	73,7	1,099	86,4	12,71	14,7	93,51
Strasbourg	0,1566537	56,2	0,415	66,1	9,878	14,94	29,14
Toulon	0,2756613	60,9	0,021	50,8	10,04	19,75	20,41
Toulouse	0,6148427	73,1	0,771	78,0	4,897	6,281	22,22

Calage du modèle prix-temps ensemble des villes

Statistiques de la régression	
Coefficient de détermination multiple R	0,88500425
Coefficient de détermination R^2	0,78323253
Coefficient de détermination corrigé R^2	0,7635264
Ecart-type	0,40575968
Observations	13

Le coefficient de détermination R^2 correspond au pourcentage de variance expliquée par la régression. Plus il est proche de 1, meilleure est la régression (car la variance résiduelle est très faible). Lorsqu'il se rapproche de 0, la régression n'explique plus rien. Le R^2 est obtenu par la formule suivante : $1 - \frac{\text{somme des carrés des résidus}}{\text{somme totale des carrés}}$ ou encore $\frac{\text{somme des carrés de la régression (expliquée)}}{\text{somme totale des carrés}}$. Le coefficient de détermination multiple R est égal à la racine de R^2 . L'écart-type = racine carré (moyenne des carrés des résidus) / nombre de degrés de liberté des résidus.

Le coefficient de détermination corrigé R^2 permet de comparer des R^2 issus de plusieurs régressions en tenant compte du nombre de variables explicatives (équations 6.19 ou 6.20, p. 181).

ANALYSE DE VARIANCE					
	Degré de liberté	Somme des carrés	Moyenne des carrés	F	Valeur critique de F
Régression	1	6,54375559	6,54375559	39,7456216	5,8016E-05
Résidus	11	1,81105009	0,16464092		
Total	12	8,35480568			

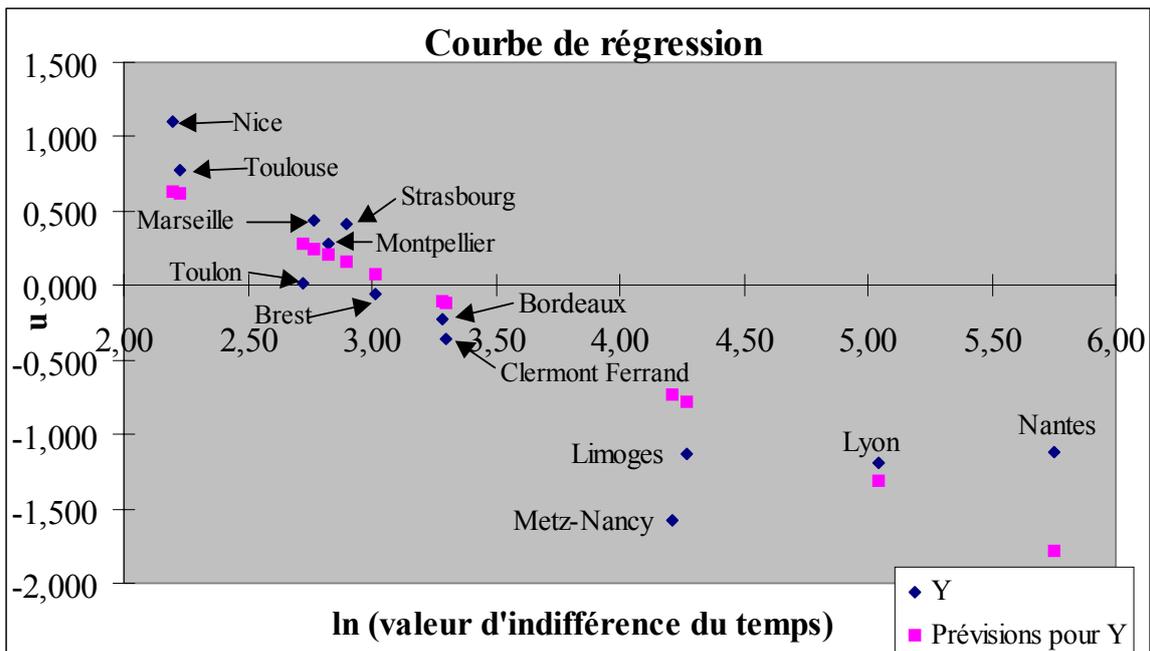
La statistique « F » de Fisher permet de tester l'hypothèse H_0 « tous les coefficients sont nuls ». Si le F observé est supérieur au F théorique (à 1 et 11 degrés de liberté) lu dans les tables de Fisher, c'est que l'un au moins des coefficients de la régression est non nul au seuil de confiance choisi (ici 5 %. Dans la pratique, on peut se contenter de lire la valeur critique du F qui donne la probabilité d'erreur lorsque l'on rejette l'hypothèse nulle, soit 5,8E-05 dans l'exemple ci-dessus). On a une seule variable explicative X_1 , le coefficient X_1 est donc significativement différent de 0. Le nombre de degrés de liberté de F = (nombre de variables explicatives - 1; nombre d'observations - nombre de variables explicatives). La valeur critique de F correspond au niveau de probabilité en dessous duquel il n'est pas possible de rejeter l'hypothèse H_0 . Cette probabilité étant inférieure à 5 %, l'hypothèse H_0 peut être rejetée au seuil de confiance de 5 %. La somme des carrés des résidus (différence entre prévision Y et valeur observée de Y) est obtenue à partir du tableau d'analyse des résidus. La moyenne des carrés des résidus = somme des carrés des résidus / nombre de degrés de liberté des résidus (et non pas nombre d'observations).

	Coefficients	Ecart-type	Statistique t	Probabilité	Limite inférieure pour seuil de confiance = 95 %	Limite supérieure
Constante	2,1322	0,3870	5,5098	1,835E-04	1,2805	2,9840
Variable X_1	-0,6818	0,1082	-6,3044	5,802E-05	-0,9199	-0,4438

La statistique t de student permet de tester l'hypothèse H_0 « le coefficient est nul ». Si le t de Student observé est supérieur au t théorique (11 degrés de liberté) lu dans les tables de student, c'est que le coefficient de la régression est non nul au seuil de confiance choisi (ici 5 %. Dans la pratique, on peut se contenter de lire la valeur de la probabilité qui donne la probabilité d'erreur lorsque l'on rejette l'hypothèse nulle, soit 5,8E-05 dans l'exemple ci-dessus pour le coefficient X_1). Le nombre de degrés de liberté = nombre d'observations - (nombre de variables explicatives + constante). La probabilité correspond au niveau de probabilité ne permettant pas de rejeter l'hypothèse H_0 . Cette probabilité étant inférieure à 5 %, l'hypothèse H_0 peut être rejetée au seuil de confiance de 5 %. Le demi-intervalle de confiance est obtenu en multipliant l'écart-type par la valeur du t de Student (au seuil de confiance choisi et pour le nombre de degrés de liberté des résidus). Les limites sont obtenues en ajoutant/retranchant le demi-intervalle de confiance à la valeur du coefficient.

ANALYSE DES RESIDUS			
Observation	Prévisions pour u (transformée de la part de marché de l'air)	Résidus	Résidus normalisés
Bordeaux	-0,1081	-0,1165	-0,2999
Brest	0,0744	-0,1344	-0,3459
Clermont-Ferrand	-0,1155	-0,2387	-0,6144
Limoges	-0,7810	-0,3527	-0,9080
Lyon	-1,3065	0,1154	0,2970
Marseille	0,2479	0,1956	0,5035
Metz/Nancy	-0,7370	-0,8455	-2,1763
Montpellier	0,2061	0,0793	0,2042
Nantes	-1,7885	0,6731	1,7326
Nice	0,6343	0,4647	1,1962
Strasbourg	0,1567	0,2586	0,6657
Toulon	0,2757	-0,2551	-0,6566
Toulouse	0,6148	0,1561	0,4019

Les résidus correspondent à la différence entre la valeur théorique de u (c'est-à-dire la valeur calculée par le modèle calé) et la valeur observée (c'est-à-dire celle résultant des données de base). Le résidu normalisé est obtenu en divisant le résidu par l'écart-type de la régression. L'analyse des résidus normalisés permet d'identifier les observations qui contribuent le plus à l'écart-type donc à la partie non expliquée de la variance. L'analyse des résidus informe sur l'écart absolu pour chacune des observations entre la prévision et l'observé, mais si on veut analyser la qualité de la prévision pour chacune des observations, l'écart relatif sera généralement plus pertinent.



Graphique annexe 1.1 : Courbe de régression - modèle prix-temps avec l'ensemble des données

Bibliographie

- Abraham C, Baumgart T, Blanchet JD (1969), Un modèle de prévision de trafic aérien intérieur, Direction des transports aériens, *Revue du Secrétariat Général de l'Aviation Civile*, N°136, pp. 15-28.
- Abraham C, Coquand R (1961), La répartition du trafic entre itinéraires concurrents, réflexions sur le comportement des usagers – application au calcul des péages, *Revue Générale des Routes et Aérodromes*, N°357, pp. 57-76.
- Ambrosini C, Patier-Marque D, Routhier J-L (1999a) *Transports de marchandises en ville - Résultats de l'enquête quantitative réalisée à Dijon*, Rapport de recherche pour la DRAST-DTT et l'ADEME, Laboratoire d'Economie des Transports, Lyon, 120p.
- Ambrosini C, Patier-Marque D, Routhier J-L (1999b), *Transports de marchandises en ville - Résultats de l'enquête quantitative réalisée à Marseille*, Rapport de recherche pour la DRAST-DTT et l'ADEME, Laboratoire d'Economie des Transports, Lyon, 120p.
- Ampt ES (1997), Response Rates - Do they matter ? in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 115-125.
- Ampt ES, Ortuzar JD, Willumsen LG (1998), Metropolitan Origin-Destination Surveys: The State of the Art, *8th World Conference on Transport Research*, Antwerp, 12-17 July 1998, 17p.
- Ampt ES, Richardson AJ & Brög W (1985) *New survey methods in transport*. VNU Science Press, Utrecht, The Netherlands.
- Ampt ES, Swanson J (1996), Pictorial presentation of SP: Does it work?, *4th International Conference on Survey Methods in Transport*, Steeple Aston, Oxford, UK, September 8-11, 1996, Proceedings vol. 2, page purple.
- Andan O, Bonnel P, Raux C (1988), *Les analyses des comportements de mobilité individuelle quotidienne, une synthèse bibliographique*, Laboratoire d'Economie des Transports, pour le compte du SERT, Ministère des Transports, Lyon, 118p.
- Andan O, Raux C (1997), L'univers de choix et les leviers de changement des comportements : L'apport d'une méthode interactive de simulation, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin, Raux R (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 417-435.
- Andan O, Tabourin E (1998), Croissance et recomposition socio-économiques des espaces urbains, in *Données Urbaines 2*, coordonné par Pumain D et Mattei M-F (éd.), collection VILLE, Anthropos, pp. 273-285.
- Ardilly P (1994), *Les techniques de sondage*, Editions Technip, Paris,
- Arduin JP, Ni J, Picq O (1994), valeur du temps, log-normalité des revenus et choix modal, *Communication à la WC RR*, SNCF, Paris.

- Armoogum J (1999), *Correction de la non-réponse et de certaines erreurs de mesures dans une enquête par sondage : application à l'enquête transports et communications 1993-94*, Thèse de doctorat Université Libre de Bruxelles, Bruxelles, 191p. + annexes
- Armoogum J, Han X-L, Madre J-L, Polak J (1997), *Improved sampling schemes for travel diaries*, deliverables D7, projet MEST, Arcueil, 22p.
- Armoogum J, Madre J-L (1997), Correction de la non-réponse dans l'enquête transports 1993-94 : l'exemple des déplacements quotidiens, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 127-143.
- Armoogum J, Coindet J-P, Madre J-L, Massot M-H, Papon F (2001), *Suivi de la mobilité en Ile-de-France : test d'une enquête longitudinale*, rapport sur convention ADEME, INRETS, Arcueil, 28p.
- Averous B (1971), *Analyse des tests dans les études de transport à moyen et long terme*, IRT, Arcueil.
- Averous B, Bieber A, Bourgin C, Currat C, Godard X, Matalon B, Orfeuill J-P, Tarrus A (1979-80), *Recherche sur la mobilité des personnes en zone urbaine ; exploitation de l'enquête de Dijon*, Rapport de recherche N°38, parus en 4 tomes, IRT, Arcueil.
- Axhausen KW (1995), *Travel diaries: an annotated catalogue*, working paper, Centre for Transport Studies, London, 128p.
- Axhausen KW (1998), Can we ever obtain the data we would like to have ?, in Gärling T, Laitila T, Westin K (eds.) (1998), *Theoretical foundations of travel choice modelling*, Pergamon, Oxford, pp. 305-323.
- Balavoine E (1981), *La prévision de la demande de transports urbains*, thèse d'Etat en Sciences Economiques, Université de Bordeaux I, Bordeaux, 591p.
- Banister D (1993), Structures démographiques et comportements sociaux, rapport introductif au 12^{ème} symposium de la CEMT Lisboa 1992 *La croissance des transports en question*, CEMT, Paris, pp. 111-156.
- Banister D (1998), *Transport policy and the environment*, E&FN SPON, Londres, 348p.
- Banister D, Bayliss D (1992), *Structural changes in population and impact on passenger transport*, ECMT Round table 88, OECD, Paris, pp. 103-142.
- Barbier de Saint Hilaire F (2001), *Manuel de référence AEL/TRIBUT, version 3.0 prototype sous Windows95...*, INRETS, Arcueil.
- Barker FN (1952), A basis for estimating traffic diversion to new highways in urban area, rapport au congrès de l'A.A.S.H.O..
- Bates J (1988), Editor, Stated preference methods in transport research, *Journal of Transport Economics and Policy*, vol. XXII, N°1, pp. 1-137.
- Bates J (1998), Reflections on Stated Preference: Theory and Practice, in *Travel Behaviour Research : updating the state of play*, Ortuzar J D, Hensher D, Jara-Diaz S (eds.), Pergamon, Elsevier, Oxford, pp. 89-103.

