

Contrat d'Objectifs Etat-IAURIF 1998

RECHERCHE SUR
LE CHOIX MODAL
EN MILIEU URBAIN

RAPPORT FINAL

Juin 2000

D5.336

Contrat d'Objectifs Etat-IAURIF 1998

RECHERCHE SUR LE CHOIX MODAL EN MILIEU URBAIN

Institut d'Aménagement et d'Urbanisme de la Région d'Ile-de-France

15, rue Falguière- 75740 PARIS CEDEX 15 - ☎: 01-53-85-76-02

Directeur général : Jean-Pierre DUFAY

Division Transports et Infrastructures - Directeur : Joseph BERTHET

Etude réalisée par :

Dany NGUYEN-LUONG, chargé d'études

email : nguyen-luong@iaurif.org

SOMMAIRE

1. Introduction	4
2. Les modèles de choix désagrégés	6
2.1. Bases théoriques	
2.2. Hypothèses de travail	
3. Constitution d'une base de données	9
3.1. Reconstitution des temps généralisés	
3.2. Structure des motifs à l'heure de pointe du matin	
3.3. Constitution de la base de données extraite de l'EGT98 pour le motif domicile->travail	
3.4. Constitution de la base de données extraite de l'EGT98 pour le motif domicile->école	
3.5. Constitution de la base de données extraite de l'EGT98 pour le motif domicile->autres motifs	
4. Estimation du modèle désagrégé	26
4.1. Motif domicile->travail	
4.2. Motif domicile->école	
4.3. Motif domicile->autres motifs	
5. Validation du modèle retenu	43
6. Conclusion	47
Annexes	49
1. Contexte de l'étude en Ile-de-France	
2. Points essentiels de la partie "Prévisions de trafic à la mise en service" (§III11) de l'expertise du CGPC sur la tangentielle nord (3 décembre 1999)	
3. Cartes du zonage du modèle global de l'IAURIF et du découpage morphologique en 5 zones	
4. Quelques modèles de choix modal aux Etats-Unis	
5. Parts modales des lignes de rocade existantes (sources : enquêtes clientèles RATP et SNCF)	
Bibliographie	72

1. INTRODUCTION

Le problème de la modélisation du choix modal en Ile-de-France se pose de manière récurrente. On sait depuis longtemps expliciter les nombreux facteurs tant quantitatifs que qualitatifs qui pourraient déterminer le choix du mode (sur cet aspect le milieu de la recherche a sans doute épuisé le sujet), mais lorsqu'il s'agit de passer à la simulation, aucun modèle existant ne donne réellement satisfaction.

Il y a actuellement trois principaux sujets d'insatisfaction en Ile-de-France :

- les modèles de choix modal existants donnent des résultats presque toujours différents pour un même projet d'infrastructure. Depuis trois ans, ce sont les études de trafic sur les tangentiels ferrés qui focalisent l'attention sur l'étape du choix modal. Dans le cas de la tangentielle nord, cela a été particulièrement évident : les reports modaux vont du simple au double, comme il est rappelé dans l'annexe 2, illustrant ainsi la fragilité des résultats. Finalement, c'est moins l'imperfection d'un modèle qui pose problème que les différences de résultats entre les différents modèles.
- les modèles existants manquent complètement de transparence. Il n'existe pas de document technique clair et complet qui explique *comment* ont été élaborées les courbes de partage modal des différents organismes en charge d'études de trafic, que ce soient les modèles de la DREIF, la SNCF, la RATP, le STP ou l'IAURIF.
- les modèles de choix modal actuellement utilisés sont difficilement transposables d'une étude à l'autre pour deux raisons : soit on s'entend dire que le modèle n'est pas actualisé car il a été calé sur des enquêtes obsolètes (EGT83 ou 91) et sur le fonctionnement des lignes radiales pour les transports en commun, soit on s'entend dire que le modèle a été conçu et développé pour un secteur d'étude donné dans un découpage spécifique et donc qu'il n'est pas applicable immédiatement à un autre secteur.

La présente étude a tenu compte des trois points précédents pour l'élaboration de la méthodologie. Elle n'a pas pour objectif de produire un rapport de recherche de plus : tout a déjà été dit et répété sur le sujet. L'étude doit être pratique et tenter de mettre d'accord les organismes en charge des évaluations de projets d'infrastructure en Ile-de-France sur un

modèle de choix modal utilisable dans des études d'une nouvelle offre de transport en commun lourd, même s'il ne prend pas en compte tous les facteurs imaginables.

Ainsi, l'étude essaie de rendre complètement transparente la démarche de modélisation. Elle s'appuie sur la dernière enquête ménages, l'EGT 97-98, dont les fichiers (version ultime) sont disponibles depuis le 15 février 2000. Le modèle utilisé est un modèle de choix discret ou modèle désagrégé qui sera estimé à partir des observations de comportement individuel de cette enquête. Ce type de modèle permet en particulier de s'affranchir des variables zones origine et destination des déplacements, c'est-à-dire du découpage en zones du secteur d'étude.

Il est probablement utopique, compte tenu de la complexité du contexte francilien (cf annexe 1) de prétendre proposer un modèle de choix modal qui s'applique à toutes les études à venir d'infrastructures de transport en Ile-de-France. L'étude se limitera donc raisonnablement à l'objectif suivant : ouvrir un champ d'études pratiques sur le choix modal désagrégé en Ile-de-France qui s'appuie sur l'Enquête Globale Transports. Elle peut être considérée comme un premier guide méthodologique pour estimer un modèle désagrégé de choix modal dans le cadre d'un modèle global de prévision de trafic.

Dans le cadre du dernier appel à propositions du PREDIT (février 2000) sur la modélisation des flux de transports, une étude similaire plus approfondie doit être réalisée sur l'estimation d'un modèle désagrégé en Ile-de-France à partir de l'EGT 97-98. Elle devrait constituer la suite logique de la présente étude.

Ce rapport final ¹ se décompose en quatre parties :

- un rappel des principes de base sur les modèles désagrégés de choix modal et l'exposé des hypothèses de travail,
- une description de la démarche de constitution de la base d'observations individuelles à partir de laquelle sera estimé le modèle désagrégé,
- les résultats de l'estimation proprement dite réalisée avec le logiciel de statistiques SAS,
- la validation du modèle sur une ligne existante de rocade en transport en commun.

¹ Deux notes d'avancement ont été rendues à la DREIF le 31 août 1999 et le 16 mars 2000.

2. LES MODELES DE CHOIX DESAGREGES

Aux Etats-Unis, dans le domaine de la modélisation du choix modal, les vingt dernières années ont vu le passage du modèle de choix discret de l'état de l'art à celui de la pratique. Le modèle standard est actuellement le modèle logit multinomial. On trouvera en annexe 4 des exemples de modélisation du choix modal dans quelques métropoles américaines. Or en Ile-de-France, la plupart des modèles de choix modal sont des modèles dits "**agrégés**", c'est-à-dire basés sur des données de zone à zone dans un découpage donné. Les courbes de choix modal (en fonction du motif de déplacement, du type de zone de résidence, ...) sont estimées au moyen de pourcentages de répartition modale entre un certain nombre de couples de zones origine/destination. Par contre, un modèle dit "**désagrégé**" est basé sur les observations de déplacements déclarés par les individus. Les paramètres du modèle sont estimés directement sur une base d'observations extraite d'une enquête ménage, indépendamment de tout découpage en zones.

A l'image de ce qui est devenu usuel aux Etats-Unis, nous avons donc opté dans le cadre de cette étude pour un modèle désagrégé de choix modal à partir de la dernière enquête globale transports, l'EGT 97-98. Il faut dire que, si le modèle désagrégé a nettement fait une percée aux Etats-Unis dans certaines métropoles, les modélisateurs américains sont confrontés aux mêmes problèmes que tout le monde : la disponibilité des données et le manque d'expérience des praticiens.

2.1. Bases théoriques

On peut rappeler succinctement les bases théoriques d'un modèle de choix discret. Pour en savoir plus, on pourra se reporter à la littérature adéquate, en particulier à l'ouvrage de référence [1] « Discrete choice analysis » (1985) du Professeur Ben Akiva. C'est la mise en œuvre et la validation de ce modèle qui sont autrement difficiles.

Le modèle logit bimodal est la forme mathématique la plus utilisée dans les modèles de choix discrets. Il est basé sur l'hypothèse qu'un individu m associe une utilité $U_{i,m}$ à chacun des modes i principaux (Véhicule Particulier ou Transport en Commun) et qu'il choisira alors le mode qui lui procure la plus grande utilité. Cette utilité se décompose en une utilité déterministe u_i et en un terme aléatoire non observable ε_i , censé prendre en compte les erreurs de mesure relatives à la partie déterministe de l'utilité et les facteurs non observés.

On peut donc exprimer l'utilité stochastique par :

$$U_{i,m} = u_{i,m} + \varepsilon_{i,m} = \beta' X_i + \varepsilon_{i,m} \quad , \quad i \in \{VP, TC\}$$

où X_i représente le vecteur des *caractéristiques* du choix du mode i et β représente le vecteur des coefficients à estimer.

En posant $V_{i,m} = 1_{U_{i,m} > 0}$, on a :

$$P_{i,m} = \text{Prob}_m(V_i=1) = \text{Prob}_m(\beta' X_i > -\varepsilon_{i,m}) = F_m(\beta' X_i), \quad F \text{ étant une fonction de répartition de } -\varepsilon_{i,m} \text{ avec } F(w) = \frac{\exp(w)}{1 + \exp(w)}.$$

La probabilité pour l'individu m de choisir le mode i est alors donnée par la formule :

$$P_{i,m} = \frac{\exp(U_{i,m})}{\exp(U_{vp,m}) + \exp(U_{tc,m})}$$

Il faut remarquer que le qualificatif "désagrégé" est parfois utilisé de façon abusive. Un *vrai* modèle désagrégé, une fois qu'il a été calé, devrait s'appliquer directement sur une base d'observations futures (obtenue par exemple par un synthétiseur de population et de déplacements). En pratique, le modèle désagrégé est le plus souvent appliqué de zone à zone. Dans ce cas, il n'estime plus la probabilité de choix d'un individu mais une part de marché de zone à zone. Le modèle apparaît sous la forme de formules mathématiques de type tangente hyperbolique ("courbe en S").

2.2. Hypothèses de travail

- On ne considèrera que deux modes, notés VP (véhicule particulier) et TC (transport en commun).
- Les modèles désagrégés décrits par MM. Ben Akiva et Lerman dans l'ouvrage de référence [1] prennent en compte de nombreuses variables liées aux caractéristiques des individus et des ménages comme le revenu, l'âge, le sexe, la possession du permis de conduire, ... Le problème qui se pose ensuite lorsqu'on veut appliquer ce modèle à un horizon futur est de faire des projections de population avec ces caractéristiques. On pourrait utiliser pour cela un programme de génération de population mais on sait

combien il est difficile à cause des nombreuses hypothèses à prendre et des discussions interminables qui s'ensuivent. Par contre, à un horizon futur, les variables caractérisant l'offre de transport sont bien connues. On supposera dans un premier temps que l'utilité pour chaque mode est donnée essentiellement par des variables d'offre caractérisant le niveau de service ², ce qui ne nous empêchera pas par la suite d'introduire dans la spécification des utilités quelques caractéristiques socio-économiques de l'individu a priori importantes tels que le taux de motorisation et la disposition d'un parking gratuit au lieu de destination.