- Beckman M, McGuire C, Winsten C (1956), *Studies in the economics of transportation*, Yale University Press, New Haven.
- Beesley M (1965), The value of time spent travelling: some new evidence, *Economica*, pp. 174-185.
- Ben-Akiva M, Lerman SR (1985), *Discrete choice analysis, theory and application to travel demand*, MIT Press, Cambridge, Massachusetts, Etats-Unis, 4^{ème} édition, 390p.
- Ben-Akiva M, François B (1983), μ *Homogeneous Generalised Extreme Value Model*, working paper, Department of Civil Engineering, MIT, Cambridge, Massachusetts.
- Benlahrech N, Bonnel P, Doby P, Mermoud F (1995), *Enquêtes déplacements sur la Région Urbaine de Lyon, étude préliminaire*, Rapport pour le compte de la Préfecture de Région Rhône-Alpes, de la Direction Régionale de l'Équipement et le Conseil Régional Rhône-Alpes ; CETE de Lyon, Laboratoire d'Économie des Transports, Ecole Nationale des Travaux Publics de l'État, Lyon, 35p. + annexes.
- Bhat C (1997), Recent methodological advances relevant to activity and travel behavior analysis, in *Recent developments in travel behavior research*, Pergamon, Oxford.
- Bhat C (2000), Flexible model structures for discrete choice analysis, in *Handbook of transport modeling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 71-90.
- Bhat C (2001), Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model, *Transportation research part B*, vol. 35B, N°7, pp. 677-693.
- Bonnafous A (1989), *Le siècle des ténèbres de l'économie*, Economica, Paris, 184p.
- Bonnafous A (1993), Structures démographiques et comportements sociaux, rapport introductif au 12^{ème} symposium de la CEMT Lisboa 1992 *La croissance des transports en question*, CEMT, Paris, pp. 21-77.
- Bonnafous A, Patier-Marque D, Plassard F (1981), *Mobilité et vie quotidienne : suivi du métro de Lyon*, Presses Universitaires de Lyon, Lyon.
- Bonnafous A, Gérardin B (1976), *La demande de transports de voyageurs en milieu urbain, méthodologie de l'analyse et de la prévision*, CEMT, Table ronde 32, Paris, 92p.
- Bonnafous A, Puel H (1983), *Physionomies de la ville*, éditions économie et humanisme, les éditions ouvrières, série initiation économique, Paris, 165p.
- Bonnafous A, Tabourin E (1995), Modèles de simulation stratégique, Laboratoire d'Économie des Transports, 7^{ème} WCTR, Sydney, 17p.
- Bonnafous A, Tabourin E (1998), Modélisation de l'évolution des densités urbaines, in *Données Urbaines 2*, coordonné par Pumain D et Mattei M-F (éd.), collection VILLE, Anthropos, pp. 167-180.
- Bonnel P (1985), *Les besoins latents de déplacement, une méthode de révélation appliquée au cas de Saint-Priest*, thèse de docteur-ingénieur en Économie des Transports, Université Aix-Marseille II, Laboratoire d'Économie des Transports, Lyon, 374p.

- Bonnel P (1989a), *Analyse de l'opérationnalité d'une méthode de simulation des changements de comportements*, Rapport final pour le compte du Service des Etudes, de la Recherche et de la Technologie du Ministère des Transports, Laboratoire d'Economie des Transports, Ecole Nationale des Travaux Publics de l'Etat, 150p. + annexes.
- Bonnel P (1989b), A method of simulating for the effects of transport projects, in *Travel Behaviour Research*, International Association for Travel Behaviour (ed.), Gower Publishing Group, Aldershot, pp. 216-224.
- Bonnel P (1992), Mobilité et contraintes, *6ème Conférence Mondiale sur la Recherche dans les Transports*, Lyon, 29 juin- 3 juillet 1992, 12p.
- Bonnel P (1993), Europe : des politiques de transports urbains très contrastées, in *Transports Urbains*, N°81, pp. 5-15.
- Bonnel P (1995a), An application of activity-based travel analysis to simulation of change in behaviour, in *Transportation*, vol. 22, N°1-2, pp. 73-93.
- Bonnel P (1995b), Urban Car Policy in Europe, in *Transport policy*, vol. 2, N°2, pp. 83-95.
- Bonnel P (1999), La qualité des données d'enquêtes déplacements : téléphone versus face à face, in *Approches nouvelles en Sondages*, Brossier G, Dussaix A-M (éd.), Dunod, Paris, pp. 315-331.
- Bonnel P (2000a), *LYON 21 Étude de faisabilité d'un système de transport radicalement différent pour l'agglomération lyonnaise*, LET, rapport final pour le compte de l'INRETS, Lyon, 84p.
- Bonnel P (2000b), Une mesure dynamique des relations entre transports collectifs, étalement urbain et motorisation, le cas de Lyon, 1976-1995, *Cahiers scientifiques du Transport*, N°38/2000, pp. 19-44.
- Bonnel P (2001), Postal, telephone and face-to-face surveys: how comparable are they?, resource paper for Workshop I: mixed mode surveys, *International conference on Transport Survey Quality and Innovation*, Kruger Park, South-Africa, 5-10 August 2001, 21p.
- Bonnel P, Cabanne I (2000), *Mesure de l'effet de différents facteurs du partage modal dans les transports urbains, étude sur l'agglomération lyonnaise*, LET, Lyon, 77p.
- Bonnel P, Cabanne I, Massot M-H (2001), *Evolution de la part de marché des transports collectifs : quelles perspectives pour les politiques de déplacements urbains ?*, Documentation Française, 57p.
- Bonnel P, Chapleau R, Lee-Gosselin M, Raux C (éd.) (1997), *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Programme Rhône Alpes, Recherches en Sciences Humaines, collection Les chemins de la recherche, actes du colloque de Lyon, 6-8 décembre 1995, 512p.
- Bonnel P, Chausse A (2000), Urban travel: Competition and pricing, *Transport reviews*, vol. 20, N°4, pp. 385-401.
- Bonnel P, Claisse G, Mignot D (2000), *Economie et politique des transports*, cours de 2^e année, Ecole Nationale des Travaux Publics de l'Etat, Vaulx-en-Velin, 176p.
- Bonnel P, Demanget S, Rabilloud J-L, Thome B (1994a), *Politiques de déplacements urbains en Europe, analyse comparative : Espagne, France, Grande Bretagne, Italie, Norvège, Suisse*, LET, Lyon, 49p.

- Bonnel P, Gabet P (1999), *Mesure de l'effet de l'évolution des localisations et de la motorisation sur la part de marché des transports collectifs*, LET pour le compte de la DRAST programme PREDIT, Lyon, 160p.
- Bonnel P, Klein, O (1999), *Séminaire TGV*, document pédagogique réalisé dans le cadre du cours Politique et Economie des Transports, ENTPE, Lyon.
- Bonnel P, Le Nir M (1998), The Quality of survey data: telephone versus face-to-face interviews, *Transportation*, special issue: Extending the scope of travel survey, Guest editors Lee-Gosselin M, Bonnel P, Raux C, vol. 25, N°2, pp. 147-167.
- Bonnel P, Le Nir M, Nicolas J-P (1994b), *Les enquêtes de déplacements urbains, réflexions méthodologiques sur les enquêtes ménages et les enquêtes régionales origine destination canadiennes*, Laboratoire d'Economie des Transports, Lyon, 133p.
- Bonnel P, Le Nir M, Nicolas J-P (1994c), Comparison of French and Canadian urban travel surveys, *seventh International Conference on Travel Behaviour*, Valle Nevado, Santiago, Chili, 13-16 juin 1994, 12p.
- Bonnel P, Thibaud S (1991), *Méthodologie des études de suivi d'une ligne de TCSP*, LET, SEMALY, rapport de synthèse pour le compte du SYTRAL, Lyon, 26p.
- Bonnel P, Thibaud S (1994), Evaluation of urban fixed-track public transport - method and content, *third International Conference on Competition and Ownership in Surface Passenger Transport*, Niagara-on-the-Lake, Ontario, Canada, September 26-29, 1993, Jean Love (ed.), Toronto, Canada, pp. 81-102.
- Bouf D (1989), *Un nouvel instrument d'analyse stratégique pour la RATP : le modèle Gros Quinquin*, thèse de l'Université Lumière Lyon 2, LET, Lyon.
- Bourgin C (1978), *Les évolutions dans l'usage des modes de transports - influence des moments de transition dans le cycle de vie*, IRT, rapport de recherche, N°36, Arcueil, 38p.
- Bouyaux P (1988), Modélisation désagrégée des transports urbains : une application à la ville de Rennes, *Revue d'Economie Régionale et Urbaine*, N°5, pp. 783-809.
- Bowker AH, Lieberman GJ (1965), *Méthodes statistiques de l'ingénieur*, Dunod, Paris.
- Bowman JL, Ben-Akiva ME (2001), Activity based disaggregate travel demand model system with activity schedules, *Transportation Research*, vol. 35A, N°1, pp. 1-28.
- Bradley M (1988), Realism and adaptation in designing hypothetical travel choice concepts, *Journal of Transport Economics and Policy*, vol. 22, N°1, pp. 121-137.
- Brémont J. (1990), *Les économistes néoclassiques de L. Walras à M. Allais, de F. Von Hayek à M. Friedman*, 2^{ème} édition, Hatier, Paris, 248p.
- Brög W (2000), The new Kontiv design, a total survey design for surveys on mobility behaviour, *International Conference on Establishment surveys – II*, June 17-21, 2000, Buffalo, New York, 13p.

- Brög W, Daesler C, Förg O, Gearay H, Hartmann E, Kammhuber S, Kloose H, Linke M (1983), An empirical test of various survey instruments for the recording of travel behaviour, *Second International Conference on Survey Methods in Transport*, Hungerford Hill, Australia, 12-16 September 1983, 26p.
- Brög W, Erl E, Raux C, Jones P (1996), *Changing daily urban mobility less or differently?*, Round table 102, OECD, Paris, 231p.
- Brög W, Meyburg AH (1983), Influence of survey methods on the results of representative travel surveys, *Transportation Research part A*, vol. 17A, N°2, pp. 149-156.
- Brownstone D (1998), Multiple imputation methodology for missing data, non-random response, and panel attrition, in Gärling T, Laitila T, Westin K (eds.) (1998), *Theoretical foundations of travel choice modelling*, Pergamon, Oxford, pp. 421-449.
- Brownstone D (2000), Discrete choice modelling for transportation, resource paper at 9th IATBR *Travel Behaviour Conference*, Australia, July 2000.
- Buisson C (1996), Analyse bibliographique et première classification des algorithmes de simulation de l'écoulement du trafic routier, *Recherche, Transports, Sécurité*, N°53, pp. 35-48.
- Bureau of Public Roads (1964), *Traffic assignment manual*, US Department of Commerce, Washington DC.
- Burrell JE (1968), Multiple road assignment and its application to capacity restraint, in Leutzbach W and Baron P (eds.), *Beiträge zur Theorie des Verkehrsflusses*, Strassenbau und Strassenverkehrstechnik Heft, Karlsruhe.
- Cabanne I, Durand S, Hennebelle P-Y, Masson S, Routhier J-L, Saliou E (2000), *Télescopage : modèle de simulation des trafics de marchandises et de personnes locaux et interurbains dans un espace région-ville*, Rapport final pour l'ADEME, PREDIT, Laboratoire d'Economie des Transports, Lyon, 171p.
- Calot G (1975), *Cours de statistique descriptive*, 2ème édition, Dunod, Paris.
- Cambridge Systematics, Barton-Aschman Associates (1996), *Travel Manual Survey*, report for U.S. Department of Transportation and U.S. Environmental Protection Agency, Washington.
- Campbell EW, McCargar RS, Objective and subjective correlates of expressway use, *Highway Research Board*, bulletin 119, pp. 17-38.
- Caron N (1996), *Les principales techniques de correction de la non-réponse, et les modèles associés*, INSEE, méthodologie statistique, document de travail N°9604, Paris, 27p.
- CEMT (1977), Motifs psychologiques qui guident les usagers, *Table ronde N°34*, Paris.
- CEMT (1982), Review of demand models forecast/recorded traffic comparisons for urban and intercity transport, *Round table 58*, Paris, 168p.
- CERAU, SETEC (1968), *Etude de la demande de transport sur l'axe Paris- Rouen – Le Havre*, Commission de la Communauté Economique Européenne.
- CERTU (1996), *Plan de déplacements urbains*, Lyon, 263p.

- CERTU (1998a), *L'enquête ménages déplacements « méthode standard »*, Collections du CERTU, éditions du CERTU, Lyon, 295p.
- CERTU (1998b), *Comportements de déplacement en milieu urbain : les modèles de choix discrets, vers une approche désagrégée et multimodale*, Dossier du CERTU mobilité-Transport, éditions du CERTU, Lyon, 133p.
- CETE, CETUR (1989), *Dossier méthodologique pour la réalisation d'une enquête ménages*, Lyon, 53p. + annexes.
- CETE de Lyon (1986), *Enquête ménages Lyon 1985 - agglomération lyonnaise 1985-1986, document technique* », CETE DE Lyon, Lyon, 119p.
- CETE de Lyon, CETE de l'Ouest, CETUR, MELATT (1986), *Modèles désagrégés, principes généraux, méthodologie, applications (Grenoble, Nantes), Journées de rencontre sur les modèles désagrégés*, Paris, 10-11 juin 1986, 76p.
- CETE de Lyon, CETUR (1986), *Grenoble 1985 - Une expérience d'enquête par téléphone*, Lyon, 81p. + annexes.
- CETE de Lyon, Conseil Général du Rhône, Conseil Régional Rhône-Alpes, COURLY, DDE du Rhône, Ministère de l'Équipement (1991), *Enquête Cordon de Lyon 1990, Rapport de synthèse, principaux résultats*, Lyon, 153p.
- CETE de Lyon, INSEE, SYTRAL (1995), *Enquête « déplacements auprès des ménages de l'agglomération lyonnaise »*, document technique, Lyon, 119p.
- CETE Méditerranée, CERTU (2000a), *CartOPERA, Logiciel de prévision et d'affectation de trafic urbain, Manuel de présentation*, CETE Méditerranée, Aix-en-Provence, 35p.
- CETE Méditerranée, CERTU (2000b), *CartOPERA, logiciel de prévision et d'affectation de trafic urbain, manuel d'utilisation*, CETE Méditerranée, Aix-en-Provence, 55p.
- CETE Normandie, CETUR, DTT, Ville du Havre (1992), *Enquête ménages sur la région havraise - Expérimentation d'une enquête par téléphone*, Rouen, 45p.
- CETUR (1977), *Éléments pour évaluer la précision obtenue dans l'estimation des indicateurs donnés par les enquêtes ménages*, CETUR, Bagnaux, 38p.
- CETUR (1985), *Les déplacements domicile-travail et domicile-école, modèles désagrégés de choix modal, application au cas de l'agglomération grenobloise*, Bagnaux, 71p.
- CETUR (1989), *Les études de prévision de trafic en milieu urbain - Pourquoi ? - Comment ?*, Bagnaux, 44p.
- CETUR (1990a), *Les études de prévision de trafic en milieu urbain - Guide technique*, Bagnaux, 79p.
- CETUR (1990b), *10 ans de mobilité urbaine, les années 80*, Bagnaux, 86p.
- CETUR (1993a), *Dossier de voirie d'agglomération, réflexions et enjeux*, Bagnaux, 135p.
- CETUR (1993b), *Dossier de voirie d'agglomération, exemples et méthodes*, Bagnaux, 139p.