- On introduira la notion de captivité d'un mode. Des expériences passées de la DREIF ont montré que la non-distinction au départ entre les captifs et les non-captifs fausse les courbes de partage modal au niveau des asymptotes.
- Le temps de parcours généralisé pour une liaison donnée est une combinaison linéaire des différents temps le constituant . Ces différents temps seront donc les *caractéristiques* de l'utilité :
 - pour un déplacement en TC : temps de rabattement au départ et à l'arrivée, temps de correspondance si nécessaire, temps d'attente, temps de roulement. Une autre variable est considérée a priori comme pertinente : le nombre de correspondances.
 - pour un déplacement en VP : temps de parcours (issus d'une affectation routière à contrainte de capacité à l'heure de pointe), temps de recherche de parking à l'arrivée.
- On introduira dans un deuxième temps des variables de coûts monétaires de déplacement pour passer des temps généralisés aux coûts généralisés. Il est souvent dit et répété que le coût monétaire n'intervient pas dans le choix du mode en milieu urbain mais qu'il est déterminant en milieu interurbain. Il est intéressant de vérifier cette affirmation.

Les coefficients de pondération de ces variables formant le vecteur β reflètent l'importance relative de chacun d'eux.

² Fonder le modèle de choix modal uniquement sur des variables d'offre est une hypothèse de travail très forte. L'expérience de nombreuses agglomérations à l'étranger a bien montré qu'une amélioration de l'offre de transports collectifs, même performante, ne provoque pas toujours les reports modaux attendus. La part des nouveaux voyageurs ("demande latente") est parfois plus importante que la part des reportés de la voiture. Le choix du mode intègre de nombreux aspects qualitatifs, d'ordre sociologique, qui sont très difficiles à modéliser et à projeter dans le futur, tels que les préférences, les habitudes, le paraître, ...

3. CONSTITUTION D'UNE BASE DE DONNEES

La difficulté dans le développement d'un modèle désagrégé réside dans la constitution du fichier des observations individuelles. Ce fichier doit comporter si possible des données récentes, un nombre suffisant d'observations (aucune étude n'a montré quel est ce nombre mais on peut penser par analogie avec d'autres domaines d'étude qu'un effectif de 500 à 1000 est un minimum) et surtout mixer à la fois des données d'enquêtes et des données reconstituées par simulation.

Dans cette partie, nous décrivons la méthode utilisée pour élaborer un fichier d'observations de qualité.

3.1. Reconstitution des temps généralisés

Pour caler un modèle de ce type, il faut disposer des temps généralisés VP et TC pour chaque liaison i-j. Le modèle de l'IAURIF en 488 zones (non compris les 34 zones externes) fonctionnant sur le logiciel MINUTP a été utilisé pour reconstituer les temps généralisés de zone à zone. La carte du découpage figure en annexe 3.

Le fichier des temps généralisés en TC contient la décomposition en temps partiels. Le fichier des temps généralisés VP contient des variables qui font référence à la zone du découpage morphologique de l'IAURIF³. Chaque fichier (VP et TC) contient donc 488*488 enregistrements, moins les temps intra-zonaux (le modèle ne calcule pas ces temps), soit 237 656 enregistrements.

Pour une description plus précise de l'offre VP et TC du modèle, il faut se reporter au document [9].

Un extrait des fichiers pour chaque mode est présenté ci-après sur les 11 premiers couples origine-destination :

³ Paris, banlieue intérieure, banlieue extérieure+villes nouvelles, franges de l'agglomération +agglomérations secondaires, rural. Voir carte en annexe 3.

1) Temps généralisés en TC :

Ces temps sont calculés par le modèle de l'IAURIF, sur un réseau TC de l'année 1996, à l'heure de pointe du matin.

ZONE_O	ZONE_D	WTT	WKI	WKJ	WKX	MTO	XFE	TTC
1	2	200	420	580	0	87	0	13
1	3	200	480	400	0	115	0	12
1	4	300	300	300	0	356	0	13
1	5	200	480	290	0	243	0	12
1	6	200	480	820	0	115	0	16
1	7	300	420	220	264	368	1	16
1	8	200	420	460	0	229	0	13
1	9	200	420	450	0	229	0	13
1	10	300	420	580	264	414	1	20
1	11	200	430	790	0	476	0	19
1	12	200	420	770	0	87	0	15

...

Notations : WTT à MTO sont des temps en centièmes de minutes, TTC est arrondi à la minute.

- ZONE_O : zone origine
- ZONE_D : zone destination
- WTT : temps d'attente total
- WKI : temps de rabattement au départ
- WKJ : temps de rabattement à l'arrivée
- WKX : temps de correspondance total
- MTO : temps de parcours dans le transport collectif
- XFE : nombre de correspondances
- **TTC : temps total en TC**

Remarques :

- Le réseau TC du modèle de l'IAURIF comprend un réseau de bus très simplifié (97 lignes dont 21 sur Paris). Il est très incomplet en grande couronne, ce qui implique que pour des déplacements entre zones voisines à l'intérieur de la grande couronne, le modèle peut trouver un temps de parcours très élevé (supérieur à 5 heures à cause de rebroussements à Paris) et affecte alors la valeur nulle (particularité du logiciel utilisé). En fait, il faut considérer ces temps de parcours sur le réseau lourd comme étant infinis, c'est-à-dire non compétitifs par rapport aux temps en voiture sur

la même liaison. Avec un graphe de bus complet sur la région, ce problème ne se rencontrerait pas.

- Les temps de rabattement du modèle de l'IAURIF sont des temps moyens de rabattement tous modes confondus calculés par le SIGR.

2) Temps généralisés en VP

Ces temps sont calculés par le modèle de l'IAURIF, sur un réseau VP de l'année 1996, à l'heure de pointe du soir. La matrice résultante des temps a ensuite été transposée, on obtient ainsi les temps à l'heure de pointe du matin.

ZONE_O	FOUC_O	ZONE_D	FOUC_D	TPS	TPS+TPARK
1	1	2	1	5	14
1	1	3	1	7	16
1	1	4	1	18	27
1	1	5	1	20	29
1	1	6	1	14	23
1	1	7	1	13	22
1	1	8	1	12	21
1	1	9	1	12	21
1	1	10	1	13	22
1	1	11	1	7	16
1	1	12	1	12	21

...

- ZONE_O : zone origine
- ZONE_D : zone destination
- FOUC_O : zone origine du découpage morphologique en 5 zones
- FOUC_D : zone destination du découpage morphologique en 5 zones
- TPS : temps réel de parcours total en minutes
- TPS+TPARK : temps réel de parcours + temps de recherche de parking fonction des variables FOUC_O et FOUC_D fixé comme suit :

Zone	ORIGINE	DESTINATION
Paris	+3	+6
Banlieue intérieure	+2	+4
Banlieue ext. urbanisée + VN	0	+1
Franges de l'agglom. + agglom. sec.	0	0
Rural	0	0

En résumé, nous disposons des temps modélisés suivants :

<p>Pour le <i>mode TC</i>, le temps total est composé de :</p> <ul style="list-style-type: none"> - WTT : temps d'attente total - WKI : temps de rabattement au départ - WKJ : temps de rabattement à l'arrivée - WKX : temps de correspondance total - MTO : temps de parcours dans le transport collectif - XFE : nombre de correspondances 	<p>Pour le <i>mode VP</i>, le temps total est composé de :</p> <ul style="list-style-type: none"> - TPS : temps de parcours - TPARK : temps de recherche de parking
---	---

3.2. Structure des motifs à l'heure de pointe du matin

L'EGT 97-98 porte au total sur un échantillon de 4285 ménages, 9681 individus et 35907 déplacements.

D'après cette enquête, la structure des motifs de déplacements VP ou TC le matin entre 7h et 9h59 (heure d'arrivée) est la suivante :

	TC	VP	TOTAL	%
DOMICILE > TRAVAIL	938 401	995 467	1 933 868	38%
DOMICILE > ECOLE	517 785	366 404	884 189	17%
DOMICILE > AUTRES MOTIFS	261 064	962 686	1 223 750	24%
TRAVAIL > DOMICILE	10 378	16 700	27 077	1%
ECOLE > DOMICILE	2 246	0	2 246	0%
AUTRES MOTIFS > DOMICILE	15 864	211 793	227 657	5%
TRAVAIL > TOUS MOTIFS (SAUF DOMICILE)	6 433	91 505	97 938	2%
TOUS MOTIFS (SAUF DOMICILE) > TRAVAIL	88 937	255 731	344 667	7%
ECOLE > TOUS MOTIFS (AUTRES QUE DOMICILE OU TRAVAIL)	4 143	8 498	12 642	0%
TOUS MOTIFS (AUTRES QUE DOMICILE OU TRAVAIL) > ECOLE	10 758	18 769	29 527	1%
AUTRES CAS	31 469	242 374	273 843	5%
	1 887 478	3 169 927	5 057 405	100%

Ce tableau montre qu'en travaillant seulement sur les trois motifs au départ du domicile, on modélise près de 80% des déplacements le matin.

On se propose donc de constituer 3 bases de données : domicile->travail, domicile->école et domicile->autres motifs.

3.3. Méthode de constitution de la base de données extraite de l'EGT 98 pour le motif domicile->travail

La constitution de la base est réalisée en huit étapes :

1) Concaténation de fichiers

La première étape a consisté à concaténer les fichiers ménages, individus et déplacements en prenant comme clé composite les variables NQ, NLE et NUMPERS. Comme c'est le mode "regroupé" (nomenclature EGT) que l'on exploite, le fichier moyens n'est pas pris en compte.

2) Succession de filtres

Filtre	Variable EGT	Modalité
Motif D->T	DRDMOD91 (motif regroupé)	1
Mode VP ou TC	DMTR91 (mode regroupé)	1 , 2
Période de pointe du matin (heure d'arrivée entre 7h00 et 9h59) : 3 heures	DTHD	7 à 12
Longueur de déplacement (renseigné)	DLD	< 999997
Durée de déplacement (renseigné)	DDU	< 9997
Origine et destination (O/D en IdF)	DZDMODMD	1 à 12

La variable DMTR91 correspond au choix du mode par l'individu.

3) Sélection de variables

Il s'agit de ne retenir que quelques colonnes de notre fichier. A priori, on ne sait pas quels sont les déterminants fondamentaux du choix du mode, il est donc préférable à cette étape de conserver un maximum d'informations. En particulier, il est intéressant de garder les variables qui permettent d'estimer le coût monétaire de chaque déplacement. L'idée est plutôt d'élaguer des variables sans rapport avec notre étude de manière à alléger le fichier. Par ailleurs, il faut garder les variables qui permettent de filtrer les captifs d'un mode.

Variables propres aux déplacements :

Libellé	Variable EGT
Communes d'origine et de destination	DDO, DCO, DDF, DCF
Carroyages d'origine et de destination	DCRO, DCRF
Durée	DDU
Longueur	DLD
Mode principal utilisé	DMT98 (la modalité VP passager est intéressante)
Transport d'animaux ou objets lourds	DANI
Motif à la destination	DMD
Nb de personnes dans le véhicule	DNPVP
Stationnement réservé employeur	DSR
Où avez-vous stationné	DST

Variables propres aux individus :

Libellé	Variable EGT
Possède permis de conduire	IPC
Dispose d'un parking gratuit sur lieu de travail	IPGT (si IPGT=96 ou 97, alors IPGT=2)
Zone utilisation carte orange	ISO
Détention abonnement titre de transport	IUCO

Variables propres aux ménages :

Libellé	Variable EGT
Nb d'actifs dans le ménage	MNACT
Nb de véhicules dans le ménage	MNV4

Comme on étudie des comportements, il est préférable de ne pas garder les poids. En effet, le fait que M.Max ait un poids égal à 2000 ne lui accorde pas plus d'importance que M.Blaise dont le poids vaut 500.

4) Ajout d'informations sur le découpage

Pour améliorer la lisibilité du fichier, on ajoute deux colonnes correspondant aux noms des communes d'origine et de destination : NOM_O et NOM_D.