- CETUR (1994), *Les enjeux des politiques de déplacement dans une stratégie urbaine*, Bagnaux, 368p.
- CETUR, DDE du Rhône (1994), *Définition d'une méthode d'étude intermodale pour les systèmes de transports urbains par l'analyse de la valeur*, Lyon, 28p.
- Chapleau R (1991), *La planification des transports urbains et le traitement désagrégé des informations à référence spatiale*, document d'appoint pour utilisation dans une perspective pédagogique, Ecole Polytechnique de Montréal, Montréal, 431p.
- Chapleau R (1993), Une carte d'utilisation du sol dérivée d'une enquête origine-destination, 28^{ème} Congrès annuel de l'AQTR, Sainte Adèle, Québec, 4-7 avril 1993, 19p.
- Chapleau R (1997), Symphonie d'usage des grandes enquêtes Origine-Destination, en totalement désagrégé majeur, opus Montréal 87 et 93, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 33-59.
- Chapleau R, Allard B, Trepanier M, Morency C (1999), Telephone Interview Origin-Destination Survey Software : the State of the Art, *AQTR-CITE 1999 joint Conference*, Montreal, 19p.
- Chapleau R, De Cea J (1983), La perception de l'offre par les usagers du transport en commun, sous la perspective d'un modèle d'affectation, Ecole Polytechnique de Montréal, *Conférence Mondiale sur la Recherche dans les Transports*, Hambourg, 26-29 avril 1983, 18p.
- Chessel D, Thioulouse J (1990), Auto-modélisation en analyse des données, in *La modélisation, confluent des sciences*, Brissaud M, Forsé M, Zighed A (éd.), Editions du CNRS, Lyon, pp. 71-86.
- Claisse G (1997), *L'abbaye des télémythes, techniques, communication et société*, Aléas, Lyon, 358p.
- Clayton RL, Werking GS (1991), Amélioration de la qualité des données à l'aide d'un modèle de collecte mixte, in *Techniques d'enquête*, Statistiques Canada, vol. 17, N°1, pp. 3-16.
- Clément L (1996), *Review of existing land/use transport models*, CERTU, Lyon, 76p.
- Cochrane RA (1975), A possible economic basis for the gravity model, *Journal of Transport Economics and Policy*, vol. 9, N°1, pp. 34-49.
- Cohen S (1990), Cours d'Ingénierie du trafic, ENPC, INRETS, Arcueil, 161p.
- Cohen S (1991), Flow variables, in *Concise encyclopedia of traffic and transportation systems*, Papageorgiou M (ed.), Pergamon press, Oxford, pp. 139-143.
- Cohen M, Pradel J (1993), *Econométrie, théorie et techniques de base - méthode d'utilisation - exercices*, LITEC économie, Paris, 318p.
- Collins M, Sykes WM (1987), Comparaison de l'interview téléphonique et de l'interview sur place au Royaume Uni, *Techniques d'enquête* Statistiques Canada, vol. 13, N°1, pp. 19-33.
- Dagnelie P (1973), *Théorie et méthodes statistiques, vol. 1*, 2^{ème} édition, Les presses agronomiques de Gembloux, Bruxelles.

- Dagnelie P (1975), *Théorie et méthodes statistiques*, vol. 2, 2ème édition, Les presses agronomiques de Gembloux, Bruxelles.
- Daly A (1982), Applicability of disaggregate models of behaviour: a question of methodology, in *Transportation research part A*, vol. 16A, N°5-6, pp. 363-370.
- Daly A (1985), *A study of transferability of disaggregate mode choice models from Grenoble to Nantes*, Cambridge Systematics, 19p.
- Daly A, Rohr C (1998), Forecasting demand for new travel alternatives, in Gärling T, Laitila T, Westin K (eds.) (1998), *Theoretical foundations of travel choice modelling*, Pergamon, Oxford, pp. 451-471.
- Data Management Group, University of Toronto, Joint program in transportation, with the participation of Transport Ministry Ontario, Regional Municipality of : Metropolitan Toronto, Durham, York, Peel, Halton, Hamilton-Wentworth; Go Transit; Toronto Transit commission (1992), "*1991 TRANSPORT TOMORROW SURVEY : Design and conduct of the survey*" Prepared for the Toronto Area Transportation Planning Data Collection Steering Committee.
- De Cea J, Fernandez E (2000), Transit assignment models, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 497-508.
- Decouflé A-C (1978) Prévision et prospective, in *Traité élémentaire de prévision et de prospective*, Decouflé A-C (éd.), PUF, Paris, pp. 21-41.
- De Leeuw E (1992), *Data Quality in Mail, Telephone and Face to Face surveys*, TT Publikaties Amsterdam, Vrije Universiteit, 168p.
- De Leeuw E, Van der Zouwen J (1992), Data quality and mode of data collection: methodology and explanatory model, in *La qualité de l'information dans les enquêtes*, Dunod, Paris, pp. 11-31.
- De Palma A, Marshal F, Nesterov Y (1997), Metropolis: a modular system for dynamic traffic simulation, *Transportation Research Record*, vol. 1607, pp. 178-84.
- De Palma A, Fontan C (2000), *Enquête MADDIF : Multimotif Adaptée à la Dynamique des comportements de Déplacements en Ile-de-France*, THEMA, Rapport pour la DRAST dans le cadre du PREDIT, Cergy-Pontoise.
- De Palma A, Fontan C (2001), Eléments d'analyse de la composante horaire des déplacements : le cas de la région Ile-de-France, *Les Cahiers Scientifiques du Transport*, N°39, pp. 55-86.
- Dial RB (1971), *A probabilistic multipath traffic assignment model which obviates path enumeration*, *Transportation Research*, vol. 5, N°2, pp. 83-112.
- Dijkstra EW (1959), Note on two problem in connection with graphs (spanning tree, shortest path), *Numerical mathematics*, vol. 1, N°3, pp. 269-271.
- Dijkstra W, Smit JH (1991), Persuasion strategies for reducing refusal rates in telephone survey, *Bulletin de méthodologie sociologique*, N°33, pp. 3-19.
- Domencich T A, Kraft G, Valette J-P (1968), *Estimation of urban passenger travel behavior: an economic demand model*, HRR 238 (HRB).

- Domencich T A, McFadden D (1975), *Urban travel demand : a behavioural analysis*, North-Holland, Elsevier, Amsterdam.
- Dorland P, Giard P, Lavigueur P, Pimpare M (1997), L'enquête téléphonique origine-destination de Montréal 1993, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 235-255.
- DREIF (1995), *Les déplacements des Franciliens en 1991-1992, enquête globale de transport*, Paris, 70p.
- DREIF (1998), *Modus 1 & 2 modèle de déplacements urbains et suburbains*, DREIF, Paris, 30p.
- DREIF (2001), *Les déplacements des Franciliens en 1997-1998, enquête globale de transport*, Paris, 71p.
- Drew JD (1991), Recherche et essais pour les méthodes d'enquêtes par téléphone à Statistiques
- Dupont F (1995), Redressements alternatifs en présence de plusieurs niveaux d'information auxiliaire, *Techniques d'enquêtes*, Statistiques Canada, vol. 21, N°2, pp. 141-150.
- Dupuy G (1975), *Une technique de planification au service de l'automobile, les modèles de trafic urbain*, Document de travail, action concertée de recherches urbaines, Paris, 201p.
- Durand D (1996), *La systémique*, Presses Universitaires de France, collection Que sais-je ? 7e édition, Paris, 126p.
- EPRI (1977), *Methodology for predicting the demand for new electricity-using goods*, EA-593, project 488-1, final report, Electric Power Research Institute, Palo Alto, Californie, Etats-Unis.
- Ettema DF, Timmermans HJP (eds.) (1997), *Activity-based approaches to travel analysis*, Pergamon, Oxford, 371p.
- Faivre d'Arcier B, Andan O, Raux C (1998), Stated adaptation surveys and choice process: some methodological issues, *Transportation*, special issue: Extending the scope of travel survey, Guest editors Lee-Gosselin M, Bonnel P, Raux C, vol. 25, N°2, pp. 169-185.
- Fellendorf M, Haupt T, Heidl U, Scherr W (1997), PTV vision: activity-based demand forecasting in daily practice, in *Activity-based approaches to travel analysis*, Ettema DF, Timmermans HJP (eds), Pergamon, Oxford, pp. 55-72.
- Fernandez JE, Friesz TL (1983), Equilibrium predictions in transportation markets: the state of the art, *Transportation Research*, vol. 17B, N°2, pp. 155-172.
- FHWA (1994), *Highway Capacity manual*, Special report N°209, 3^e édition, Transportation Research Board, Washington.
- Fichelet R (1976), *Approche psychologique des problèmes urbains, sur quelques concepts méthodologiques*, SERES, Paris.
- Fisk CS (1977), Note on the ML calibration of Dial's assignment model, *Transportation Research* vol. 11, pp. 67-68.

- Focas C (1989), *Determining the minimum sample size required for a large household based transport survey: the case of the 1991 London area transport survey*, London Research Centre, Londres, 7p.
- Frank M, Wolfe P (1956), An algorithm for quadratic programming, *Naval Research Logistic Quarterly*, vol. 3, pp. 95-110.
- Fréjean M, Panzani J-P, Tassi P (1990), Les ménages inscrits en liste rouge et les enquêtes par téléphone in *Journal de la Statistique de Paris*, vol. 131, N°3-4, pp. 86-102.
- Friesz TL, Bernstein D (2000), Analytical Dynamic Traffic Assignment Models, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 181-195.
- Gallez C (1995), Une nouvelle perspective pour la projection à long terme des comportements d'équipement et de motorisation, *Recherche Transport Sécurité*, N°48, pp. 3-14.
- Gärling T, Laitila T, Westin K (eds.) (1998), *Theoretical foundations of travel choice modelling*, Pergamon, Oxford, 498p.
- Gayda S (1993) *Réalisation d'une enquête « stated preference » sur le stationnement près des zones d'emploi*, en Région bruxelloise, séminaire AGIR, N°26, faculté universitaire de Namur, Belgique, 14 juin 1993.
- Géraud P (1977), *Les modèles de demande en transport interurbain de voyageurs*, Institut de Recherche des Transports, note d'information N°10, Arcueil, 58p.
- Géronneau (1966), *Réseaux de transport et Urbanisation*, note du SERC, Paris
- GESMAD (2000), *Evaluation des modèles de prévision de trafic, note de synthèse*, rapport pour le PREDIT, Paris, 125p.
- Ghiglione R, Matalon B (1985), *Les enquêtes sociologiques - théories et pratiques*, Armand Collin, Collection U, Paris, 301p.
- Golob TF, Kitamura R, Long L (eds.) (1997), *Panels for transportation planning : methods and applications*, Kluwer, Boston.
- Goodwin PB (1984), *Evolution de la motivation des usagers en matière de choix modal*, rapport de la 68^{ème} table ronde de la CEMT, Paris.
- Goodwin PB (1998), The end of Equilibrium, in *Theoretical foundations of travel choice modelling*, Gärling T, Laitila T, Westin K (eds.), Pergamon, Elsevier, Oxford, pp. 103-132.
- Groves RM, Biemer PP, Lyberg LE, Massey JT, Nichols II WL, Waksberg J (1988), *Telephone Survey Methodology*, John Wiley&Sons, New-York, 581p.
- Guitton H (1964), *Statistiques et économétrie*, Dalloz, Paris.
- Gunn HF (1985), *Value of time for evaluation purposes: the state of the art*, report N°421-01, Hague Consulting Group, The Hague (Netherlands).
- Gunn HF (2000), An introduction to the valuation of Travel-time Savings and Losses, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 433-448.

- Gunn H, Van der Hoorn T (1998), The predictive power of operational demand models - a case study, *World Conference on Transport Research*, Anvers, 12-17 juillet 1998, 17p.
- Hägerstrand T (1974), *Les effets du transport sur la qualité de la vie*, 5^e symposium international sur la théorie et la pratique dans l'économie des transports, les transports dans la décennie 1980-1990, vol. 1, Athènes, octobre 1973, CEMPT, Paris.
- Hausman J, McFadden D (1984), Specification tests for the multinomial logit model, *Econometrica*, N°52, pp. 1219-1240.
- Henderson JM, Quandt RE (1980), *Microeconomic theory a mathematical approach*, McGraw-Hill International editions, economics series, 3^e édition, 420p.
- Henn V (2001), *Information routière et affectation du trafic : vers une modélisation floue*, Thèse de Doctorat de l'Université de Saint-Etienne, St.-Etienne, 239p.
- Hensher DA (1994), Stated preference Analysis of travel choices : the state of practice, *Transportation*, vol. 21, N°2, pp. 107-133.
- Hensher DA, Button KJ (eds.) (2000), *Handbook of transport modelling*, Pergamon, Elsevier, 666p.
- Hensher DA (ed.) (2001), *Travel behaviour research, the leading edge*, Pergamon, Amsterdam.
- Hivert L, Orfeuil J-P, Troulay P (1988), *Modèles désagrégés de choix modal : réflexions méthodologiques autour d'une prévision de trafic*, INRETS, rapport N°67, Arcueil, 65p.
- Horowitz AJ (1982), Specification tests for probabilistic choice models, *Transportation research part A*, vol. 16 N°5-6, pp. 383-94.
- Horowitz A J (1993), *Reference Manual, Quick Reference System II for windows, version 3.6*, Centre for transportation studies, University of Wisconsin, Milwaukee, 207p.
- IAURP (1963), *Modèles de trafic – analyse bibliographique*, IAURP, Paris.
- INSEE, INRETS (1997), Spécial enquête Transports et Communications, *Recherche, Transports, Sécurité*, N°56-57, Dunod, Paris.
- INRETS (1982), *Les effets d'une nouvelle station de métro sur le commerce avoisinant*, rapport de recherche N°57, INRETS, Paris.
- INRETS (1985), *Aspects méthodologiques des études de suivi. Application aux métros lyonnais et lillois*, rapport de recherche, N°77, INRETS, Paris, 112p.
- INRETS (1989), *Un milliard de déplacements par semaine*, La Documentation Française, Paris, 293p.
- INSEE, SES (1998), *Les comptes des Transports en 1997*, Paris.
- International Association for Travel Behaviour (ed.) (1989), *Travel Behaviour Research*, Gower Publishing Group, Aldershot.
- Jara-Diaz SR (2000), Allocation and Valuation of Travel-time Savings, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 303-319.

- Javeau C (1985) *L'enquête par questionnaire - manuel à l'usage des praticiens*, Editions de l'Université de Bruxelles, Bruxelles, 138p.
- Jones P (1997), Using computer-aided interactive surveys to explore choice sets and decision processes, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux, C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 437-450.
- Jones PM, Dix MC, Clarke MI, Heggie IG (1980), *Understanding travel behaviour*, Oxford studies of Transport, Gower, 241p. et annexes.
- Junod A (1997), Raisons du passage du mode épistolaire au mode téléphonique pour l'enquête suisse sur les transports, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux, C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 161-176.
- Karsten HM, Koning HJ (1985), *Differences in estimate due to change in methods of data collection*, Centraal Bureau voor de Statistiek, The Netherlands, 17p.
- Kaufmann V, Guidez J-M (1996), *Les citoyens face à l'automobilité – les déterminants du choix modal*, rapport du FIER N°19, Paris.
- Kaufmann V, Guidez J-M (2000), *Vers de nouvelles dynamiques urbaines écomobiles ?*, Paris, Lyon, Strasbourg, Aix-en-Provence, Rapport de recherche 148 pour le PREDIT, EPFL-IREC, CERTU, DRAST, Paris, 164p. + annexes.
- Kitamura R (2000), Longitudinal methods, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 113-129.
- Kostyniuk LP, Kitamura R (1987), Effects of ageing and motorisation on travel behaviour: An exploration, *Transportation Research Record 1135*, pp. 31-36.
- Kraft G (1963), *Demand for intercity passenger travel in the Washington-Boston corridor*, US Department of Commerce, Washington.
- Koppelman ES, Sethi V (2000), Closed-form discrete-choice models, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 211-227.
- Kulka RA & Weeks MF (1988), Toward the development of optimal calling protocols for telephone surveys: a conditional probabilities approach, *Journal of Official Statistics*, vol. 4, N°4, pp. 319-332.
- Lamb GM, Havers G (1970), Introduction to transportation planning: Treatment of networks, *Traffic Engineering and Control*, vol. 11, N°11, pp. 486-489.
- Lave C (1970), The demand for urban mass transit, *Review of economics and statistics*, N°52, pp. 320-323.
- Lee-Gosselin MEH (1997), Portée et potentiel des méthodes de collecte de données de type « réponses déclarées interactives » (RDI), in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux, C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 287-318.

- Lee-Gosselin M, Bonnel P, Raux C (1998), Extending the scope of travel surveys, *Transportation*, vol. 25, N°2, Special issue.
- Legay J-M (1990), De la complexité des objets à la méthode des modèles - synthèse des travaux du colloque, in *La modélisation, confluent des sciences*, Brissaud M, Forsé M, Zighed A (éd.), Editions du CNRS, Lyon, pp. 235-239.
- Lejeune M (2001), *Traitements des fichiers d'enquête, redressements, injections de réponses, fusions*, Lejeune (éd.), Presses Universitaires de Grenoble, collection Libre Cours, Grenoble, 127p.
- Lenntorp B (1978), Les déplacements considérés comme une part de la vie : un cadre conceptuel pour l'analyse de la distribution des possibilités de déplacements au sein d'une population, colloque *la mobilité dans la vie urbaine*, Arc-et-Senans, IRT, Paris, pp. 155-172.
- Le Moigne J-L (1994), *La théorie du système général - Théorie de la modélisation*, 4ème édition, PUF, Paris, 338p.
- Le Nir M (1991), *Les modèles de prévision de déplacements urbains*, thèse d'Université d'Economie des Transports, Université Lumière Lyon 2, Lyon, 313p.
- Lerman SR (1984), Recent advances in disaggregate demand modelling, in Florian M (ed.), *Transportation planning models*, North Holland, Amsterdam.
- Leurent F (1993), Cost versus time equilibrium over a network, *European Journal of Operational Research*, N°71, pp. 205-221.
- Leurent F (1995a), *Constitution d'une boîte-à-outils pour opérer diverses affectations statiques du trafic routier*, note de travail 95.2, INRETS, Arcueil, 120p.
- Leurent F (1995b), Un algorithme pour résoudre plusieurs modèles d'affectation du trafic : la procédure d'égalisation par transvasement, *Les Cahiers Scientifiques du Transport*, N°30, pp. 31-49.
- Leurent F (1995c), *Une famille d'algorithmes d'équilibrage pour des modèles d'affectation statique du trafic routier*, note de travail 95.1, INRETS, Arcueil, 19p.
- Leurent F (1997), Analyse et mesure de l'incertitude dans un modèle de simulation, les principes, une méthode et l'exemple de l'affectation bicritère du trafic, Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, Paris, 301p.
- Lichère V, Raux C (1997a), *Développement d'un modèle stratégique de simulation des déplacements, présentation générale*, SEMALY, LET, Lyon, 28p. + annexes.
- Lichère V, Raux C (1997b), *Développement d'un modèle stratégique de simulation des déplacements, guide de l'utilisateur*, SEMALY, LET, Lyon, 33p.
- Lindsey R, Verhoef E (2000), Congestion modeling, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 353-373.
- Lisco T (1967), *The value of commuter travel time: a study in urban transportation*, University of Chicago, Ph.D. Dissertation.

- Liss S (1999), The effects of survey methodology changes in the NPTS, *TRB conference on personal travel : the long and short of it*, Washington, June-July 1999, 15p.
- Lois et Décrets : publication au Journal Officiel : www.legifrance.gouv.fr
- Louvière JL (1988), Conjoint analysis modelling of stated preferences, *Journal of Transport Economics and Policy*, vol. 22, N°1, pp. 93-119.
- Louvière JL (1992), editor, Special issue on experimental choice analysis, *Journal of Business Research*, vol.24, N°2, pp. 89-189.
- Luce R (1959), *Individual choice behaviour: a theoretical analysis*, Wiley, New York
- Luce R, Suppes P (1965), Preference, utility and subjective probability, in *Handbook of Mathematical psychology*, vol. 3, Luce R, Bush R, Galanter E (eds.), Wiley, New York.
- Mackinder IH, Evans SE (1981), *The predictive accuracy of British studies in urban areas*, Transport and Road Research Laboratory, Crowthorne, 34p.
- McFadden D (1978), Modelling the choice of residential location, in *Spatial Interaction Theory and Planning Models*, Karlqvist A, Lundqvist L, Snickars F, Weibull J (eds), North Holland, Amsterdam, pp. 75-96.
- McFadden D (2000), Disaggregate behavioral travel demand's RUM side, a 30-year retrospective, in pre-print *9th International Association for Travel Behaviour research Conference*, Gold Coast, Queensland, Australia, 2-7 July 2000, vol. 1, 38p.
- McFadden D, Tye W, Train K (1977), An application of diagnostic tests for the irrelevant alternatives property of the multinomial logit model, *Transportation Research Record*, N°637, pp. 39-46.
- McLynn JM, Woronka T (1969), *Passenger demand and modal split models*, report 230, Northeast Corridor Transportation project, US Department of Transport, Washington.
- Madre J-L (1980), *Ajustement et extrapolation de tableaux statistiques*, thèse de doctorat spécialité statistiques – mathématiques à l'Université Pierre et Marie Curie Paris 6, Paris, 105p.
- Madre J-L (1997), L'enquête transports, *Recherche, Transports, Sécurité*, N°56, Dunod, Paris, pp. 6-8.
- Madre J-L, Armoogum J (1996), *Motorisation et mobilité des franciliens aux horizons 2010 - 2020*, Rapport INRETS, N° 209, Arcueil : INRETS, 125p.
- Madre J-L, Lambert T (1989), *Prévisions à long terme du trafic automobile*. Paris : CREDOC, Collection des rapports.
- Mahmassani HS (2000), Trip timing, in *Handbook of transport modelling*, D A Hensher, K J Button (eds.), Pergamon, Elsevier, pp. 393-407.
- Manheim ML (1984), *Fundamentals of transportation systems analysis, vol. 1: basic concepts*, 4^e édition, MIT Press, Cambridge, Massachusetts, Etats-Unis, 658p.
- Marche R, Dumontier JB (1974), *Le modèle de demande TRIP*, ensemble de notes de travail de la SETEC et du SAEI, Paris.

- Margail F (1989), *Vers une prospective des déplacements urbains*, Mémoire de DEA, ENTPE, Université Lumière Lyon 2, Lyon, 176p.
- Massot M-H, Orfeuil J-P (1989), *Offre et demande de transport en commun dans les villes françaises sans métro - Usages et usagers des services de transport en commun urbain*, Rapport INRETS N° 100, premier tome, Arcueil : INRETS, 63p.
- Massot M-H, Orfeuil J-P (1990), *Offre et demande de transport en commun dans les villes françaises sans métro - Structures de l'offre de transport*, Rapport INRETS, N° 121, deuxième tome, Arcueil : INRETS, 73p.
- Massot M-H, Orfeuil J-P (1991), *Offre et demande de transport en commun dans les villes françaises sans métro - Sensibilité de l'usage à l'offre*, Rapport INRETS, N° 148, troisième tome, Arcueil : INRETS, 76p.
- Matalon B (1978), Les limites de la prévision scientifique, in *Traité élémentaire de prévision et de prospective*, Decouflé A-C (éd.), PUF, Paris, pp. 105-130.
- Mayberry J-P (1973), Structural requirements for abstract-mode models of passenger transportation, In Quandt R E (ed.), *The demand for travel: Theory and Measurement*, D C Heath and Co., Lexington, Mass.
- Mercadal G (1970), *Contribution à une psychosociologie des comportements urbains, choix du mode de transport*, rapport introductif à une table ronde de la CEMT, Publications de recherches urbaines, Ministère de l'Équipement et du Logement, Paris.
- Moore EF (1957), The shortest path through a maze, *Proceedings International Symposium on the Theory of Switching*, Harvard University Press, Cambridge, pp. 285-292.
- Morellet O, Berland D, Boulnois D, Bialka M, Hervé J, Le Blanc F, Lesage M-A, Lunel A, Postel D, Siméone R (1980), *Un modèle de choix des usagers entre itinéraires concurrents*, SETRA, rapport d'étude, Bagneux, 50p.
- Morellet O, Julien H (1990), *M.A.T.I.S.S.E. : Un modèle intégrant étroitement induction et partage modal fin du trafic*, Rapport INRETS, Paris, 139p.
- Morellet O, Marchal P (1995), *M.A.T.I.S.S.E. : Un modèle intégrant étroitement contexte socio-économique et offre de transport*, Rapport INRETS, N° 203, Paris, 250p.
- Morellet O, Marchal P (1999), Formulation théorique de l'induction de trafic : garantir à la fois le signe du résultat et la cohérence vis-à-vis du partage modal, in *les Cahiers Scientifique du Transport*, N°35, pp. 73-102.
- Morikawa T (1994), Correcting stated dependence and serial correlation in the RP/SP combined estimation method, *Transportation*, vol.21, N°2, pp. 153-165.
- Moritz G, Brög W (1999), Redesign of the Dutch travel surveys : response improvement, *TRB conference on personal travel : the long and short of it*, Washington, June-July 1999, 13p.
- Morlok EK (1978), *Introduction to transportation engineering and planning*, McGraw-Hill, Etats-Unis, 767p.