On ajoute aussi deux autres colonnes : la zone origine et la zone destination du découpage en 488 zones du modèle de l'IAURIF⁴ : ZONE_O et ZONE_D.

⁴ Voir carte en annexe 3.

Un filtre supplémentaire s'avère nécessaire à cette étape : supprimer les observations où $ZONE_O = ZONE_D$. En effet, en vue de l'étape suivante, les temps de parcours intrazonaux n'ont pas été modélisés.

5) Ajout de "pseudo-variables" concernant le mode alternatif

C'est l'étape délicate de la constitution du fichier. En effet, pour chaque observation du fichier (un déplacement effectué dans un certain mode), il manque les données sur le mode alternatif qui n'a pas été choisi par l'individu. Il est donc nécessaire de compléter le fichier. Comment ? A l'aide des matrices de temps en VP/TC reconstituées par le modèle de l'IAURIF tournant sur le logiciel MINUTP. Le contenu de ces matrices a déjà été décrit dans le paragraphe précédent.

Donc à chaque observation, on ajoute les données sur le mode alternatif à partir de la matrice de temps : si l'observation est un déplacement en VP, on lui ajoute les 6 variables du temps TC modélisé ; sinon (l'observation est un déplacement TC), on lui ajoute les 2 variables du temps VP modélisé.

Il convient de passer un petit filtre à ce stade : compte tenu de la remarque faite en 3.1 sur les temps modélisés TC, il faut supprimer les observations où le temps modélisé TC est nul.

6) Filtre sur les temps déclarés

Rappelons le problème inhérent à la variable DDU (temps déclaré) : pour une même origine-destination, il existe une forte disparité des temps déclarés, qui s'explique par trois facteurs :

- un temps déclaré est un temps perçu et non réel, dépendant des caractéristiques de l'individu.
- le temps déclaré par un individu comprend le temps de rabattement et le temps de diffusion : ces deux temps introduisent une grande variabilité dans le temps total pour une même origine-destination car ils sont en réalité particuliers à chaque déplacement. Les temps de parcours modélisés sont quant à eux des *temps moyens* de centroïde à centroïde.
- Les temps déclarés que l'on a retenus sont compris entre 7h00 et 9h59. Sur les flancs de cette période de pointe de 3 heures, la circulation pour le mode VP est moins dense que en pleine heure de pointe. Ceci introduit donc une variabilité des temps déclarés pour une même origine-destination.

Exemples flagrants extraits de l'EGT97-98 illustrant ce problème des temps déclarés à l'heure de pointe du matin :

- Pavillons-sous-Bois vers Paris 1^{er} **en VP** : 45 mn (contre 97 mn en temps simulé)
- Meudon vers Paris 15^{ème} (quartier Montparnasse⁵) **en VP** : 7 mn (contre 38 mn en temps simulé)
- Paris 12^{ème} vers Paris 15^{ème} **en TC** : 75 mn (contre 45 mn en temps simulé)
- Saint-Denis vers Paris 16^{ème} **en TC** : 85 mn (contre 47 mn en temps simulé)

Les temps simulés sont des temps moyens de zone à zone alors que les temps déclarés sont des temps individuels. Pour pouvoir comparer des temps moyens VP/TC, il faut donc retenir parmi les temps déclarés pour un mode que ceux qui sont "moyens". Il convient donc de procéder à ce stade à un nouveau filtrage du fichier : retenir seulement les observations dont le temps déclaré dans un mode donné ne diffère que de -20 mn ou +20 mn **et** de -40% ou +40% du temps simulé pour le même mode :

$$\Delta_{absolu} \in [-20mn, +20mn]$$

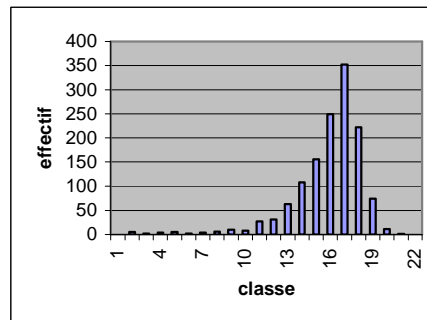
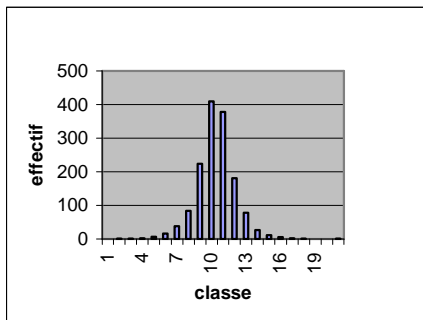
$$\Delta_{relatif} \in [-40\%, 40\%]$$

On a regardé la distribution des écarts absolus et des écarts relatifs :

⁵ Paris est découpé en 80 zones (quartiers) dans notre modèle.

intervalle	Classe	effectif
-90	1	0
-80	2	1
-70	3	1
-60	4	2
-50	5	7
-40	6	16
-30	7	38
-20	8	84
-10	9	224
0	10	410
10	11	378
20	12	181
30	13	78
40	14	27
50	15	12
60	16	6
70	17	2
80	18	1
90	19	0
100	20	0
110	21	1
120	22	0

intervalle	Classe	effectif
-300%	1	0
-280%	2	5
-260%	3	2
-240%	4	4
-220%	5	5
-200%	6	2
-180%	7	4
-160%	8	6
-140%	9	10
-120%	10	8
-100%	11	27
-80%	12	31
-60%	13	63
-40%	14	108
-20%	15	156
0%	16	249
20%	17	352
40%	18	222
60%	19	74
80%	20	11
100%	21	1
120%	22	0



7) Filtres sur les captifs VP ou TC

Il s'agit de faire l'estimation du modèle logit (pour le motif domicile-travail) sur un fichier de non-captifs, c'est-à-dire ceux qui ont réellement le choix du mode. On considère qu'un passager n'est pas un captif TC car il a eu le choix entre être accompagné en voiture ou prendre seul le bus.

Les captifs TC sont dans l'un des 2 cas suivants :

- individu sans permis de conduire et non-passager : DMTR91=1 et IPC=2 et DMT98≠17,
- individu sans véhicule et non-passager : DMTR91=1 et MNV4=0 et DMT98≠17,

Remarquons qu'il existe une autre catégorie de captif TC mais qui est difficilement extractable : c'est l'actif qui n'a pas accès à l'unique voiture du ménage, le conjoint l'utilisant déjà.

Par ailleurs, on considère qu'un individu ne disposant pas d'un parking disponible et gratuit sur lieu de travail n'est pas un captif (IPGT=2) car il a encore le choix du stationnement

payant. En conservant ainsi la variable IPGT, on pourra estimer le poids de la variable "disposition d'un parking gratuit au lieu de travail".

Les captifs VP sont dans l'un des quatre cas suivants :

- transport d'objets lourds : DMTR91=2 et DANI=1,
- motif à la destination étant un service passager : DMTR91=1 et DMD=15,
- motif intermédiaire au retour du travail vers le domicile nécessitant d'avoir une voiture particulière (motif type achat hebdomadaire).
- pas de gare à proximité de son domicile (par exemple dans un rayon de 15 km) ou à proximité de son lieu de travail

Les variables de l'EGT permettent de filtrer aisément les 2 premières catégories de captifs VP. Par contre, pour la troisième catégorie, cela demande une exploitation très lourde du fichier : la reconstitution des chaînes de déplacement. De même pour la quatrième catégorie, il faudrait utiliser les variables de carroyage et les exporter dans le Système d'Information Géographique Régional (SIGR) de façon à identifier les observations dont les carroyages d'origine ou de destination ne sont pas compris dans un cercle de rayon 15 km autour d'une gare. Ces deux dernières catégories de captifs VP ne seront donc dans une première approche pas filtrées.

8) Ajout de variables sur les coûts de déplacement

Afin d'intégrer la variable de coût monétaire dans le modèle, on ajoute pour toutes les observations les zones carte orange à l'origine et à la destination. On dispose pour cela d'une table de correspondance entre le zonage du modèle de l'IAURIF et le zonage carte orange. Deux nouvelles variables ont donc été créées : CO_O et CO_D.

Deux coûts de transport sont calculés pour chaque mode, l'un étant un coût ressenti, l'autre un coût que l'on qualifiera de tutelle. Le terrain de ce sujet est comme tout le monde sait très glissant, nous donnons seulement ci-après des formules ... :

- pour le mode TC,
 - 1^{er} coût (coût ressenti par l'utilisateur) : si l'individu possède un abonnement carte orange (IUCO >1), c'est le prix de la carte orange mensuelle (en tenant compte du remboursement employeur, on prend la moitié seulement) de zone à zone que divise un nombre moyen de déplacements mensuels (50) par le moyen de cette carte. Si IUCO=1, on considère que le coût de déplacement est le triple de celui effectué avec la carte orange.

- 2^{ème} coût (coût tutélaire) : on multiplie la longueur de déplacement (longueur à vol d'oiseau DLD+30%) par le coût kilométrique moyen donné par le Compte Voyageurs en Ile-de-France (0,43 Francs en 1997).
- Pour le mode VP,
 - 1^{er} coût (coût ressenti par l'usager) : on multiplie la longueur de déplacement (longueur à vol d'oiseau DLD+10%) par le coût moyen de carburant (6F pour 10 km), auquel on peut ajouter un coût de stationnement journalier moyen si l'individu ne dispose pas d'un parking au lieu de travail (IPGT=2) . Dans une simplification d'ordre méthodologique, on a estimé que ce coût de stationnement dépend de la zone de destination (65 F pour Paris, 12 F pour la petite couronne, 0 F pour la grande couronne). On pourrait affiner en prenant le découpage morphologique en cinq zones de l'IAURIF ou mieux encore en liant le coût de stationnement à la densité d'emploi de chaque zone.
 - 2^{ème} coût (coût tutélaire) : on multiplie la longueur de déplacement (longueur à vol d'oiseau DLD+20%) par le coût kilométrique moyen donné par le Compte Voyageurs en Ile-de-France (1,84 Francs en 1997).

Mise en oeuvre

Les huit étapes ont donc été déroulées à partir des fichiers originaux. Les filtres successifs réduisent progressivement le nombre d'observations :

Des 35904 observations au départ, le fichier final pour le domicile->travail ne comporte plus que 1003 observations. En effet, les filtres des étapes 1 à 5 fournissent 1749 observations. Puis les filtres sur les temps déclarés de l'étape 6 ne laissent que 1163 observations. Enfin, les filtres sur les captifs de l'étape 7 fournissent 1003 observations. Si l'on souhaite travailler sur 2 heures de pointe au lieu de 3, un filtre sur l'heure d'arrivée (DTHD compris entre 8 et 11) ne retient que 872 observations. On peut considérer que cet échantillon est suffisant pour estimer un modèle logit. Cependant, lorsque l'on s'attaquera aux autres motifs, il semblera nécessaire de procéder à des regroupements de sous-motifs.

Sur la page suivante figure un extrait de la base qui comporte 63 champs : les 18 premiers enregistrements.

Feuille "extrait base dt.xls"

3.4. Constitution de la base de données extraite de l'EGT 98 pour le motif domicile->école

Les déplacements domicile->école comprennent les déplacements domicile->école primaire, école secondaire, université. Les déplacements effectués en VP sont soit des déplacements d'un enfant déposé à l'école en voiture, soit des déplacements vers son lieu d'étude d'un individu (lycéen ou étudiant) ayant son permis de conduire. Rappelons que le déplacement d'un parent déposant son enfant à l'école est classé comme motif accompagnement et non école.

Quatre problèmes se posent pour le motif domicile->école :

- On travaille dans le cadre de cette étude avec des individus "moyens". Autant pour le motif domicile->travail, il n'y avait a priori pas de segmentation fondamentale à faire entre les actifs, autant pour le motif domicile->école on se retrouve avec une grande hétérogénéité des individus : enfants, collégiens, lycéens, étudiants. Les comportements vis à vis du choix du mode entre VP et TC sont en général des comportements de captifs d'un mode dû à des caractéristiques individuelles (âge, statut). On atteint ici la limite de la modélisation du choix modal par les seules variables d'offre sur des individus dont les disparités sont immenses.
- Comme l'étude traite uniquement du choix bimodal, elle délaisse les modes "doux" qui sont très utilisés pour le motif D-E. C'est donc un modèle très limité, destiné uniquement à la simulation et pas du tout explicatif, que nous sommes en train d'élaborer.
- D'après l'EGT97-98, le taux de déplacements à l'intérieur d'une commune à la période de pointe du matin est de plus d'un tiers (près de 60% pour la VP et de 77% pour le motif école primaire). Ces chiffres sont a fortiori encore plus élevés dans un zonage regroupant des communes, comme dans le zonage du modèle de l'IAURIF en 522 zones. Or comme ces déplacements doivent être retirés du fichier d'observations car les temps de parcours intrazonaux ne sont pas modélisés, ils réduisent très fortement la taille de la base.
- De nombreux déplacements des élèves sont effectués en bus. Une bonne modélisation du choix modal nécessite de disposer d'un réseau de bus complet ce qui n'est pas encore le cas pour le modèle global de l'IAURIF qui ne comporte qu'un graphe schématique de 97 lignes dont 21 sur Paris.

Les tableaux suivants de l'EGT97-98 fournissent quelques chiffres à la période de pointe du matin (7h-9h59) sur les parts modales par statut et par type de liaison.

	Région			à l'intérieur d'une commune		
	TC	VP	TOTAL	TC	VP	TOTAL
primaire	19 287	205 251	224 538	12 252	160 223	172 475
secondaire	314 796	110 251	425 047	71 169	46 843	118 012
étudiants	153 556	39 419	192 975	8 252	3 185	11 437
apprentis	15 130	1 387	16 516	-	-	-
Total	502 769	356 307	859 076	91 674	210 250	301 924

	Région			à l'intérieur d'une commune		
	TC	VP	TOTAL	TC	VP	TOTAL
primaire	9%	91%	100%	7%	93%	100%
secondaire	74%	26%	100%	60%	40%	100%
étudiants	80%	20%	100%	72%	28%	100%
apprentis	92%	8%	100%	0%	0%	0%
Total	59%	41%	100%	30%	70%	100%

	parts à l'intérieur d'une commune		
	TC	VP	TOTAL
primaire	64%	78%	77%
secondaire	23%	42%	28%
étudiants	5%	8%	6%
apprentis	0%	0%	0%
Total	18%	59%	35%

A la fin, si nous parvenons à élaborer une formule de partage modale, il faudra garder à l'esprit les 4 problèmes rencontrés pour le motif D->E et donc utiliser avec réserve le modèle logistique qui aura été estimé.

Par rapport à la base d'observations pour le motif domicile-travail, les différences tiennent au choix de nouvelles variables ou modalités :

- Les déplacements domicile->école sont codés par DRDMOD91=2.
- La définition des captifs TC pour le motif D->T (individu sans permis de conduire et non-passager ou individu sans véhicule et non-passager) ne tient plus pour le D->E car les déplacements VP des élèves sont tous effectués en motif accompagnement par le conducteur, donc on est obligé de les garder dans le fichier. On constate là encore les limites d'un modèle de choix modal qui ne prend pas en compte les chaînes de déplacements, du type domicile-service passager-travail, le motif intermédiaire étant un déterminant du choix modal.
- La variable IPGT (disposition d'un parking gratuit sur lieu de travail) ne sert plus. Par contre, on a récupéré la variable IOP (occupation principale ou statut : élève du primaire, élève du secondaire, étudiant), dans le but d'élaborer quelques tableaux croisés.
- L'abonnement carte orange d'un élève ou d'un étudiant n'est pas subventionné (en 1998, il n'y avait pas encore la carte Imagin'R).

- Dans le coût de déplacement en VP, on n'intègre plus le coût du stationnement. En effet, l'EGT montre que les individus se déplaçant en VP pour le motif D->E savent pour la plupart où stationner (DSSAV=1) et qu'ils se sont garés gratuitement.

Mise en oeuvre

Les huit étapes ont donc été déroulées à partir des fichiers originaux. Les filtres successifs réduisent progressivement le nombre d'observations :

Des 35904 observations au départ, le fichier final pour le domicile->école ne comporte plus que 335 observations à la période de pointe du matin. La taille de cette base interpelle sur deux points : d'une part on peut légitimement se demander si l'effectif est suffisant pour estimer un modèle logit, d'autre part il y a confirmation que l'on ne peut décemment pas segmenter la base d'observations en trois groupes (élèves du primaire, élèves du secondaire, étudiants) au risque de se retrouver avec des fichiers inexploitable.

3.5. Constitution de la base de données extraite de l'EGT 98 pour le motif domicile->autres

Les déplacements domicile->autres comprennent les déplacements domicile->affaires personnelles, affaires professionnelles⁶, achats, santé, loisirs, visites, service passager.

Les tableaux suivants de l'EGT97-98 fournissent quelques chiffres à la période de pointe du matin (7h-9h59) sur les parts modales par sous-motif :

⁶ Il aurait été sans doute plus judicieux de regrouper domicile-travail et domicile-affaires professionnelles.

	Région			à l'intérieur d'une commune		
	TC	VP	TOTAL	TC	VP	TOTAL
affaires pro.	131 124	211 923	343 047	10 134	19 209	29 342
achats	25 326	111 284	136 610	12 989	43 165	56 154
loisirs	14 102	37 057	51 159	4 810	21 776	26 587
visites	10 070	24 325	34 395	-	4 706	4 706
santé	10 964	26 293	37 257	1 989	5 986	7 975
affaires perso.	56 417	69 117	125 534	9 325	29 637	38 963
service passager	13 061	482 687	495 748	5 201	294 277	299 479
TOTAL	261 064	962 686	1 223 750	44 448	418 757	463 205

	Région			à l'intérieur d'une commune		
	TC	VP	TOTAL	TC	VP	TOTAL
affaires pro.	38%	62%	100%	35%	65%	100%
achats	19%	81%	100%	23%	77%	100%
loisirs	28%	72%	100%	18%	82%	100%
visites	29%	71%	100%	0%	100%	100%
santé	29%	71%	100%	25%	75%	100%
affaires perso.	45%	55%	100%	24%	76%	100%
service passager	3%	97%	100%	2%	98%	100%
Total	21%	79%	100%	10%	90%	100%

	parts à l'intérieur d'une commune		
	TC	VP	TOTAL
Affaires pro.	8%	9%	9%
achats	51%	39%	41%
loisirs	34%	59%	52%
visites	0%	19%	14%
santé	18%	23%	21%
affaires perso.	17%	43%	31%
service passager	40%	61%	60%
Total	17%	43%	38%

Ces résultats à la période de pointe du matin appellent quelques remarques :

- Comme pour le D-E, il y a une certaine hétérogénéité des comportements modaux, la part VP allant de 55% à 97% suivant le sous-motif.
- La part des déplacements intracommunaux est encore plus importante (38%) que pour le motif domicile->école (35%).
- Le volume de déplacements des services passagers est très important. Il constitue pratiquement 50% des déplacements VP pour domicile->autres motifs. Or c'est un motif qui est en général "oublié" dans les modèles de simulation. On sait relativement bien modéliser les déplacements achats, santé, ... lorsque les données (surfaces commerciales, capacités en lits des établissements hospitaliers, ...) sont disponibles mais comment modéliser l'accompagnement à la gare, à l'école, ... ? D'où l'intérêt de regrouper les sous-motifs dans la base d'observations.

Par rapport à la base d'observations pour le motif domicile-travail, les différences tiennent au choix de nouvelles variables ou modalités :

- Les déplacements domicile->autres motifs sont codés par DRDMOD91=3.
- La variable IPGT (disposition d'un parking gratuit sur lieu de travail) ne sert plus.

Mise en oeuvre

Les huit étapes ont donc été déroulées à partir des fichiers originaux. Les filtres successifs réduisent progressivement le nombre d'observations :

Des 35904 observations au départ, le fichier final pour le domicile->autres motifs ne comporte plus que 406 observations à la période de pointe du matin.

4. ESTIMATION DU MODELE DESAGREGÉ

Pour estimer les paramètres d'un tel modèle, il y a aujourd'hui plusieurs logiciels commerciaux, soit spécialisés tels que ALOGIT, LIMDEP, HIELOW, soit de statistique générale avec des procédures logit, tels que STATA ou SAS.

Ils sont tous basés sur le même principe, l'optimisation des valeurs des paramètres par la méthode du maximum de vraisemblance.

Nous avons choisi pour cette étude d'utiliser la procédure LOGISTIC du logiciel SAS, sachant qu'il y en a d'autres : CATMOD, PHREG.

Il n'y a pas d'autre méthode que l'exploration ou la méthode par essais-erreurs ou encore "à taton". Nous avons testé une trentaine de modèles différents.

La procédure LOGISTIC ajuste le modèle sur la probabilité de la modalité la plus faible, c'est-à-dire $P(Y=0)$ soit $P(Y=TC)$.

Le code du programme PROC LOGISTIC tient en quelques lignes :

```
data tablei;
infile "dthpmi.csv" delimiter=";";
input choix var1 var2 ... varn;
run;
title 'régression logistique numéro i' ;
proc logistic data=tablei;
model choix= var1 var2 ...varn /maxiter=25 ctable pprob=0.50;
run ;
```

Pour chaque motif, nous reproduisons ci-dessous les sorties SAS des régressions logistiques les plus intéressantes. Le test global du chi-2 a permis de rejeter l'hypothèse nulle (modèle sans variable explicative) de chacun des modèles, on ne fera donc pas figurer ci-dessous le résultat du test. Ensuite, pour chaque variable, on a dans l'ordre : l'estimateur du paramètre (β_i) et l'écart type σ_i de ce paramètre, puis la statistique de Wald (carré de la statistique de Student) et la probabilité de significativité du paramètre. Les deux dernières colonnes correspondent à l'estimateur standardisé et à l'odds ratio, elles ne seront pas utilisées.

Attachons nous au test de nullité d'un paramètre (y compris sur la constante INTERCEPT) donné par la statistique de Wald⁷ et la probabilité de significativité p :

⁷ Cette statistique suit une loi du chi² à 1 degré de liberté. L'hypothèse de la nullité d'un coefficient est rejetée lorsque la statistique de Wald dépasse un certain seuil, environ 4 pour une significativité à 5%.

- si stat Wald > 4, β_i a 95% de chances de se situer entre $\hat{\beta}_i - 2\hat{\sigma}_i$ et $\hat{\beta}_i + 2\hat{\sigma}_i$
- si $p < 0.05$, β_i est significativement non nul.

Pour évaluer la capacité prédictive du modèle, on regardera aussi les indices de corrélation entre prédiction et observation. Plus le nombre de paires concordantes est élevé, meilleur est le modèle. Il en est de même pour les 4 indices D, Gamma, Tau-a et c (voir la documentation technique de SAS pour plus de détails sur ces indices [10]).

Cette exploration a bénéficié des remarques de Cédric FONTAN, chercheur au département THEMA de l'Université de Cergy-Pontoise.

4.1. Motif domicile->travail

Nous disposons d'un fichier de 1003 observations qui comprend 553 déplacements en TC et 450 déplacements en VP.

La variable à expliquer, DMTR91 (choix du mode), a été recodée en 0 (TC) ou 1 (VP). De même, la variable IPGT a été recodée en 0 (NON) et 1 (OUI).