- Moskowitz K, California method of assigning diverted traffic to proposed freeway, *Highway Research Board*, bulletin 130, pp. 1-26.
- Nathan G, Eliav T (1988), Comparison of Measurement Errors for Telephone Interviewing and Home Visits by Misclassification Models, *Journal of Official Statistics*, Sweden, vol. 4, N°4, pp. 363-374.
- Nicolas J-P (1996), *Villes, transports et environnement. Contribution relatives des paramètres du trafic routier affectant la pollution sonore et atmosphérique en milieu urbain*, Université Lumière Lyon 2, thèse de doctorat, Lyon, 420p.
- Nicolas J-P, Pochet P, Poimboeuf H (2001), *Indicateurs de mobilité durable sur l'agglomération lyonnaise, méthodes et résultats*, Recherche pour le compte de Renault, APDD, LET, Lyon, 235p.
- Offner J-M (1985), Aspects méthodologiques d'un suivi exploratoire du VAL sur deux quartiers lillois, Acte des journées CETUR, sur *Les suivis des grandes infrastructures de déplacements urbains*, 13 et 14 juin 1985, Bagneux, 317p.
- Orfeuil J-P (1992), *Structural changes in population and impact on passenger transport*, ECMT Round Table 88, OECD, Paris, pp. 43-102.
- Orfeuil J-P (1999), *La mobilité : analyses, représentations, controverses*, mémoire de thèse d'habilitation à diriger les recherches, Université Paris XII, IUP, Paris, 345p.
- Orfeuil J-P (2000), *Stratégies de localisation, ménages et services dans l'espace urbain*, La Documentation Française, Collection le point sur, Paris, 75p.
- Ortuzar J de D, Hensher DA, Jara-Diaz SR (eds.) (1998), *Travel Behaviour Research: Updating the State of Play*, Pergamon, Elsevier, Oxford.
- Ortuzar J de D, Willumsen LG (1994), *Modelling transport*, 2^e édition, John Wiley & Sons, UK.
- Papageorgiou M (ed.) (1991), *Concise encyclopedia of traffic and transportation systems*, Pergamon press, Oxford.
- Plassard F (1988), Le réseau TGV et les transformations de l'espace - la Région Rhône-Alpes, *Les Annales de la Recherche Urbaine*, N°39.
- Platak R, Pierre-Pierre F K, Stevens P (1985), Elaboration et conception des questionnaires d'enquête, Statistique Canada, Division des méthodes de recensement et d'enquêtes-ménages, Ottawa, 131p.
- Pochet P (1995), *Mobilité quotidienne des personnes âgées en milieu urbain : évolutions récentes et perspectives*, thèse de doctorat de sciences économiques, Université Lumière Lyon 2, Lyon, 321p.
- Polak J, Ampt E (1996), An analysis of response wave and non-response effects in travel diary survey's, proceedings 4th *International Conference on Surveys Methods in Transport*, Steeple Aston, Oxford, UK, September 8-11 1996, pp. 232-247.
- Polak J, Jones P (1997), Using stated-preference methods to examine travellers preferences and responses, In Stopher P, Lee-Gosselin M (eds.) *Understanding travel behaviour in an era of change*, Pergamon, pp. 177-207.
- Quandt RE, Baumol WJ (1966), The demand for abstract transport modes : theory and measurement, *Journal of Regional Science*, vol. 6, N°2, pp. 13-26.

- Quinet E (1982), *La demande de transport, de la modélisation de la demande de transport à l'appréhension des besoins*, Presse de l'ENPC, Paris, 323p.
- Quinet E (1998), *Principes d'économie des transports*, Economica, 419p.
- RATP, Cambridge Systematics (1982), *Etudes de politiques de transports en région Ile-de-France, mise au point et utilisation de modèles désagrégés de choix modal*, Paris, 91p.
- Raux C (1983), *Modèles et prévision des comportements de mobilité quotidienne*, Thèse pour le doctorat de docteur-ingénieur en économie des transports, Université Lumière Lyon 2, ENTPE, Lyon, 256p.
- Raux C (1996), Réduire ou repenser la mobilité urbaine quotidienne, CEMT table ronde 102, Paris, pp. 89-138.
- Raux C, Godinot C, Masson S (1996), *Développement d'un modèle stratégique de simulation des déplacements, vingt ans de rétrospective à travers les enquêtes déplacements de l'agglomération lyonnaise (1976-1985-1995), rapport d'étape*, LET, SEMALY, Lyon, 159p.
- Raux C, Lichère V, Masson, S (1998), *Modèle stratégique de déplacements de l'agglomération lyonnaise, tests de sensibilité, erreurs et incertitudes liées à la prévision*, LET, SEMALY, Lyon, 64p.
- Raux C, Tabourin E (1991), *Les investissements en transports collectifs dans l'agglomération lyonnaise : simulation des effets et risques financiers*, LET, Lyon, 70p.
- Renaud J-F, Tabourin E (1998), *Les grandes fonctions économiques, la production et les dépenses*, ellipses, Paris, 348p.
- Richardson AJ (1982), Search models and choice set generation, *Transportation Research part A*, vol. 16 N°5/6, pp. 403-19.
- Richardson AJ (2000), Behavioural Mechanisms of Non-Response in Mailback Travel Surveys, 79th *Transportation Research Board*, Washington, 18p.
- Richardson AJ, Ampt E (1993), *The Victoria Integrated travel, activities and land-use toolkit*, VITAL working paper VWP93/1, Transport Research Centre, University of Melbourne.
- Richardson AJ, Ampt ES, Meyburg AH (1995), *Survey methods for transport planning*, Eucalyptus press, University of Melbourne, Australia, 459p.
- Ritschard G (1990), Détection de données atypiques, in *La modélisation, confluent des sciences*, Brissaud M, Forsé M, Zighed A (éd.), Editions du CNRS, Lyon, pp. 95-110.
- Rousseau J, Saut C (1997), Un outil de simulation de politiques de transport: impact 3, *Revue générale des chemins de fer*, pp. 77-83.
- Roy G, Vanheuverzwyn A (2000), Le téléphone mobile dans les enquêtes par sondage, VIIèmes *Journées de Méthodologie Statistique de l'INSEE*, Paris, 4-5 décembre 2000, Médiamétrie, 11p.
- Salenius (1972), Un modèle de distribution de trajets avec opportunité d'intervention et voyageurs concurrents, *Transportation Research*, vol. 6-2, pp. 169-185.

- Salomon I, Bovy P, Orfeuil J-P (eds.) (1993), *A billion trips a day, tradition and transition in European travel patterns*, Kluwer academic publishers, The Netherlands, 460p.
- Sammer G (1995), Problems and solutions in Urban Travel Survey, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux, C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 145-159.
- Schéou B (1997), *Modélisation des déplacements domicile-travail en milieu péri-urbain : le cas de la région lyonnaise*, thèse de doctorat, Université Lumière Lyon 2, Lyon, 317p.
- Schittenhelm H (1990), On the integration of an effective assignment algorithm with path-flow management in a combined trip distribution and traffic assignment algorithm, *PTRC, 18th Summer Annual Meeting*, proceedings of Seminar H.
- Schneider M (1959), Gravity models and trip distribution theory, *papers and proceedings of the Regional Science Association*, V(1), pp. 51-56.
- Ségalou E, Routhier J-L, Durand S (2001), *Comment mesurer l'impact du transport de marchandises en ville ? - Le modèle de simulation FRETURB (V.1)*, DRAST-ADEME, document N°2, Lyon, 90p.
- Ségonne C (1998), *Comportements de choix d'itinéraires, le cas des automobilistes marseillais confrontés au péage urbain de financement*, Thèse de doctorat en Economie des Transports, Université Lumière Lyon 2, Lyon, 250p.
- SEMALY (1996), *Evaluation de la modélisation des trafics sur la ligne D du métro de Lyon*, rapport pour le compte de la DRAST, Lyon, 106p.
- SEMALY (2000), *TERESE, modèle d'affectation de voyageurs dans les études de transport collectif*, document pédagogique, cours analyse et prévision de la demande de transport du DESS Transports Urbains et Régionaux de Personnes, ENTPE, Université Lumière Lyon 2, Lyon.
- SETRA (1973), *Etudes préliminaires d'Infrastructures de transport - références méthodologiques*, SETRA, Bagnoux.
- SETRA, CETUR (), *Caractéristiques des déplacements en milieu urbain, dossier 3 : distribution des déplacements de personnes*, SETRA, Bagnoux.
- Sheffi Y (1985), *Urban Transportation Networks*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Small K, Hsiao C (1982), *Multinomial logit specification tests*, working paper, Departments of Economics, Princeton University, Princeton, Etats-Unis.
- Smit JH, Dijkstra W (1991), Persuasion strategies for reducing refusal rates in telephone survey, *Bulletin de Méthodologie Sociologique*, N°33, pp. 3-19.
- SMTC (1992), *Enquête ménages de l'agglomération grenobloise 1992, document technique*, SMTC, Grenoble, 55p. + annexe.
- Sossiau AB, Hassam AB, Carter MM, Wickstromc GV (1978), *Quick-response urban travel estimation techniques and transferable parameters, user's guide*, National Cooperative Highway Research Program, report 187, TRB, Washington D.C.

- Srinivasan KK, Mahmassani HS (2000), Kernel logit method for the longitudinal analysis of discrete choice date: some numerical experiments, Paper presented at the 9th *International Association of Travel Behaviour Research Conference*, vol. 3, Method workshop 2, methodological developments in travel behaviour modelling, July 2-7 2000, Gold Coast, Australia, 25p.
- Stopher PR (1998), A review of separate and joint strategies for the use of data on revealed and stated choices, *Transportation*, special issue: Extending the scope of travel survey, Guest editors Lee-Gosselin M, Bonnel P, Raux C, Vol25, N°2, pp. 187-205.
- Stopher PR, Lee-Gosselin MEH (eds.) (1997), Understanding travel behaviour in an era of change, Pergamon, Oxford, 583p.
- Stouffer A (1940), Intervening opportunities: a theory relating mobility and distance, *American Sociological Review*, vol. 5, N°6, pp. 845-867.
- Supernak J (1983), Transportation modelling: lessons from the past and tasks for the future, *Transportation*, N°12, pp. 79-90.
- Supernak J, Talvitie AP, Dejohn A (1983), Person category trip generation modelling, *Transportation Research Record* 944, pp. 74-83.
- Swait J (2001), Choice set generation within the generalized extreme value family of discrete choice models, *Transportation Research, part B*, vol. 35B, N°7, pp. 643-666.
- SYTRAL, COURLY, Conseil Général du Rhône, Ministère des transports (1990), Comment se déplacent les lyonnais, Lyon, 29p.
- Tabourin E (1989), *Un modèle de simulation du financement des transports collectifs urbains à l'an 2000 : le modèle QuinQuin, application à l'agglomération lyonnaise*, thèse de l'Université Lumière Lyon 2, LET, Lyon.
- Tabourin E (1995), Les formes de l'étalement urbain, la logique du modèle de Bussière appliquée à l'agglomération lyonnaise, *Les annales de la Recherche Urbaine*, N°67, Densités et Espacements, Pir-Ville - Plan Urbain, pp. 32-42.
- Tassi P (1992), La qualité dans les enquêtes téléphoniques, l'échantillon des répondants, in *La qualité de l'information dans les enquêtes*, Dunod, Paris, pp. 33-54.
- Thomas R (1991), *Traffic assignment techniques*, Aveburry Technical, Gower House, Aldershot, 385p.
- Thurstone L (1927), A law of comparative judgment, *Psychological review*, N°34, pp. 273-286.
- Timmermans H, Arentze T, Hofman F, Kalfs N (1997), Data needs, data collection and data quality requirements of activity-based transport demands models, *International Conference on Transport Survey Quality and Innovation*, Grainau, Germany, May 24-30, 31p.
- Van Evert H, Moritz G (2000), The new Dutch travel survey, *9th International association for travel behaviour conference* (proceedings), Gold Coast, Queensland, Australia, July 2-7 2000, 16p.
- Van Vliet D (1978), Improved shortest path algorithm for transport networks, *Transportation Research*, vol. 12, N°1, pp. 7-20.

- Van Vliet D, Dow PC (1979), Capacity restrained road assignment, *Traffic Engineering and Control*, vol. 20, N°6, pp. 296-305.
- Viaurox C (2001), *Essai sur la demande de transport en milieu urbain*, Thèse de doctorat en Sciences Economiques, Université de Toulouse 1, Toulouse, 271p.
- Viégas JM (1997), Mobility surveys in Porto and Lisbon: data production methodologies and critical evaluation, in *Les enquêtes de déplacements urbains : mesurer le présent, simuler le futur*, Bonnel P, Chapleau R, Lee-Gosselin M, Raux C (éd.), collection Programme Rhône-Alpes Recherches en Sciences Humaines, Lyon, pp. 217-234.
- Villé J (1970), *Le paradoxe de la congestion*, note ronéotée, Ministère français de l'équipement.
- Volle M (1989), *Analyse des données*, Economica, Collection Economie et statistiques avancées, Paris.
- Wachs M (1993), Learning from Los Angeles: transport, urban form and air quality, *Transportation* 20, pp. 329-354.
- Wardrop JG (1952), Some theoretical aspects of road traffic research, *Proceedings of the Institution of Civil Engineers*, part II-1, pp. 325-378.
- Wen C-H, Koppelman FS (2001), The generalized nested logit models, *Transportation Research, part B*, vol. 35B, N°7, pp. 627-641.
- Wermuth M, Sommer C, Kreitz M (2001), Impact of new technologies in travel surveys, resource paper for Workshop B-3: impact of new technologies, *International conference on Transport Survey Quality and Innovation*, Kruger Park, South-Africa, 5-10 August 2001, 29p.
- White R (1963), Volume speed and concentration, in Davies E (ed.), *Traffic Engineering Practice*.
- Wiel M (1999), *La transition urbaine ou le passage de la ville pédestre à la ville motorisée*, MARDAGA, Col. Architecture + Recherches, Sprimont (Belgique), 149p.
- Wildert S (1998), Stated Preferences Studies: The Design Affects the Results, in *Travel Behaviour Research : updating the state of play*, Ortuzar J D, Hensher D, Jara-Diaz S (eds.), Pergamon, Elsevier, Oxford, pp. 105-121.
- Wilson AG (1974), *Urban and regional models in geography and planning*, John Wiley and Sons, Londres.
- Wonnacott TH, Wonnacott RJ (1991), *Statistiques* (traduction française de l'ouvrage *Introductory statistics for business and economics*, 4^e édition, John Wiley & Sons, New York), Economica, Paris.
- Zhang C, Mercier S, Bonnel P, Clément L, Morain J-E (2001), *Définition d'une typologie de la mobilité adaptée à la modélisation de la demande, partie I*, Programme de recherche Prédit, DRAST, ISIS, LET, Lyon.