Listons les variables potentiellement explicatives (elles ont été renommées de façon à être parlantes). Seules deux variables liées aux caractéristiques individuelles ont été retenues car elles sont les moins difficiles à projeter dans le futur.

TVP	Temps total de parcours en VP simulé (TPS+TPARK)
TTC	Temps total de parcours en TC simulé (WKI+WKJ+WTT+WKX+MTO)
DELTAT	TVP-TTC
CVP1	Coût de déplacement VP ressenti directement (carburant seulement)
CVP11	Coût de déplacement VP ressenti (carburant +stationnement)
CTC1	Coût de déplacement TC ressenti (carte orange)
CVP2	Coût de déplacement VP tutélaire (1,84 F/km)
CTC2	Coût de déplacement TC tutélaire (0,43 F/km)
DELTAC1	CVP1-CTC1
DELTAC11	CVP11-CTC1
DELTAC2	CVP2-CTC2
TPARK	Temps de recherche de parking simulé en VP
TRAB	Temps de rabattement en TC simulé (WKI+WKJ)
TATT	Temps d'attente en TC simulé (WTT)

TCOR	Temps de correspondance en TC simulé (WKX)
NBCOR	Nombre de correspondances en TC simulé (XFE)
DPARK	Disposition d'un parking gratuit sur lieu de travail (IPGT)
TAUM	Nombre de véhicules du ménage (MNV4)

1) Variables : TVP et TTC

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.4751	0.2106	5.0878	0.0241	.	.
TVP	1	0.0870	0.00569	233.8425	0.0001	1.497801	1.091
TTC	1	-0.0719	0.00638	126.9595	0.0001	-0.919262	0.931

Association of Predicted Probabilities and Observed Responses

Concordant = 88.2%	Somers' D = 0.765
Discordant = 11.7%	Gamma = 0.766
Tied = 0.1%	Tau-a = 0.379
(248850 pairs)	c = 0.882

Le taux de concordance est satisfaisant. C'est la statistique de Wald pour la constante qui est à améliorer. Des variables sont peut-être omises.

2) Variables : DELTAT et DELTAC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	0.1048	0.0993	1.1152	0.2909	.	.
DELTAT	1	0.0840	0.00556	227.7901	0.0001	1.340798	1.088
DELTAC1	1	-0.0214	0.0147	2.1141	0.1459	0.079717	1.022

Association of Predicted Probabilities and Observed Responses

Concordant = 87.9%	Somers' D = 0.758
Discordant = 12.0%	Gamma = 0.759
Tied = 0.1%	Tau-a = 0.375
(248850 pairs)	c = 0.879

Si l'on travaille seulement sur les variables de différences, le modèle est mal spécifié. Il est préférable de travailler directement avec les variables de temps globaux que les différences. Introduisons toutes les variables d'offre sur le 1^{er} modèle.

3) Variables : TVP, TTC, TPARK, TRAB, TATT, TCOR, NBCOR

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-1.3521	0.4289	9.9398	0.0016	.	.
TVP	1	0.0691	0.00653	112.1241	0.0001	1.190318	1.072
TTC	1	-0.0404	0.0114	12.4841	0.0004	-0.515900	0.960
TPARK	1	0.1727	0.0436	15.6910	0.0001	0.287067	1.189
TRAB	1	-0.0411	0.0169	5.8979	0.0152	-0.181913	0.960
TATT	1	0.00453	0.0292	0.0241	0.8767	0.014028	1.005
TCOR	1	-0.0357	0.0488	0.5366	0.4638	-0.056606	0.965
NBCOR	1	0.0892	0.1329	0.4500	0.5023	0.045363	1.093

Association of Predicted Probabilities and Observed Responses

Concordant = 89.0%	Somers' D = 0.781
Discordant = 10.9%	Gamma = 0.782
Tied = 0.1%	Tau-a = 0.387
(248850 pairs)	c = 0.890

Les temps décomposés ne jouent apparemment pas. Ajoutons à ce modèle des variables que tout le monde dit et répète être déterminantes dans le choix modal : la disposition d'un parking gratuit au lieu de travail et le nombre de véhicules du ménage.

4) Variables : TVP, TTC, DPARK, TAUM, TPARK, TRAB, TATT, TCOR, NBCOR

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	1.1147	0.5204	4.5883	0.0322	.	.
TVP	1	0.0801	0.00749	114.1573	0.0001	1.378217	1.083
TTC	1	-0.0457	0.0128	12.7387	0.0004	-0.584312	0.955
DPARK	1	-1.1327	0.1980	32.7130	0.0001	-0.312257	0.322
TAUM	1	-1.3709	0.1529	80.4154	0.0001	-0.605097	0.254
TPARK	1	0.0441	0.0500	0.7775	0.3779	0.073318	1.045
TRAB	1	-0.0336	0.0188	3.2032	0.0735	-0.149031	0.967
TATT	1	0.0298	0.0323	0.8524	0.3559	0.092306	1.030
TCOR	1	-0.0475	0.0548	0.7525	0.3857	-0.075332	0.954
NBCOR	1	0.0721	0.1484	0.2358	0.6273	0.036664	1.075

Association of Predicted Probabilities and Observed Responses

Concordant = 92.6%	Somers' D = 0.853
Discordant = 7.3%	Gamma = 0.854
Tied = 0.1%	Tau-a = 0.423
(248850 pairs)	c = 0.927

Les Wald chi-square des 5 dernières variables sont inférieurs à 4. On ne retiendra donc finalement aucune de ces variables dans un modèle intégrant déjà les temps totaux de déplacement et sans prise en compte des coûts monétaires. La constante n'est pas satisfaisante. Les 2 variables DPARK et TAUM sont en effet très pertinentes.

5) Variables : TVP, TTC, DPARK, TAUM

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	1.1917	0.2786	18.3028	0.0001	.	.
TVP	1	0.0822	0.00603	186.1977	0.0001	1.415467	1.086
TTC	1	-0.0521	0.00676	59.4278	0.0001	-0.665711	0.949
DPARK	1	-1.1523	0.1937	35.3761	0.0001	-0.317653	0.316
TAUM	1	-1.3788	0.1482	86.5203	0.0001	-0.608621	0.252

Association of Predicted Probabilities and Observed Responses

Concordant = 92.4%	Somers' D = 0.849
Discordant = 7.5%	Gamma = 0.850
Tied = 0.1%	Tau-a = 0.421
(248850 pairs)	c = 0.925

Ce modèle sans facteur coût semble bien spécifié.

6) Variables : TVP, CVP1, TTC, CTC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-2.1694	0.3911	30.7617	0.0001	.	.
TVP	1	0.0755	0.00629	144.0542	0.0001	1.299975	1.078
CVP1	1	0.00432	0.0326	0.0175	0.8947	0.017061	1.004
TTC	1	-0.0722	0.00913	62.5201	0.0001	-0.923116	0.930
CTC1	1	0.6545	0.1121	34.1026	0.0001	0.750557	1.924

Association of Predicted Probabilities and Observed Responses

Concordant = 89.8%	Somers' D = 0.797
Discordant = 10.1%	Gamma = 0.798
Tied = 0.1%	Tau-a = 0.395
(248850 pairs)	c = 0.899

L'introduction des coûts monétaires améliore le Wald chi-square pour la constante. Par contre le paramètre de CVP1 n'est pas significatif. Ceci peut s'expliquer en disant que le coût VP évalué uniquement par la dépense de carburant n'influence pas ceux qui peuvent utiliser leur voiture. Ceux qui utilisent les TC ont quelque part intégré le coût du ticket dans leur choix du mode. Peut-être que la variable CVP1 (coût VP du carburant seulement) ne semble pas la variable pertinente pour le coût VP. On va essayer avec la variable CVP11 (carburant + stationnement).

7) Variables : TVP, TTC, CVP11, CTC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-2.6134	0.3726	49.1945	0.0001	.	.
TVP	1	0.0598	0.00599	99.7084	0.0001	1.029947	1.062
TTC	1	-0.0563	0.00650	75.0548	0.0001	-0.719635	0.945
CVP11	1	0.0306	0.00472	42.1099	0.0001	0.506853	1.031
CTC1	1	0.5920	0.1095	29.2070	0.0001	0.678929	1.808

Association of Predicted Probabilities and Observed Responses

Concordant = 91.3%	Somers' D = 0.827
Discordant = 8.6%	Gamma = 0.828
Tied = 0.1%	Tau-a = 0.409
(248850 pairs)	c = 0.913

Le modèle devient vraiment intéressant au vu de toutes les statistiques et les indices.

Considérons ce modèle comme le modèle de référence et introduisons alors toutes les autres variables.

Remarque : On peut en profiter pour calculer la valeur du temps des automobilistes pour le motif D-T. Elle est donnée par le ratio des paramètres de TVP et CVP11 : $0,0598/0,0306*60 = 117$ F/h. C'est une valeur dont l'ordre de grandeur s'accorde bien avec des valeurs du temps connues et nous conforte donc dans le bien fondé de la spécification de notre modèle.

8) Variables : TVP, TTC, CVP11, CTC1, DPARK, TAUM, TPARK, TRAB, TATT, TCOR, NBCOR

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
----------	----	--------------------	----------------	-----------------	-----------------	-----------------------	------------

INTERCPT	1	-0.7204	0.6081	1.4034	0.2362	.	.
TVP	1	0.0691	0.00803	74.2150	0.0001	1.190307	1.072
TTC	1	-0.0504	0.0137	13.5002	0.0002	-0.643543	0.951
CVP11	1	0.0194	0.00677	8.2016	0.0042	0.320977	1.020
CTC1	1	0.7993	0.1286	38.6413	0.0001	0.916620	2.224
DPARK	1	-0.5925	0.2690	4.8514	0.0276	-0.163331	0.553
TAUM	1	-1.6641	0.1766	88.7953	0.0001	-0.734555	0.189
TPARK	1	-0.0512	0.0561	0.8314	0.3619	-0.085064	0.950
TRAB	1	-0.0340	0.0201	2.8611	0.0907	-0.150628	0.967
TATT	1	0.0302	0.0336	0.8100	0.3681	0.093490	1.031
TCOR	1	-0.1019	0.0607	2.8193	0.0931	-0.161434	0.903
NBCOR	1	0.2031	0.1522	1.7794	0.1822	0.103303	1.225

Association of Predicted Probabilities and Observed Responses

Concordant = 94.3%	Somers' D = 0.886
Discordant = 5.7%	Gamma = 0.887
Tied = 0.1%	Tau-a = 0.439
(248850 pairs)	c = 0.943

Les Wald chi-square des 5 dernières variables sont inférieurs à 4. On ne retiendra donc aucune de ces variables. La variable DPARK n'est pas pertinente dans ce modèle. Ceci s'explique par le fait qu'elle est corrélée à la variable CVP11.

9) Variables : TVP, TTC, CVP11, CTC1, TAUM

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-1.4395	0.4027	12.7759	0.0004	.	.
TVP	1	0.0599	0.00643	86.8150	0.0001	1.030819	1.062
TTC	1	-0.0449	0.00697	41.4878	0.0001	-0.573675	0.956
CVP11	1	0.0260	0.00501	27.0071	0.0001	0.430978	1.026
CTC1	1	0.7143	0.1167	37.4484	0.0001	0.819188	2.043
TAUM	1	-1.6006	0.1676	91.1843	0.0001	-0.706509	0.202

Association of Predicted Probabilities and Observed Responses

Concordant = 93.9%	Somers' D = 0.879
Discordant = 6.0%	Gamma = 0.880
Tied = 0.1%	Tau-a = 0.435
(248850 pairs)	c = 0.940

En introduisant le taux de motorisation dans le modèle de référence, on améliore encore la qualité prédictive du modèle. Mais la variable INTERCEPT est moins "constante" que dans le modèle spécifié sans la variable TAUM.

10) Variables : TVP, TTC, CVP1, CTC1, DPARK, TAUM

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.3655	0.4550	0.6453	0.4218	.	.
TVP	1	0.0651	0.00677	92.3743	0.0001	1.120145	1.067
TTC	1	-0.0624	0.0102	37.6278	0.0001	-0.796896	0.940
CVP1	1	0.0672	0.0364	3.4099	0.0648	0.265641	1.070
CTC1	1	0.7629	0.1220	39.1082	0.0001	0.874911	2.145
DPARK	1	-1.1151	0.2077	28.8140	0.0001	-0.307403	0.328
TAUM	1	-1.6070	0.1677	91.8642	0.0001	-0.709353	0.200

Association of Predicted Probabilities and Observed Responses

Concordant = 93.9%	Somers' D = 0.879
Discordant = 6.0%	Gamma = 0.880
Tied = 0.1%	Tau-a = 0.435
(248850 pairs)	c = 0.940

On réintroduit la disposition d'un parking mais en prenant le coût VP dû au carburant seulement. On constate que la variable CVP1 n'est alors pas pertinente.

11) Variables : TVP, TTC, CVP2, CTC2, TAUM

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.7319	0.4371	2.8030	0.0941	.	.
TVP	1	0.0757	0.00647	136.7688	0.0001	1.302803	1.079
TTC	1	-0.0696	0.00982	50.1437	0.0001	-0.889129	0.933
CVP2	1	0.1504	0.8179	0.0338	0.8541	1.822412	1.162
CTC2	1	-0.4770	3.2282	0.0218	0.8825	-1.462954	0.621
DPARK	1	1.1992	0.1961	37.3975	0.0001	0.330591	3.318
TAUM	1	-1.4170	0.1506	88.5106	0.0001	-0.625451	0.242

Association of Predicted Probabilities and Observed Responses

Concordant = 92.6%	Somers' D = 0.853
Discordant = 7.3%	Gamma = 0.854
Tied = 0.1%	Tau-a = 0.423
(248850 pairs)	c = 0.927

Ici, ce sont les coûts de déplacements avec les valeurs de tutelle qui sont introduits. Au vu des statistiques, les variables CVP2 et CTC2 ne sont pas pertinentes. On prendra donc les coûts ressentis.

En **conclusion**, on retiendra 4 modèles pour le motif domicile-travail : les numéros 1, 5, 7 et

9. Les notations sont les suivantes :

- P la probabilité de prendre les transports en commun pour aller de la zone i à la zone j :
P = proba (choix = TC) = part TC en %
- T_{vp} : temps global de parcours VP de i à j . Ce temps inclut un temps fixé de recherche de parking.
- T_{tc} : temps global de parcours TC de i à j.
- C_{vp} : coût de déplacement ressenti prenant en compte le coût du carburant et le coût du stationnement à la destination.
- C_{tc} : coût de déplacement ressenti prenant en compte le coût de la carte orange et le remboursement employeur
- DPARK : disposition d'un parking gratuit au lieu de travail.
- ϑ_m : taux de motorisation par ménage

A) Modèle le plus simple : temps seulement

$$\text{Logit (P)} = \log (P/1-P) = -0,475 + 0,087 T_{vp} - 0,072 T_{tc}$$

Donc

$$P = \frac{1}{1 + \exp[-(-0,475 + 0,087 T_{vp} - 0,072 T_{tc})]}$$

B) Modèle sans coûts :

$$\text{Logit (P)} = \log (P/1-P) = 1,19 + 0,082 T_{vp} - 0,052 T_{tc} - 1,15 \text{ DPARK} - 1,39 \vartheta_m$$

Donc

$$P = \frac{1}{1 + \exp[-(1,19 + 0,082 T_{vp} - 0,052 T_{tc} - 1,15 \text{ DPARK} - 1,39 \vartheta_m)]}$$

C) Modèle avec coûts et taux de motorisation ϵ_m :

$$\text{Logit (P)} = \log (P/1-P) = -1,44 + 0,060 T_{vp} - 0,045 T_{tc} + 0,026 C_{vp} + 0,714 C_{tc} - 1,60 \epsilon_m$$

Donc

$$P = \frac{1}{1 + \exp[-(-1,44 + 0,060 T_{vp} - 0,045 T_{tc} + 0,026 C_{vp} + 0,714 C_{tc} - 1,60 \epsilon_m)]}$$

D) Modèle avec coûts et sans taux de motorisation :

$$\text{Logit (P)} = \log (P/1-P) = -2,61 + 0,060 T_{vp} - 0,056 T_{tc} + 0,031 C_{vp} + 0,59 C_{tc}$$

Donc

$$P = \frac{1}{1 + \exp[-(-2,61 + 0,060 T_{vp} - 0,056 T_{tc} + 0,031 C_{vp} + 0,59 C_{tc})]}$$

Commentaires

- Considérons les signes des paramètres :
 - Plus le temps ou le coût en VP augmente, plus la probabilité de prendre les TC est grande. Les signes de T_{vp} et C_{vp} sont effectivement positifs.
 - Inversement pour le temps en TC. Le signe de T_{tc} est effectivement négatif.
 - La variable de coût TC qui a été retenue est le coût du ticket. Le signe positif s'explique par le fait que le coût de la carte orange est liée à la distance entre domicile et lieu de travail. Si elle est chère, cela signifie que l'on habite loin du lieu de travail, et tout voyageur a pleinement conscience que sur une longue distance le coût de déplacement en TC est nettement moindre que le coût en VP. Et plus la carte orange est chère, plus on désire amortir le prix et donc on se déplace en TC. Par ailleurs le fait d'avoir augmenté régulièrement depuis 10 ans le prix de la carte orange au dessus de l'inflation n'a pas provoqué une désaffection des usagers des TC. Ainsi, comme le montre le signe positif de C_{TC} , si le prix augmente, la part TC ne diminue pas.

- Plus la possibilité de garer sa voiture gratuitement au lieu de travail est grande, moins on a tendance à prendre les TC : le signe de DPARK est effectivement négatif.
- Plus le taux de motorisation est élevé, moins on prend les TC : le signe de ε_m est effectivement négatif.

Ainsi, les signes des paramètres sont conformes à l'intuition.

- Quel modèle retenir ? Cela dépend des données dont on dispose dans le cadre d'une étude. Il y a un compromis à trouver entre la précision du modèle en sortie et la précision des données en entrée. Le taux de motorisation à un horizon futur peut être obtenu par un autre modèle qu'il faut faire tourner pour chaque zone du modèle. Plus simplement, on peut se contenter de faire des "hypothèses d'expert" par type de zone. Par contre, la donnée DPARK n'est pas facile à recueillir. C'est une donnée à fixer pour chaque zone du modèle, allant de 0 (NON) à 1 (OUI) . On pourrait par exemple prendre pour Issy-les-Moulineaux DPARK= 1, pour Paris 1^{er} DPARK=0, et pour Paris 15^{ème} DPARK=0,3. On voit bien que ces valeurs vont donner lieu à des discussions sans fin. Mais cette variable est nécessaire dans un modèle sans coût. Si l'on retient un modèle avec coût, la difficulté réside dans le calcul des coûts de déplacements à un horizon lointain. Le zonage carte orange actuel existera-t-il encore à 2015 ?

4.2. Motif domicile->école

Nous disposons d'un fichier de 335 observations qui comprend 242 déplacements en TC et 93 déplacements en VP.

Sur les 93 déplacements en VP, 38 sont des passagers accompagnés dans le cadre d'une boucle de déplacements (de type Domicile-Ecole-Travail), 24 sont des conducteurs, 31 sont des passagers accompagnés à l'école puis l'accompagnateur retourne directement au domicile. Il faudrait normalement retirer de la base les 38 déplacements faisant partie d'une boucle car le choix modal du passager dépend de celui de l'accompagnateur dans sa chaîne de déplacements. Ici, nous avons choisi de les garder pour avoir un échantillon suffisant et équilibré en VP/TC.

Listons les variables potentiellement explicatives (elles ont été renommées de façon à être parlantes).

TROULVP	Temps de roulement en VP (TPS)
TVP	Temps total de parcours en VP simulé (TPS+TPARK)
TTC	Temps total de parcours en TC simulé (WKI+WKJ+WTT+WKX+MTO)
DELTAT	TVP-TTC
CVP1	Coût de déplacement VP ressenti directement (carburant seulement)
CTC1	Coût de déplacement TC ressenti (carte orange)
DELTAC1	CVP1-CTC1
TPARK	Temps de recherche de parking simulé en VP
TRAB	Temps de rabattement en TC simulé (WKI+WKJ)
TATT	Temps d'attente en TC simulé (WTT)
TCOR	Temps de correspondance en TC simulé (WKX)
NBCOR	Nombre de correspondances en TC simulé (XFE)
TAUM	Nombre de véhicules du ménage (MNV4)

1) Variables : TROULVP, TTC

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	0.9993	0.3098	10.4023	0.0013	.	.
TROULVP	1	0.0481	0.00912	27.8023	0.0001	0.718997	1.049
TTC	1	-0.0344	0.00837	16.9039	0.0001	-0.405952	0.966

Association of Predicted Probabilities and Observed Responses

Concordant = 69.5%	Somers' D = 0.393
Discordant = 30.2%	Gamma = 0.395
Tied = 0.4%	Tau-a = 0.158
(22506 pairs)	c = 0.697

2) Variables : TVP, TTC

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	0.7415	0.3180	5.4370	0.0197	.	.
TVP	1	0.0440	0.00812	29.3480	0.0001	0.681280	1.045
TTC	1	-0.0293	0.00767	14.5960	0.0001	-0.345926	0.971

Association of Predicted Probabilities and Observed Responses

Concordant = 70.0%	Somers' D = 0.406
--------------------	-------------------

Discordant = 29.4%	Gamma = 0.409
Tied = 0.6%	Tau-a = 0.163
(22506 pairs)	c = 0.703

La comparaison des statistiques des deux modèles montre qu'il vaut mieux prendre le temps de parcours en VP sans y ajouter le temps de recherche de parking. Mais en basant le modèle uniquement sur des temps, on arrive à 70% de paires concordantes, contre près de 90% pour le motif D->T. Introduisons les variables de coûts.

3) Variables : TROULVP, TTC, CVP1, CTC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	2.9484	0.4989	34.9245	0.0001	.	.
TROULVP	1	0.0555	0.0139	15.9775	0.0001	0.830123	1.057
TTC	1	-0.0339	0.00961	12.4612	0.0004	-0.400344	0.967
CVP1	1	-0.0452	0.0555	0.6642	0.4151	-0.154163	0.956
CTC1	1	-0.1845	0.0272	46.0358	0.0001	-0.513483	0.831

Association of Predicted Probabilities and Observed Responses

Concordant = 81.1%	Somers' D = 0.624
Discordant = 18.7%	Gamma = 0.626
Tied = 0.3%	Tau-a = 0.251
(22506 pairs)	c = 0.812

Le coût VP ne joue pas contrairement au coût TC.