Liste des tableaux

Tableau 1.1 : Tests de cohérence dans les études de transport au cours des années 1960	18
Tableau 1.2 : Mise en place d'une procédure de planification urbaine	20
Tableau 1.3 : Echelles spatio-temporelles et modélisation	28
Tableau 2.1 : Comparaison modèle théorique, données d'enquête	58
Tableau 3.1 : Impact des localisations sur l'évolution de la part de marché des transports collectifs (sur l'ensemble VP + TC)	86
Tableau 3.2 : Impact des effets simples autres que localisation sur l'évolution de la part de marché des transports collectifs (sur l'ensemble VP + TC)	87
Tableau 3.3 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 18 ans et plus	94
Tableau 3.4 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 60 ans et plus	95
Tableau 3.5 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 39 à 59 ans	95
Tableau 3.6 : Décomposition de la croissance du nombre de déplacements automobiles chez les personnes de 18 à 38 ans	95
Tableau 4.1 : Principales enquêtes déplacements disponibles en milieu urbain en France	117
Tableau 4.2 : Comparaison des résultats d'affectation à partir d'adresses en coordonnées x-y ou par zone (Chapleau, De Cea, 1983)	128
Tableau 4.3 : Taxinomie des approches d'enquêtes de Réponses Déclarées	136
Tableau 4.4 : Variables étudiées dans l'exercice « <i>stated preference</i> », sur les trains à longue distance entre Stockholm, Göteborg et Sundsvall (Wildlert, 1998)	137
Tableau 6.1 : Répartition des déplacements par motifs	170
Tableau 6.2 : Nombre de déplacements attirés selon les caractéristiques de la zone	174
Tableau 6.3 : Nombre de déplacements émis selon le revenu ou la motorisation de la zone	174
Tableau 6.4 : Jeu de données	179
Tableau 7.1 : Distribution des déplacements : notations	191
Tableau 7.2 : Matrice de référence	195
Tableau 7.3 : Matrice à l'horizon d'étude avec un facteur de croissance uniforme (= 1,2)	195
Tableau 7.4 : Facteurs de croissance en ligne et en colonne	196
Tableau 7.5 : Résultat de l'algorithme de Furness (4 ^e itération)	197
Tableau 7.6 : Résultat de l'algorithme de Fratar (3 ^e itération)	198
Tableau 8.1 : Répartition modale selon le nombre de voitures à disposition, Lyon 1995	224
Tableau 8.2 : Nombre de déplacements quotidiens et population enquêtée selon le nombre de voitures à disposition, Lyon 1995	225
Tableau 8.3 : Exemple de grille de part modale pour les transports en commun pour le motif domicile-travail (CETUR, 1990a)	226
Tableau 9.1 : Illustration des résultats de la propriété IIA pour une population non homogène	259
Tableau 9.2 : Calcul de l'élasticité pour des fonctions courantes	277
Tableau 9.3 : Calcul de dérivée pour des fonctions courantes	277
Tableau 9.4 : Procédure de sélection des variables	279
Tableau 10.1 : Exemples de zonages pour différentes études	311
Tableau 10.2 : Classification des modèles d'affectation	329
Tableau annexe 1.1 : Trafic ferroviaire et aérien en 1996	362
Tableau annexe 1.2 : Temps de parcours ferroviaire et aérien en 1996	363
Tableau annexe 1.3 : Tarifs ferroviaire et aérien en 1996	363

Liste des illustrations

Graphique 1.1 : Evolution du taux de motorisation dans les pays de l'OCDE	14
Schéma 1.1 : Groupe A - faible interaction	30
Schéma 1.2 : Groupe B - prise en compte partielle des interactions entre les sous-systèmes urbains	31
Schéma 1.3 : Groupe C - interactions prépondérantes entre les trois sous-systèmes	32
Schéma 2.1 : Procédure de prévision	34
Graphique 2.1 : Courbe de coût généralisé fer-avion	52
Graphique 2.2 : Forme de la fonction de densité log-normale	53
Graphique 2.3 : Forme de la fonction de répartition log-normale	53
Graphique 2.4 : Calage de la droite de Henry	55
Graphique 2.5 : Application du modèle prix-temps sur le partage air-fer - courbe de régression	59
Schéma 3.1 : Diagramme espace-temps de recueil de la mobilité quotidienne	68
Graphique 3.1 : Nombre de déplacements tous modes par jour et par personne	70
Graphique 3.2 : Nombre de déplacements voiture particulière par jour et par personne	71
Graphique 3.3 : Nombre de déplacements transports collectifs par jour et par personne	72
Graphique 3.4 : Nombre de déplacements marche à pied par jour et par personne	73
Graphique 3.5 : Nombre de déplacements deux roues par jour et par personne	74
Graphique 3.6 : Distance quotidienne par personne selon le mode	75
Carte 3.1 : Périmètre de l'enquête ménages déplacements et découpage en 7 zones	77
Graphique 3.7 : Part de marché des TC sur (TC+VP) pour les déplacements ayant une origine ou une destination dans la zone..., périmètre 77	79
Graphique 3.8 : Evolution des flux ayant une origine ou une destination dans la zone..., enquête ménages déplacements 85/ enquête ménages déplacements 77, périmètre 77 (% d'évolution)	80
Graphique 3.9 : Evolution des flux ayant une origine ou une destination dans la zone..., enquête ménages déplacements 95/ enquête ménages déplacements 85, périmètre 85 (% d'évolution)	80
Graphique 3.10 : Evolution de la motorisation selon la zone de résidence et la définition de la motorisation (périmètre 1977)	81
Graphique 3.11 : Evolution des temps généralisés voiture et transports collectifs de 1977 à 1995 pour les déplacements ayant une extrémité dans la zone ... (périmètre 1977)	83
Graphique 3.12 : Poids des personnes âgées dans la population française entre 1950 et 2040	89
Graphique 3.13 : Nombre de déplacements quotidiens en voiture comme conducteur selon l'âge	89
Graphique 3.14 : Evolution du nombre de déplacements quotidiens des personnes de 60 ans et plus	90
Graphique 3.15 : Evolution du nombre de déplacements quotidiens des personnes de 18 à 59 ans	91
Graphique 3.16 : Evolution du nombre de déplacements à pied des femmes par cohorte	92
Graphique 3.17 : Evolution du nombre de déplacements automobiles par cohorte	93
Graphique 3.18 : Evolution du nombre de déplacements automobiles par cohorte des conducteurs potentiels	93
Schéma 4.1 : Méthodologie de la production des données pour l'analyse	101
Schéma 4.2 : Méthodologie de la production des données pour la modélisation	104
Schéma 4.3 : Schéma de définition d'une enquête par questionnaire	105
Graphique 4.1 : Loi mobilité - motorisation dans le cadre des EPIT	140
Schéma 4.4 : Cadre conceptuel dans le cadre des EPIT	140
Schéma 4.5 : Cadre conceptuel prenant en compte la perception individuelle de l'offre	141

Schéma 4.6 : Cadre conceptuel considérant les déplacements comme demande dérivée des activités	142
Graphique 5.1 : Forme générale de la fonction d'utilité totale	146
Graphique 5.2 : Forme générale de la fonction d'utilité marginale	146
Graphique 5.3 : Illustration graphique de la courbe de demande en fonction du prix de X	148
Graphique 5.4 : Forme générale des courbes de coûts	148
Graphique 5.5 : Fonction d'offre J	150
Graphique 5.6 : Fonction de demande D	150
Graphique 5.7 : Equilibre offre-demande	151
Graphique 5.8 : Incidence d'une amélioration du système de transport	151
Schéma 5.1 : Approche agrégée	162
Schéma 5.2 : Approche agrégée catégorielle	162
Schéma 5.3 : Démarche agrégée / désagrégée	163
Schéma 6.1 : Génération des déplacements : émission et attraction par zone	166
Schéma 6.2 : Déplacement, chaîne, sortie	167
Schéma 6.3 : Emissions et attractions, définition pour le motif travail	171
Schéma 6.4 : Représentation du calcul des émissions et attractions pour les déplacements basés au domicile selon la méthode de calcul	172
Graphique 6.1 : Distribution générale de Y selon X donné	175
Graphique 6.2 : Hypothèses sur la distribution générale de Y selon X dans le modèle linéaire	176
Graphique 6.3 : Illustration de la décomposition des écarts et de l'interprétation graphique du coefficient R^2	180
Graphique 6.4 : Illustration du choix entre un modèle de régression unique ou un modèle catégoriel	182
Schéma 7.1 : Distribution des déplacements	192
Graphique 7.1 : Forme de la distribution des distances (temps généralisés) des déplacements	194
Graphique 7.2 : Formes de la courbe de résistance	202
Graphique 8.1 : Part de marché des TC sur (TC+VP) pour les déplacements ayant une origine ou une destination dans la zone..., périmètre 76	221
Schéma 8.1 : Exemple de grille de part modale pour la voiture pour le motif domicile-travail (extrait de CETE Méditerranée, CERTU 2000b)	227
Graphique 8.2 : Exemple de courbe de partage modal voiture-transports collectifs pour le motif domicile-travail (extrait de CETUR, 1990a)	227
Graphique 8.3 : Illustration de la forme générique des courbes de partage modal en S	230
Graphique 8.4 : Illustration de la forme générique des fonctions de densité logistique	231
Schéma 8.2 : Structures de modèles logit multinomiaux	234
Graphique 8.5 : Part de marché des modes 1 et 2 en fonction des valeurs des niveaux de service composites des modes 1 et 2	239
Graphique 9.1 : Fonction de densité uniforme	249
Graphique 9.2 : Fonction de probabilité du modèle de distribution linéaire	250
Graphique 9.3 : Fonction de distribution logit ($\mu = 1$) et probit ($\sigma = \frac{\pi}{\sqrt{3}}$)	254
Graphique 9.4 : Fonction de probabilité logit ($\mu = 1$) et probit ($\sigma = \frac{\pi}{\sqrt{3}}$)	255
Graphique 9.5 : Fonctions de densité de la loi de Gumbel (centrée, $\mu = 1$) et de la loi normale (centrée réduite)	257
Schéma 9.1 : Structure de modèle logit emboîté	263
Graphique 9.6 : Illustration de la fonction de vraisemblance	272
Graphique 9.7 : Illustration de la valeur minimum de ρ^2 relativement au modèle avec constante seule pour le logit binomial	281
Graphique 9.8 : Illustration du problème de l'agrégation en prévision dans les modèles désagrégés	296

Schéma 10.1 : Séquence des étapes lors du calage du modèle à quatre étapes	304
Schéma 10.2 : Séquence des étapes lors de l'utilisation du modèle à quatre étapes en prévision	305
Schéma 10.3 : Exemples de codification de réseau	314
Schéma 10.4 : Exemples de codification d'un carrefour	316
Schéma 10.5 : Exemples de codification de lignes de bus	317
Graphique 10.1 : Exemple de courbe débit-vitesse de forme parabolique (équation 10.1, pour une capacité de 1800 uvp/heure et une vitesse à vide de 90km/heure)	321
Graphique 10.2 : Exemple de courbe débit-temps (équation 10.1, pour une capacité de 1800 uvp/heure)	322
Graphique 10.3 : Courbe débit-coût du <i>Bureau of Public Roads</i> des Etats-Unis (1964, pour un coût à vide = 1)	323
Graphique 10.4 : Courbe débit-coût du logiciel DAVISUM (pour un coût à vide = 1)	324
Graphique 10.5 : Courbe débit-coût du logiciel CartOPERA (pour un coût à vide = 1)	324
Schéma 10.6 : Arbre de plus court chemin depuis l'origine h	331
Graphique 10.6 : Sélection des chemins efficaces dans le modèle prix-temps	340
Schéma 10.7 : Illustration de la sensibilité à la définition du réseau pour l'affectation de Dial	341
Graphique 10.7 : Illustration de l'équilibre selon le premier principe de Wardrop	345
Schéma 10.8 : Illustration du paradoxe de Braess	352
Graphique annexe 1.1 : Courbe de régression - modèle prix-temps avec l'ensemble des données	368

Liste des abréviations

ADEME	Agence De l'Environnement et de la Maîtrise de l'Energie
CAPI	<i>Computer-Aided Personal Interview</i> , enquête en face à face assistée par ordinateur
CATI	<i>Computer-Aided Telephone Interview</i> , enquête par téléphone assistée par ordinateur
CETE	Centre d'Etudes Techniques de l'Equipement
CETUR	Centre d'Etudes des Transports URbains
CERTU	Centre d'Etudes sur les Réseaux, les Transports, l'Urbanisme et les constructions publiques
CNIL	Commission Nationale Informatique et Liberté
CNL	<i>Cross-Nested Logit</i>
DDE	Direction Départementale de l'Equipement
DREIF	Direction Régionale de l'Equipement de l'Ile de France
DVA	Dossier de Voiries d'Agglomération
ENTPE	Ecole Nationale des Travaux Publics de l'Etat
EPIT	Etude Préliminaire d'Infrastructure de Transport
GART	Groupement des Autorités Responsables des Transports
GEV	<i>Generalised Extreme Value (models)</i> , (modèle de la) valeur extrême généralisée
GNL	<i>Generalised Nested Logit</i> , (modèle) logit emboîté généralisé
HB	<i>home-based trips</i> , déplacements basés au domicile
HL	<i>Hierarchical Logit</i> , (modèle) logit hiérarchique ou logit emboîté
HMNL	<i>Heteroscedastic Multinomial Logit</i> , (modèle) logit multinomial hétéroscédastique
IATBR	<i>International Association of Travel Behaviour Research</i> , Association Internationale de Recherche sur les Comportements de Déplacements
IAURP	Institut d'Aménagement et d'Urbanisme de la Région Parisienne
IIA	<i>Independence from Irrelevant Alternatives</i> , Indépendance vis-à-vis des autres alternatives
IID	<i>Independent and Identically Distributed</i> , distribution indépendante et de même écart-type
INRETS	Institut National de Recherche sur les Transports et leur Sécurité
INSEE	Institut National de la Statistique et des Etudes Economiques
LET	Laboratoire d'Economie des Transports
LOF	Loi d'Orientation Foncière
LOTI	Loi d'Orientation des Transports Intérieurs
MCO	méthode des Moindre Carrés Ordinaires
MMNL	<i>Mixed Multinomial Logit</i> , (modèle) logit multinomial mixte
MNL	<i>Multinomial Logit</i> , (modèle) logit multinomial
MNP	<i>multinomial probit</i> , (modèle) probit multinomial
MSA	<i>Method of Successive Averages</i> , méthode des moyennes successives
NHB	<i>non home-based trips</i> , déplacements secondaires
NL	<i>Nested Logit</i> , (modèle) logit emboîté
OCDE	Organisation pour la Coopération et le Développement Economique
OGEV	<i>Ordered GEV</i>
OTU	Omnium Technique Urbain
PCL	<i>Paired Combinatorial Logit</i>
PCS	Position et Catégorie Socio-professionnelle
PDU	Plan de Déplacements Urbains
POS	Plan d'Occupation des Sols (aujourd'hui baptisé PLU Plans Locaux d'Urbanisme)
PTU	Périmètre des Transports Urbains
RER	Réseau Express Régional
RP	<i>Revealed Preferences</i> , préférences révélées
RUM	<i>Random Utility Maximisation</i> , (théorie de la) maximisation de l'utilité aléatoire