4) Variables : TROULVP, TTC, CTC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	3.0790	0.4781	41.4717	0.0001	.	.
TROULVP	1	0.0474	0.00917	26.7423	0.0001	0.708701	1.049
TTC	1	-0.0374	0.00889	17.6860	0.0001	-0.441152	0.963
CTC1	1	-0.1852	0.0271	46.6876	0.0001	-0.515369	0.831

Association of Predicted Probabilities and Observed Responses

Concordant = 80.8%	Somers' D = 0.620
Discordant = 18.9%	Gamma = 0.621
Tied = 0.3%	Tau-a = 0.249

(22506 pairs) c = 0.810

5) Variables : TROULVP, TTC, CTC1, DLD

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	2.9485	0.4987	34.9503	0.0001	.	.
TROULVP	1	0.0555	0.0139	15.9938	0.0001	0.830224	1.057
TTC	1	-0.0339	0.00961	12.4625	0.0004	-0.400286	0.967
CTC1	1	-0.1846	0.0272	46.0462	0.0001	-0.513548	0.831
DLD	1	-0.00003	0.00004	0.6663	0.4144	-0.154334	1.000

Association of Predicted Probabilities and Observed Responses

Concordant = 81.1% Somers' D = 0.624
 Discordant = 18.6% Gamma = 0.626
 Tied = 0.3% Tau-a = 0.251
 (22506 pairs) c = 0.812

6) Variables : TROULVP, TTC, DLD

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	0.8314	0.3331	6.2296	0.0126	.	.
TROULVP	1	0.0591	0.0133	19.6023	0.0001	0.882856	1.061
TTC	1	-0.0293	0.00906	10.4402	0.0012	-0.345465	0.971
DLD	1	-0.00005	0.000038	1.4785	0.2240	-0.219840	1.000

Association of Predicted Probabilities and Observed Responses

Concordant = 69.7% Somers' D = 0.397
 Discordant = 30.0% Gamma = 0.399
 Tied = 0.3% Tau-a = 0.160
 (22506 pairs) c = 0.699

Les modèles que nous retenons sont finalement les numéros 1 et 4.

Avec les notations suivantes :

- P est la part TC de la zone i (domicile) à la zone j (école).
- T_{vp}^R : temps de parcours (roulement pur) en VP de i à j, sans temps de recherche de parking.
- T_{tc} : temps de parcours en TC de i à j.
- C_{tc} : coût de déplacement ressenti en TC de i à j (sans subvention)

A) Modèle sans coût

$$\text{Logit}(P) = \log(P/1-P) = 1 + 0,048 T_{vp}^R - 0,034 T_{tc}$$

Donc

$$P = \frac{1}{1 + \exp[-(1 + 0,048 T_{vp}^R - 0,034 T_{tc})]}$$

B) Modèle avec coût

$$\text{Logit}(P) = \log(P/1-P) = 3,08 + 0,047 T_{vp}^R - 0,037 T_{tc} - 0,185 C_{tc}$$

Donc

$$P = \frac{1}{1 + \exp[-(3,08 + 0,047 T_{vp}^R - 0,037 T_{tc} - 0,185 C_{tc})]}$$

Commentaires

- Considérons les signes des paramètres :
 - Les signes de T_{vp}^R et C_{vp} sont positifs, conformément à ce qu'on pouvait prévoir.
 - A l'inverse de la formule pour le motif D->T, le signe de C_{TC} est négatif : moins l'abonnement est cher, plus la probabilité de prendre les TC est grande. En effet, la distance de déplacement D-E est en moyenne moins importante que la distance D-T, le facteur distance ne joue pas pour le motif D-E.
- Le signe négatif de C_{TC} pour les élèves et les étudiants inspire la réflexion suivante : pour augmenter la part TC pour les déplacements D-E, il faudrait diminuer le coût de l'abonnement carte orange ou le subventionner. Ceci est confirmé par la mise en place de la carte Imagine "R" (abonnement réservé aux scolaires et étudiants de moins de 26 ans, près de 500 000 cartes vendues en 1999) en 1998 qui a fait augmenter le trafic TC des jeunes.

4.2. Motif domicile->autres motifs

Nous disposons d'un fichier de 406 observations qui comprend 131 déplacements en TC et 275 déplacements en VP.

Listons les variables potentiellement explicatives (elles ont été renommées de façon à être parlantes).

TROULVP	Temps de roulement en VP (TPS)
TVP	Temps total de parcours en VP simulé (TPS+TPARK)
TTC	Temps total de parcours en TC simulé (WKI+WKJ+WTT+WKX+MTO)
DELTAT	TVP-TTC
CVP1	Coût de déplacement VP ressenti directement (carburant seulement)
CTC1	Coût de déplacement TC ressenti (carte orange)
DELTAC1	CVP1-CTC1
TPARK	Temps de recherche de parking simulé en VP
TRAB	Temps de rabattement en TC simulé (WKI+WKJ)
TATT	Temps d'attente en TC simulé (WTT)
TCOR	Temps de correspondance en TC simulé (WKX)
NBCOR	Nombre de correspondances en TC simulé (XFE)
TAUM	Nombre de véhicules du ménage (MNV4)

1) Variables : TROULVP, TTC

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-0.7853	0.2540	9.5613	0.0020	.	.
TROULVP	1	0.0499	0.00656	57.8166	0.0001	0.754363	1.051
TTC	1	-0.0441	0.00819	29.0613	0.0001	-0.624891	0.957

Association of Predicted Probabilities and Observed Responses

Concordant = 75.3%	Somers' D = 0.510
Discordant = 24.3%	Gamma = 0.512
Tied = 0.3%	Tau-a = 0.224
(36025 pairs)	c = 0.755

2) Variables : TVP, TTC

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-1.1208	0.2628	18.1874	0.0001	.	.
TVP	1	0.0516	0.00640	65.0495	0.0001	0.803473	1.053
TTC	1	-0.0443	0.00810	29.8634	0.0001	-0.626434	0.957

Association of Predicted Probabilities and Observed Responses

Concordant = 77.9%	Somers' D = 0.560
Discordant = 21.9%	Gamma = 0.562
Tied = 0.3%	Tau-a = 0.245
(36025 pairs)	c = 0.780

De ces deux premiers modèles, on peut conclure que la prise en compte d'un temps de recherche de parking dans le temps de parcours VP améliore le modèle.

3) Variables : TVP, TTC, CVP1, CTC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	0.2755	0.4650	0.3512	0.5535	.	.
TVP	1	0.0706	0.00883	63.9160	0.0001	1.098013	1.073
TTC	1	-0.0241	0.0112	4.5951	0.0321	-0.341137	0.976
CVP1	1	-0.1074	0.0459	5.4817	0.0192	-0.408116	0.898
CTC1	1	-0.1658	0.0237	48.8640	0.0001	-0.541259	0.847

Association of Predicted Probabilities and Observed Responses

Concordant = 85.5%	Somers' D = 0.712
Discordant = 14.4%	Gamma = 0.712
Tied = 0.1%	Tau-a = 0.312
(36025 pairs)	c = 0.856

Avec l'introduction des coûts de déplacements, le modèle est mal spécifié au niveau de la significativité des coefficients de la constante, du temps et du coût VP malgré un taux de reconnaissance plus élevé qu'un modèle sans coûts.

4) Variables : TVP, TTC, CTC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
----------	----	--------------------	----------------	-----------------	-----------------	-----------------------	------------

INTERCPT	1	0.8318	0.3948	4.4399	0.0351	.	.
TVP	1	0.0597	0.00731	66.6825	0.0001	0.928755	1.062
TTC	1	-0.0424	0.00885	22.9705	0.0001	-0.600552	0.958
CTC1	1	-0.1666	0.0230	52.3454	0.0001	-0.543989	0.847

Association of Predicted Probabilities and Observed Responses

Concordant = 84.8%	Somers' D = 0.698
Discordant = 15.0%	Gamma = 0.699
Tied = 0.1%	Tau-a = 0.306
(36025 pairs)	c = 0.849

L'écart-type de la constante est encore trop élevé.

5) Variables : TROULVP, TTC, CTC1

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	1.2503	0.3959	9.9716	0.0016	.	.
TROULVP	1	0.0590	0.00760	60.2493	0.0001	0.892117	1.061
TTC	1	-0.0437	0.00899	23.6804	0.0001	-0.619225	0.957
CTC1	1	-0.1683	0.0230	53.4220	0.0001	-0.549542	0.845

Association of Predicted Probabilities and Observed Responses

Concordant = 83.4%	Somers' D = 0.669
Discordant = 16.5%	Gamma = 0.670
Tied = 0.2%	Tau-a = 0.293
(36025 pairs)	c = 0.835

Toutes les statistiques de Wald sont acceptables. Le temps VP à retenir est un temps de roulement pur.

Les modèles que nous retenons sont finalement les numéros 2 et 5.

Avec les notation suivantes :

- P est la part TC de la zone i (domicile) à la zone j (autres motifs).
- T_{vp}^R : temps de parcours (roulement pur) en VP de i à j, sans temps de recherche de parking.

- T_{vp} : temps global de parcours VP de i à j . Ce temps inclut un temps fixé de recherche de parking.
- T_{tc} : temps de parcours en TC de i à j.
- C_{tc} : coût de déplacement ressenti en TC de i à j (sans subvention)

A) Modèle sans coût

$$\text{Logit (P)} = \log (P/1-P) = -1,12 + 0,052 T_{vp} - 0,044 T_{tc}$$

Donc

$$P = \frac{1}{1 + \exp[-(1,12 + 0,052 T_{vp} - 0,044 T_{tc})]}$$

B) Modèle avec coût

$$\text{Logit (P)} = \log (P/1-P) = 1,25 + 0,059 T_{vp}^R - 0,044 T_{tc} - 0,168 C_{tc}$$

Donc

$$P = \frac{1}{1 + \exp[-(1,25 + 0,059 T_{vp}^R - 0,044 T_{tc} - 0,168 C_{tc})]}$$

5. VALIDATION DU MODELE RETENU

La validation du modèle de choix devrait être effectuée sur des lignes existantes de façon agrégée.

Un recueil de données auprès de la RATP et la SNCF a été effectué sur les trafics voyageurs de trois lignes de rocade existantes (T1 Tramway Saint-Denis - Bobigny, T2 Tram Val-de-Seine, liaison St Quentin-en-Yvelines - La Défense). Les détails sont fournis en annexe 5.

Nous n'effectuerons le test que sur une seule ligne. On a choisi la **ligne Saint-Quentin - La Défense (SQDF)** car c'est une ligne longue distance, qui a une zone d'influence très large, et qui induit peu de trafic de cabotage.

On attend du modèle de choix modal qu'il nous retrouve la part des reportés de la voiture sur le trafic de voyageurs de la ligne SQDF.

▪ Comptages et enquête clientèle SNCF sur la ligne SQDF

La dernière enquête clientèle de la SNCF montre que, à la journée, 8% des usagers prenaient avant leur voiture. Or comme nous travaillons en simulation à la période de pointe du matin, cette part est plus importante. Si on considère dans le cas extrême que le taux de reportés à l'heure creuse est de 0%, et sachant que la part de trafic sur 2 heures de pointe du matin par rapport à la journée est de 55%, alors le taux de reportés est de 15% (ce chiffre permet de retomber sur 8% en moyenne à la journée). Si on prend plutôt un taux de 3% en période creuse, alors le taux de reportés en période de pointe du matin est de **12,1%**. C'est ce chiffre que nous retiendrons pour la suite.

La ligne La Verrière-La Défense dans le sens ouest-est draine 19 860 voyageurs (somme des montants depuis la Verrière jusqu'à Puteaux). Le trafic à la période de pointe (2 heures le matin) est de 11 519 voyageurs, dont 12,1% de reportés de la voiture, soit environ **1400 voyageurs**.

Il y a 2 scénarios : l'un sans le projet SQDF, l'autre avec projet SQDF

▪ **Matrice de déplacements tous motorisés en 488 zones calculée par le modèle**

Pour les 2 scénarios, on part de la fin de l'étape de distribution, c'est-à-dire de la matrice jour de déplacements tous motorisés. Pour cela, le modèle de demande de l'IAURIF est utilisé en l'état, sans adaptation particulière pour cette étude. Il y a donc quelques différences mineures dans la définition des motifs de déplacement et qui ne devraient pas avoir d'incidence sensible sur le résultat final. En effet, on travaille en période de pointe du matin et le domicile est le motif origine principal.