SDAU	Schéma Directeur d'Aménagement et d'Urbanisme
SETRA	Services Techniques d'Etudes des Routes et Autoroutes
SP	<i>Stated Preferences</i> , préférences déclarées
SR	<i>Stated Response</i> , réponses déclarées
SRU	Solidarité et Renouvellement Urbains
SUE	<i>Stochastic User Equilibrium Assignment</i> , affectation à l'équilibre avec une approche probabiliste de l'utilité
TC	Transports Collectifs
TCSP	Transports Collectifs en Site Propre
UVP	Unité de Voiture Particulière
VP	Voiture Particulière
WCTR	<i>World Conference on Transport Research</i> , Conférence Mondiale sur la Recherche dans les Transports
ZAC	Zone d'Aménagement Concerté

Table des matières

Sommaire	3
Remerciements	5
Introduction	7
Chapitre 1 – Modélisation et planification	11
1. POURQUOI DOIT-ON PREVOIR ?	11
1.1. <i>Coût des investissements</i>	<i>11</i>
1.2. <i>Délai de construction et durée de vie des investissements importants</i>	<i>12</i>
1.3. <i>Anticipation des évolutions et des dysfonctionnements</i>	<i>12</i>
1.4. <i>Accompagnement par les transports du développement économique et social</i>	<i>13</i>
1.5. <i>Evolution rapide de certains indicateurs de mobilité</i>	<i>13</i>
2. HISTORIQUE DE LA MODELISATION	15
2.1. <i>Années 50 : les premiers modèles</i>	<i>15</i>
2.1.1. <i>Mise en place du processus de transfert des modèles américains</i>	<i>15</i>
2.1.2. <i>Premiers modèles : émergence d'une nouvelle problématique</i>	<i>16</i>
2.2. <i>Années 60 : l'acclimatation des modèles de trafic</i>	<i>17</i>
2.3. <i>Début des années 1970 : la mise en place d'une procédure de planification urbaine</i>	<i>19</i>
2.4. <i>1975 - 1980 : le tournant</i>	<i>19</i>
2.5. <i>Années 1980 à nos jours : vers une approche globale des problèmes de déplacements</i>	<i>21</i>
2.5.1. <i>Instruments de planification</i>	<i>21</i>
2.5.2. <i>Evolution des objectifs des politiques de déplacements urbains : conséquences pour la modélisation</i>	<i>23</i>
3. QUELS MODELES ? POUR QUELS DOMAINES D'UTILISATION ?	25
3.1. <i>Une décomposition du système urbain en sous-systèmes</i>	<i>26</i>
3.2. <i>Dimensions de l'exercice : espace-temps</i>	<i>27</i>
3.3. <i>Domaine d'utilisation des modèles à quatre étapes</i>	<i>29</i>
3.3.1. <i>Interactions faibles : modèle analogique, élasticité...</i>	<i>29</i>
3.3.2. <i>Interactions univoques : modèle à quatre étapes</i>	<i>30</i>
3.3.3. <i>Actions-rétroactions : modèle stratégique, transport-urbanisme...</i>	<i>31</i>
Chapitre 2 - Cadrage méthodologique	33
1. SCHEMA DE LA PROCEDURE DE PREVISION	33
1.1. <i>Phase 1 : objectifs, enjeux, définition du modèle</i>	<i>36</i>
1.2. <i>Phase 2 : recueil et analyse d'informations</i>	<i>38</i>
1.3. <i>Phase 3 : formalisation, calage, validation du modèle</i>	<i>39</i>
1.4. <i>Phase 4 : prévision</i>	<i>41</i>
2. LE MODELE : RESULTAT D'UN DOUBLE PROCESSUS DE SIMPLIFICATION	42
2.1. <i>Première simplification : la construction théorique</i>	<i>42</i>
2.2. <i>Deuxième simplification : la formalisation mathématique</i>	<i>44</i>
3. OPERATIONNALITE ET PROBLEMATIQUE DU MODELE	45
3.1. <i>Pertinence</i>	<i>46</i>
3.2. <i>Cohérence</i>	<i>46</i>

3.3. <i>Mesurabilité</i>	47
3.4. <i>Problématique du modèle</i>	48
3.5. <i>Classification des erreurs de modélisation</i>	50
4. EXEMPLE : CALAGE DU MODELE PRIX-TEMPS	51
4.1. <i>Présentation du modèle prix-temps et du calage du modèle</i>	51
4.1.1. <i>Principe du modèle</i>	51
4.1.2. <i>Ajustement du modèle</i>	54
4.2. <i>Présentation des phases du schéma de procédure de prévision</i>	55
4.2.1. <i>Phase 1 : objectifs, enjeux, définition du modèle</i>	55
4.2.2. <i>Phase 2 : recueil et analyse d'informations</i>	57
4.2.3. <i>Phase 3 : formalisation, calage, validation du modèle</i>	57
4.2.4. <i>Phase 4 : prévision</i>	60
4.3. <i>Le modèle : résultat d'un double processus de simplification</i>	61
4.3.1. <i>Première simplification : la construction théorique</i>	61
4.3.2. <i>Deuxième simplification : la formalisation mathématique</i>	61
4.4. <i>Opérationnalité et problématique du modèle</i>	62
4.4.1. <i>Pertinence</i>	62
4.4.2. <i>Cohérence</i>	62
4.4.3. <i>Mesurabilité</i>	62
4.4.4. <i>Problématique du modèle</i>	63
5. CONCLUSION	64
Chapitre 3 - Analyse de la mobilité	67
1. MOBILITE : UNE DEFINITION EN EVOLUTION	67
2. DESCRIPTION DE LA MOBILITE ET DE SON EVOLUTION	69
2.1. <i>Une absence de tendance claire pour le nombre de déplacements quotidiens...</i>	69
2.2. <i>... Mais une forte croissance de l'usage de la voiture...</i>	71
2.3. <i>... Un accroissement plus modéré pour les transports collectifs...</i>	72
2.4. <i>... Un effondrement de la marche...</i>	73
2.5. <i>... Et une quasi-disparition des deux-roues ?</i>	73
2.7. <i>Choix de l'indicateur de description de la mobilité</i>	74
3. TENDANCES STRUCTURELLES D'EVOLUTION DE LA MOBILITE : EXEMPLE D'ANALYSE, EXEMPLE DE QUANTIFICATION	75
3.1. <i>Quantification des effets de l'évolution de la localisation des déplacements, de la motorisation et de l'offre de transport sur la part de marché des transports collectifs</i>	76
3.1.1. <i>Principe de mesure et terrain d'étude</i>	76
3.1.2. <i>Evolution et définition des facteurs étudiés</i>	78
3.1.2.1. <i>Eclatement des flux et définition du facteur localisation</i>	79
3.1.2.2. <i>Facteur motorisation</i>	80
3.1.2.3. <i>Offre de transport</i>	82
3.1.2.4. <i>« Autres facteurs »</i>	83
3.1.3. <i>Choix d'une fonction de partage modal</i>	84
3.1.4. <i>Des effets simples importants</i>	86
3.1.5. <i>Conclusion : précautions méthodologiques</i>	87
3.2. <i>Vieillesse de la population : perspective d'évolution de la mobilité automobile des personnes âgées</i>	88
3.2.1. <i>Vieillesse de la population... solution des problèmes de congestion ?...</i>	88
3.2.2. <i>... ou amplificateur de la congestion ?... dû à l'usage de la voiture par les hommes ?...</i>	89
3.2.3. <i>... A moins que ce soit à cause des femmes ?...</i>	90
3.2.4. <i>... Ou les pièges de l'analyse de la mobilité</i>	91
3.2.5. <i>Effet d'âge ou de génération ?</i>	91

3.2.6. <i>Séparation et quantification des effets</i>	94
4. CONCLUSION : LES PIEGES DE L'ANALYSE DE LA MOBILITE	96
4.1. <i>Corrélation ou causalité</i>	96
4.2. <i>Dépendance entre les facteurs « explicatifs »</i>	97
4.3. <i>Prise en compte du temps</i>	97
4.4. <i>Quelle mesure ?</i>	97
Chapitre 4 - Production de données	99
1. POURQUOI A-T-ON BESOIN DE DONNEES ?	99
2. METHODOLOGIE DE PRODUCTION DES DONNEES	101
2.1. <i>Méthodologie de la production des données pour l'analyse</i>	101
2.2. <i>Méthodologie de la production des données pour la modélisation</i>	103
2.3. <i>Définition de la méthode d'enquête</i>	104
2.3.1. <i>Définition de l'objectif</i>	105
2.3.2. <i>Définition de la population cible</i>	105
2.3.3. <i>Définition de la méthode d'échantillonnage</i>	106
2.3.4. <i>Définition du mode d'enquête</i>	108
2.3.5. <i>Rédaction du questionnaire</i>	109
2.3.6. <i>Organisation du terrain</i>	109
2.3.7. <i>Codage, saisie des données</i>	110
2.3.8. <i>Définition du traitement des données</i>	110
2.4. <i>Précision statistique</i>	110
2.4.1. <i>Non-réponse</i>	111
2.4.2. <i>Calcul de la précision</i>	112
2.4.2.1. <i>Sondage aléatoire simple</i>	113
2.4.2.2. <i>Sondage stratifié</i>	114
2.4.2.3. <i>Effet de grappe</i>	115
3. PRINCIPALES DONNEES DISPONIBLES EN MILIEU URBAIN EN FRANCE	116
3.1. <i>Enquêtes ménages déplacements</i>	117
3.1.1. <i>Définition de l'objectif</i>	117
3.1.2. <i>Définition de la population cible</i>	119
3.1.3. <i>Définition de la méthode d'échantillonnage</i>	119
3.1.4. <i>Définition du mode d'enquête</i>	120
3.1.5. <i>Rédaction du questionnaire</i>	120
3.1.6. <i>Organisation du terrain</i>	120
3.1.7. <i>Codage, saisie des données</i>	121
3.2. <i>Enquête Cordon et bord de route</i>	121
3.2.1. <i>Définition de l'objectif</i>	121
3.2.2. <i>Définition de la population cible</i>	122
3.2.3. <i>Définition de la méthode d'échantillonnage</i>	122
3.2.4. <i>Définition du mode d'enquête</i>	122
3.2.5. <i>Rédaction du questionnaire</i>	122
3.2.6. <i>Organisation du terrain</i>	122
3.2.7. <i>Codage, saisie des données</i>	123
3.3. <i>Comptages routiers</i>	123
3.3.1. <i>Définition de l'objectif</i>	123
3.3.2. <i>Définition de la population cible</i>	123
3.3.3. <i>Définition de la méthode d'échantillonnage</i>	123
3.4. <i>Bases de données transports collectifs</i>	124
3.5. <i>Recensements de l'INSEE</i>	124
3.6. <i>Conclusion</i>	125
4. LIMITES DU DISPOSITIF FRANÇAIS D'ENQUETES SUR LES DEPLACEMENTS URBAINS POUR LA MODELISATION	125

4.1. <i>Des déplacements exclus du dispositif d'enquête</i>	125
4.2. <i>Des périmètres d'enquêtes différents</i>	126
4.3. <i>Des dates d'enquêtes très différentes</i>	126
4.4. <i>Des découpages différents et un codage trop grossier des localisations</i>	127
4.5. <i>Une taille d'échantillon trop faible pour l'enquête ménages déplacements</i>	128
4.6. <i>Proposition pour une enquête ménages déplacements téléphonique « allégée »</i>	129
4.6.1. <i>Objectifs : un choix nécessaire</i>	130
4.6.2. <i>Un questionnaire réduit</i>	130
4.6.3. <i>Une enquête téléphonique</i>	131
4.6.4. <i>Un périmètre élargi</i>	132
4.6.5. <i>Une codification plus fine des adresses</i>	132
4.6.6. <i>Mais un dispositif à tester</i>	133
5. LES ENQUETES DE PREFERENCES DECLAREES	134
5.1. <i>Principaux domaines d'applications : situations nouvelles et estimations de certains indicateurs</i>	135
5.2. <i>Classification des approches de réponses déclarées (SR, « Stated Response »)</i>	135
5.3. <i>Enquêtes de préférences déclarées</i>	136
5.4. <i>Précautions méthodologiques</i>	137
6. EXEMPLE D'APPLICATION DU SCHEMA METHODOLOGIQUE DE PRODUCTION DES DONNEES : EXEMPLE DE LA CONSTRUCTION DU CADRE CONCEPTUEL	139
6.1. <i>Premiers modèles dans le cadre des EPIT</i>	139
6.2. <i>Enrichissement du cadre conceptuel</i>	140
7. CONCLUSION	142

Chapitre 5 - Fondements théoriques des modèles de prévision de la demande de déplacements 145

1. QUELQUES ELEMENTS DE THEORIE MICRO-ECONOMIQUE NEO-CLASSIQUE DU CHOIX DU CONSOMMATEUR	145
1.1. <i>Analyse du comportement du consommateur</i>	145
1.2. <i>Optimum du consommateur</i>	147
1.3. <i>Construction de la courbe de demande</i>	147
1.4. <i>Analyse du producteur</i>	148
1.5. <i>Généralisation de l'analyse : l'équilibre</i>	149
2. APPLICATION A LA SPHERE DES TRANSPORTS	149
2.1. <i>Equilibre offre-demande de Manheim</i>	149
2.2. <i>Interprétation au sein du système urbain</i>	151
2.3. <i>Formalisation de la fonction d'offre : le coût généralisé</i>	152
2.4. <i>Formalisation de la fonction de demande</i>	154
2.4.1. <i>Approche déterministe ou stochastique</i>	156
2.4.1.1. <i>Approche probabiliste avec utilité constante</i>	157
2.4.1.2. <i>Approche probabiliste avec utilité aléatoire</i>	157
2.4.2. <i>Interprétation du terme aléatoire ε</i>	158
2.5. <i>Réflexions sur l'équilibre offre-demande appliqué à la modélisation de la demande</i>	159
3. CLASSIFICATION DES « MODELES A QUATRE ETAPES »	160
3.1. <i>Modèle direct - séquentiel</i>	160
3.2. <i>Modèle agrégé - désagrégé</i>	162
3.3. <i>Modèle déterministe - stochastique</i>	163
3.4. <i>Combinaison des dimensions</i>	163
4. CONCLUSION	164

Chapitre 6 – Génération de la demande de déplacements	165
1. DEFINITIONS ET CHOIX DE L'UNITE DE MESURE	165
1.1. <i>Emission, Attraction</i>	165
1.2. <i>Déplacements, activités, chaînes, sorties, schéma d'activités</i>	166
1.3. <i>Jour ou période de pointe</i>	168
1.4. <i>Déplacements des individus ou des ménages</i>	168
1.5. <i>Segmentation par motifs</i>	169
1.6. <i>Principales variables explicatives de la génération des déplacements</i>	172
2. PRINCIPALES FAMILLES DE MODELES DE GENERATION	173
2.1. <i>Modèles normatifs</i>	173
2.2. <i>Modèles par régression linéaire</i>	174
2.2.1. <i>Hypothèses du modèle linéaire général</i>	175
2.2.2. <i>Estimation de β</i>	176
2.2.3. <i>Estimation de α</i>	178
2.2.4. <i>Test F sur le modèle complet</i>	178
2.2.5. <i>Coefficient de détermination R^2</i>	179
2.2.6. <i>Prévision de Y pour une valeur donnée de X</i>	180
2.2.7. <i>Régression multiple</i>	181
2.3. <i>Modèle catégoriel</i>	182
3. PRINCIPALES LIMITES DE L'ETAPE DE GENERATION	183
3.1. <i>Hypothèses du modèle linéaire</i>	184
3.1.1. <i>Linéarité</i>	184
3.1.2. <i>Homoscédasticité</i>	184
3.1.3. <i>Multicolinéarité</i>	185
3.2. <i>Divergence entre les émissions et les attractions</i>	185
3.3. <i>Indépendance vis-à-vis de l'offre de transport et des autres étapes</i>	185
3.4. <i>Stabilité des paramètres de génération</i>	187
3.5. <i>Choix et prévision des variables explicatives</i>	187
3.6. <i>Qualité des données</i>	188
Chapitre 7 – Distribution des déplacements	191
1. DEFINITIONS ET CONCEPTS DE BASE	191
1.1. <i>Notation</i>	191
1.2. <i>Contraintes aux marges</i>	192
1.3. <i>Coût généralisé, temps généralisé</i>	193
1.4. <i>Répartition des distances de déplacements</i>	193
1.5. <i>Echanges et transit</i>	194
2. PRINCIPALES FAMILLES DE MODELES DE DISTRIBUTION	194
2.1. <i>Modèles à facteurs de croissance</i>	195
2.1.1. <i>Facteur uniforme</i>	195
2.1.2. <i>Définition du facteur de croissance</i>	195
2.1.3. <i>Méthode de Furness</i>	196
2.1.4. <i>Méthode de Fratar</i>	198
2.1.5. <i>Avantages et limites des facteurs de croissance</i>	199
2.2. <i>Modèles gravitaires</i>	199
2.2.1. <i>Principe</i>	199
2.2.2. <i>Transposition à l'urbain</i>	200
2.2.3. <i>Fonction de résistance ...</i>	201
2.2.4. <i>... ou fonction de demande</i>	203
2.2.5. <i>Justification économique du modèle gravitaire</i>	203
2.2.6. <i>Calage du modèle gravitaire</i>	205
2.3. <i>Maximisation de l'entropie</i>	207

2.3.1. Principe général	207
2.3.2. Génération du modèle gravitaire	208
2.4. Modèles d'opportunité	210
3. PRINCIPALES LIMITES DE L'ETAPE DE DISTRIBUTION	213
3.1. Déplacements intrazonnes	213
3.2. Pertinence de la distance comme critère de choix de la destination	214
3.3. Stabilité des coefficients	215
3.4. Segmentation	215
3.5. Zones externes	216
3.6. Cellules nulles	216
3.7. Quelle valeur pour le coût généralisé ?	216
3.8. Ordre des étapes et cohérence avec le choix du mode	217
3.9. Qualité des données	218
Chapitre 8 – Répartition modale – approche agrégée	219
1. DEFINITION ET VARIABLES EXPLICATIVES	219
1.1. Choix du mode : une étape négligée en France	219
1.2. Variables explicatives du choix du mode	220
1.2.1. Déterminants liés à l'offre de transport	220
1.2.2. Déterminants liés aux caractéristiques du déplacement	221
1.2.3. Déterminants liés aux caractéristiques individuelles	222
1.3. Coût généralisé de la période de pointe ou de la journée ?	223
1.4. Quelle segmentation ?	224
1.5. Choix du mode : déplacements ou sorties	225
2. PRINCIPALES FAMILLES DE MODELES DE REPARTITION MODALE	226
2.1. Modèles normatifs	226
2.2. Modèles logit ou régression logistique	227
2.2.1. Fondements théoriques	228
2.2.2. Distribution logitisque	229
2.2.3. Logit multinomial et propriété IIA	233
2.2.4. Logit emboîté (nested logit)	234
2.2.5. Calage des modèles logit	235
2.3. Modèles combinés distribution – répartition modale	236
2.4. Modèles directs	237
3. PRINCIPALES LIMITES DE L'ETAPE DE REPARTITION MODALE	240
3.1. Pertinence des variables explicatives du choix du mode	240
3.2. Prise en compte des problématiques de planification actuelles	241
3.3. Formalisation de la fonction d'utilité	241
3.4. « Choice axiom » et propriété IIA	242
3.5. Stabilité des coefficients de calage	242
3.6. Quelle segmentation de la demande	242
3.7. Quelle valeur pour le coût généralisé ?	242
3.8. Qualité des données	243
Chapitre 9 - Modèles désagrégés de choix discrets	245
1. FONDEMENTS THEORIQUES	246
1.1. Utilité aléatoire	246
1.2. Choix discrets	247
2. PRINCIPALES FAMILLES DE MODELES DESAGREGES	248
2.1. Modèle de probabilité linéaire	249
2.2. Formulation générale	250
2.3. Modèle probit	251
2.4. Modèle logit	252

2.4.1. Hypothèses du modèle logit	253
2.4.2. Formulation du modèle logit	254
2.4.3. Loi de Gumbel	255
2.4.4. Démonstration du modèle logit	257
2.4.5. Propriété IIA d'indépendance vis-à-vis des autres alternatives	258
2.5. Modèle logit emboîté (Nested Logit (NL))	259
2.5.1. Hypothèses du logit emboîté	259
2.5.2. Formulation du logit emboîté	261
2.5.3. Application du logit emboîté au choix d'alternatives partageant des dimensions communes	263
2.6. Modèles de la valeur extrême généralisée (Generalised Extreme Value models, GEV)	265
2.6.1. Principe, formalisation et démonstration	265
2.6.2. Dérivation du logit multinomial	267
2.6.3. Dérivation du logit emboîté	267
2.6.4. Famille des modèles de la valeur extrême généralisée	268
2.7. Développements récents : modèle logit multinomial mixte (Mixed Multinomial Logit (MMNL))	268
2.7.1. Structure spécifique du terme d'erreur	269
2.7.2. Distribution aléatoire des coefficients de l'utilité	270
3. ESTIMATION DU MODELE DESAGREGE	270
3.1. Maximum de vraisemblance	271
3.2. Tests statistiques	274
3.2.1. Signe des coefficients et élasticités	274
3.2.2. Quasi-t test	278
3.2.3. Intervalle de confiance sur les coefficients	279
3.2.4. Qualité de l'estimation : ρ^2	279
3.2.5. Test du ratio de vraisemblance (Likelihood ratio Test ou LR)	281
3.2.6. Test de la propriété IIA	283
3.2.7. Test de variation de perception (test of taste variation)	284
3.2.8. Conclusion	286
4. SPECIFICATION DES MODELES DESAGREGES	286
4.1. Choix des alternatives	287
4.2. Choix des variables, de leur spécification et de leur forme fonctionnelle	288
4.2.1. Détermination des constantes	288
4.2.2. Variables spécifiques ou génériques	288
4.2.3. Forme fonctionnelle et transformations de variables	289
4.3. Structure du modèle	291
5. PREVISION ET AGREGATION	293
5.1. Prévision et biais d'agrégation	294
5.2. Méthodes d'agrégation	295
6. PRINCIPALES LIMITES DE L'APPROCHE DESAGREGEE	297
6.1. Approche agrégée ou désagrégée	297
6.1.1. Pertinence des variables explicatives du choix du mode	297
6.1.2. Prise en compte des problématiques de planification actuelles	298
6.1.3. Formalisation de la fonction d'utilité	298
6.1.4. « Choice axiom » et propriété IIA	298
6.1.5. Stabilité des coefficients de calage	298
6.1.6. Quelle segmentation de la demande ?	298
6.1.7. Quelle valeur pour le coût généralisé ?	299
6.1.8. Qualité des données	299
6.2. « Taste variation », variation de perception	299
6.3. Approche désagrégée et choix du zonage	299

6.4. <i>Quelle structure de modèle ?</i>	300
6.5. <i>Prévision</i>	301
Chapitre 10 – Affectation, zonage et codification des réseaux	303
1. POSITIONNEMENT DE L’AFFECTION DANS LA CONSTRUCTION ET L’UTILISATION D’UN MODELE A QUATRE ETAPES	303
2. PERIMETRE ET ZONAGE DE L’AIRE D’ETUDE	307
2.1. <i>Périmètre d’étude</i>	307
2.2. <i>Zonage de l’aire d’étude</i>	308
2.2.1. <i>Prises en compte des objectifs de l’étude</i>	308
2.2.2. <i>Cohérence interne : homogénéité des zones, prises en compte des frontières « naturelles »...</i>	309
2.2.3. <i>Prise en compte des réseaux</i>	309
2.2.4. <i>Disponibilité des données, compatibilité des découpages...</i>	310
2.2.5. <i>Coût et temps de calcul.</i>	311
3. CODIFICATION DES RESEAUX	312
3.1. <i>Niveau d’agrégation</i>	312
3.2. <i>Caractéristiques des liens et des nœuds</i>	313
3.2.1. <i>Définition des éléments composant un réseau</i>	313
3.2.2. <i>Attributs des liens</i>	315
3.3. <i>Codification des lignes de transports collectifs</i>	316
4. CONCEPTS DE BASE ET ETAPES PRELIMINAIRES	319
4.1. <i>Coût d’un itinéraire et notations</i>	319
4.2. <i>Courbe débit-vitesse ou débit-coût</i>	321
4.3. <i>Coût généralisé</i>	325
4.4. <i>Conversion véhiculaire et passage à la pointe</i>	327
4.4.1. <i>Conversion véhiculaire</i>	327
4.4.2. <i>Passage à la période de pointe</i>	328
5. PRINCIPALES FAMILLES DE MODELES D’AFFECTION	328
5.1. <i>Classification des modèles d’affectation</i>	328
5.2. <i>Affectation tout-ou-rien ou plus court chemin</i>	329
Exemple d’algorithme de plus court chemin	330
5.3. <i>Affectation stochastique ou multichemin</i>	332
5.3.1. <i>Fondements théoriques de l’affectation stochastique</i>	332
5.3.2. <i>Affectation multichemin de Dial</i>	333
Exemple d’algorithme pour l’affectation multichemin de Dial	334
5.3.3. <i>Affectation multichemin de Burrell</i>	336
Exemple d’algorithme pour l’affectation multichemin de Burrell	336
5.3.4. <i>Méthode d’Abraham ou méthode Ariane : logit logarithmique</i>	337
5.3.5. <i>Affectation prix-temps</i>	339
5.3.6. <i>Principales limites de l’affectation stochastique ou multichemin</i>	340
5.4. <i>Affectation à l’équilibre ou avec contrainte de capacité</i>	342
5.4.1. <i>Affectation à l’équilibre de Wardrop</i>	342
5.4.1.1. <i>Principes généraux de l’affectation selon le premier principe de Wardrop</i>	343
5.4.1.2. <i>Un algorithme de résolution de l’équilibre selon le premier principe de Wardrop : l’algorithme de Frank-Wolfe</i>	345
5.4.1.3. <i>Principe de l’algorithme fondé sur la procédure d’égalisation par transvasement</i>	348
5.4.1.4. <i>Affectation prix-temps selon le premier principe de Wardrop</i>	348
5.4.1.5. <i>Second principe de Wardrop</i>	349
5.4.1.6. <i>Paradoxes de Braess et de Villé</i>	351
5.4.2. <i>Affectation par tranche ou incrémentale</i>	352
Exemple d’algorithme pour l’affectation par tranche ou incrémental	352

5.5. <i>Affectation à l'équilibre avec utilité stochastique</i>	354
Exemple d'algorithme pour la méthode des moyennes successives	354
6. PRINCIPALES LIMITES DE L'ETAPE D'AFFECTION	355
6.1. <i>Calage et validation</i>	356
6.2. <i>Zonage</i>	359
6.3. <i>Définition des réseaux</i>	359
6.4. <i>Hypothèses des modèles d'affectation</i>	359
6.5. <i>L'équilibre dans l'affectation</i>	360
6.6. <i>Convergence des modèles</i>	360
6.7. <i>Prévision</i>	361
Annexe 1 - Données nécessaires au calage du modèle « Prix-temps » et corrigé	362
1. TRAFICS EN 1996	362
2. TEMPS DE PARCOURS EN MINUTES EN 1996	363
3. TARIFS DE BASE AUTOMNE-HIVER 1996 (CONVERTI EN EQUIVALENT EUROS 96)	363
4. CORRIGE DE L'EXERCICE DE CALAGE DU MODELE PRIX-TEMPS - ENSEMBLE DES DONNEES	364
Bibliographie	371
Liste des tableaux	393
Liste des illustrations	395
Liste des abréviations	399
Table des matières	401