Le modèle de l'IAURIF considère 9 motifs de déplacements (voir document [9]). Pour les 3 motifs de déplacements traités dans le cadre de cette étude, la correspondance est la suivante :

EGT	Modèle de demande de l'IAURIF
D->T	D->T
D->E	Tous motifs->E (sauf école primaire)
D->AM	Tous motifs -> AM (sauf loisirs)

L'EGT 97-98 donne les coefficients de passage de la journée à la période de pointe qui sont pris en compte dans le modèle :

- D->T : 58% , 1 702 518 déplacements
- D->E : 74% , 883 744 déplacements
- D->AM : 18% , 992 652 déplacements

Le modèle reconstitue les trois matrices tous motorisés (une par motif), sur 2 heures de pointe du matin :

- D->T: 1 767 893 déplacements
- tous motifs -> E (sauf primaire) *assimilé à D->E* : 855 832 déplacements
- tous motifs -> AM (sauf loisirs) *assimilé à D->AM* : 1 108 167 déplacements

▪ **Matrices des temps de déplacement par mode en 488 zones à l'heure de pointe du matin**

Pour chaque scénario, la matrice des temps de parcours pour le mode TC est calculée. Quant à la matrice des temps VP, elle est identique pour les 2 scénarios.

Les deux missions ajoutées au scénario sans projet sont les missions La Verrière-La Défense en pointe et contre-pointe du matin.

En format MinUTP, on a :

```

$ mission 21 : La Verrière → La Défense
LINE BSL95*,MO=3,H1=15,ST=0.75,RT=35,VC=05,N=1645*-1644*-1643*-1642*
-1641*-1724-1722-1723-1713*-1712*-1711*-1710*-1709*
$ mission 20 : La Défense → La Verrière
LINE CSL95*,MO=3,H1=15,ST=0.75,RT=36,VC=05,N=1709*-1710*-1711*-1712*
-1713*-1723-1722-1724-1641*-1642*-1643*-1644*-1645*

```

▪ Calcul des matrices de parts TC en 488 zones

Nous avons utilisé les formules calculées dans la partie 3.

Pour le motif D->T, les formules A (temps seulement) et B (temps + disposition d'un parking gratuit + taux de motorisation) sont utilisées.

Pour les motifs D->E et D->AM, les formules A (temps seulement) ont été utilisées.

Pour le cas de la formule B du D->T, nous avons utilisé des valeurs moyennes par couronne.

Pour α_m , la source est l'EGT 97-98.

	Paris	Petite couronne	Grande couronne
DPARK (zone destination)	0,2	0,5	0,9
α_m (zone origine)	0,52	0,94	1,26

▪ Calcul des matrices de déplacements TC en 488 zones

Il suffit pour chaque motif de multiplier case à case la matrice de déplacements tous motorisés par la matrice de parts TC.

On obtient donc une matrice de déplacements TC par motif.

▪ Résultats (sur deux heures de pointe le matin)

Le nombre de reportés de la VP vers la nouvelle ligne SQDF correspond à la différence des matrices de chaque scénario. On devrait en principe trouver environ 1400 reportés.

Cas 1 : formule A pour le D->T

	sans projet	avec projet	delta
D->T	799 274	800 776	1 502
D->E	456 053	456 350	297
D->AM	239 503	239 858	355
	1 494 830	1 496 984	2 154

Le modèle trouve donc 2154 reportés.

Cas 2 : formule B pour le D->T

	sans projet	avec projet	delta
D->T	897 978	899 052	1 074
D->E	456 053	456 350	297
D->AM	239 503	239 858	355
	1 593 534	1 595 260	1 726

Le modèle trouve 1726 reportés.

Comme on pouvait s'y attendre, le cas 2 est plus proche des comptages. Le modèle surestime quelque peu le nombre de reportés (+326, soit + 23%) mais l'ordre de grandeur est respecté.

Dans le cas A, la différence est sensible : + 754 reportés (+54%) par rapport au nombre donné par l'enquête clientèle . La formule A pour le motif D->T n'est pas satisfaisante. Il vaut mieux prendre en compte les variables DPARK (disposition d'un parking gratuit sur le lieu de travail) et ε_m (taux de motorisation), même si l'on doit se contenter de valeurs moyennes par couronne sujettes à caution, que pas du tout.

6. CONCLUSION

Cette étude a essayé d'aborder le problème de la modélisation du choix modal en Ile-de-France d'un point de vue pratique. Elle a abouti à la mise en évidence de formules mathématiques donnant des parts de marché en transport en commun de zone à zone. Le modèle de choix modal a donné relativement satisfaction dans la reconstitution des reports modaux de la ligne Saint-Quentin - La Défense.

Les **caractéristiques de ces formules** sont les suivantes :

- elles sont calibrées sur la période de pointe du matin
- elles dépendent de trois motifs de déplacement : domicile-travail, domicile-école, domicile-autres motifs
- elles sont indépendantes de tout type de liaison. On s'affranchit ainsi d'un modèle classique de choix modal avec des courbes de répartition modale par type de liaison (par exemple, de zone carte orange à zone carte orange)
- elles prennent en compte principalement comme variables les temps de parcours par mode de "porte à porte", la décomposition en portions de temps étant inutile. D'autres variables ont été introduites suivant le motif : le coût de déplacement par mode, le temps de recherche de parking, la disposition d'un parking gratuit sur le lieu de travail et le taux de motorisation par ménage. Pour chaque motif, plusieurs formules sont proposées : elles se différencient par la présence ou non des différentes variables.

En ce qui concerne la **méthodologie d'étude**, ce rapport décrit la démarche de spécification d'un modèle désagrégé de choix modal. La phase de calibrage des paramètres d'un modèle logit bimodal est relativement simple grâce aux logiciels de statistique du marché (un tableur ne suffit pas), par rapport à la phase de constitution de la base d'observations. Pour constituer cette base, on a à la fois besoin de données d'enquêtes et de données de simulation, et toute la difficulté consiste à mélanger ces deux types de données. L'idée principale consiste à adjoindre à chaque observation individuelle une pseudo-observation simulée du mode concurrent et à ne garder que les observations individuelles dont les temps déclarés sont proches des temps moyens simulés. Outre la disponibilité de ces données, il faut dire aussi que les conditions de réussite de développement d'un modèle désagrégé sont d'ordres humain et organisationnel. En effet, les compétences requises sont diverses : connaissance et capacité à exploiter l'EGT, pratique d'un logiciel de statistiques, pratique d'un modèle de prévision de trafic, pratique d'un système d'information géographique,

programmation. Or dans une organisation, ces compétences sont en général complètement dispersées.

La **limite principale** du modèle de choix modal développé dans le cadre de cette étude réside dans la non-différentiation des individus. On a travaillé sur des individus "moyens". Or il est évident qu'une segmentation de la population permettrait d'affiner les formules du modèle, en particulier pour le motif domicile-école. Ceci pose alors un autre problème : segmenter la population revient à introduire des variables de caractéristiques individuelles (statut, revenu, ...). Ceci nécessite donc des enquêtes ménages beaucoup plus lourdes que la dernière EGT, ceci afin de trouver des échantillons d'observations en nombre significatif dans chaque segment de population. De plus, les variables de caractéristiques individuelles ne sont jamais simples à projeter dans le futur.

Les **suites** possibles à donner à cette étude sont nombreuses :

- Il serait intéressant d'étudier la transposabilité des formules d'un modèle à l'autre. Autrement dit, sont-elles vraiment indépendantes du niveau de découpage d'un modèle ? En appliquant la même méthodologie, on pourrait refaire le travail sur le modèle de la DREIF en 1180 zones, élaborer les formules de parts modales TC et les comparer à celles de la présente étude.
- Il faudrait, à partir des formules, effectuer des calculs d'élasticité du report modal aux différentes variables d'offre, en particulier la vitesse en TC et la disposition d'un parking gratuit à destination.
- Il conviendrait de valider le modèle sur les deux autres rocade existantes de proche banlieue : tramway Saint-Denis - Bobigny et tramway Val-de-Seine.
- Il faudrait ensuite le tester sur les fameux projets de tangentielle.
- Sur les aspects théoriques, il faudrait approfondir la notion de coût monétaire de déplacement en introduisant la notion de valeur du temps individuelle par motif. Au niveau de la définition des coûts VP, il faudrait s'intéresser au péage.

Finalement, cette étude pourrait ouvrir la voie en Ile-de-France à deux champs d'études : **sur le plan méthodologique**, l'affinement de la modélisation désagrégée du report modal dans les modèles globaux ; **sur le plan pratique**, l'introduction des formules présentées dans ce rapport dans les études de projet de transports collectifs qui s'avèreraient éventuellement nécessaires pour la révision du Contrat de Plan à mi-parcours.

ANNEXE 1

CONTEXTE DE L'ETUDE EN ILE-DE-FRANCE

Toute évaluation socio-économique d'un projet de transport nécessite l'utilisation d'un modèle de trafic. Or les modèles en milieu urbain utilisés à l'heure actuelle ne donnent pas toujours satisfaction alors que leur fiabilité conditionne largement celle des évaluations. Dans ce contexte, le rapport Boiteux sur l' « évaluation de la politique des transports en Ile-de-France » (février 1999) a identifié six points qui méritent une attention particulière :

- la prise en compte de la diversité des usagers et de leurs motifs de déplacements,
- la modélisation de la congestion sur un tronçon routier lorsque l'on sort du domaine de validité des courbes débit-vitesse,
- la modélisation des effets de congestion ou de décongestion engendrés par une liaison nouvelle sur le réseau secondaire,
- la modélisation du trafic induit à chaque étape et la rétroaction sur l'étape d'affectation,
- la prise en compte du trafic poids lourds,
- et enfin, la modélisation du choix modal entre voiture particulière et transport en commun.

Comme le montre une rapide revue de la littérature, de nombreux travaux de recherche tentent d'apporter en permanence des améliorations sur tous ces points. En particulier, dans un processus de modélisation qui comporte d'une manière ou d'une autre les « quatre étapes » (génération, distribution, choix modal et affectation), la phase de choix modal qui est encore grandement perfectible a fait l'objet de nombreuses études et recherches aux Etats-Unis. En Ile-de-France, quatre études récentes ont traité de ce sujet (voir bibliographie, de [5] à [8]).

Il semble qu'à la lecture de ces rapports, aucun aspect n'ait échappé au traitement du choix du mode de transport : statistiques descriptives, analyses approfondies, détermination et explication des facteurs principaux, théorisations plus ou moins poussées avec parfois des approches sociologiques ou épistémologiques, modélisations en tous genres (logit, probit, dogit). Il est alors étonnant de constater que malgré toutes ces recherches, l'étude de référence pour les modélisateurs concernant le choix du mode en Ile-de-France reste encore celle réalisée en 1966 par l'IAURP (« Les transports urbains et leurs usagers en région de Paris : choix du moyen de transport par les usagers »). Cette étude réalisait pour la première fois en France une recherche de base sur les coefficients de pénibilité à affecter aux divers facteurs du choix modal.⁸ Ces recherches anciennes n'ont jamais été réactualisées.

⁸ Il serait d'ailleurs intéressant de s'interroger sur les raisons de la position de référence occupée par cette étude depuis 33 ans. Sans vouloir trop développer (cela dépasserait notre propos), on peut invoquer des raisons

Aujourd'hui, force est de constater que les modèles de choix modal des quatre principaux modèles agrégés de trafic utilisés en Ile-de-France sont tous différents, que ce soit :

- « Chaîne » et « Modus » : les modèles de la DREIF,
- « Modèle global » : le modèle de l'IAURIF,
- « Global » : le modèle de la RATP,
- « Magellan » : le modèle de MTI Conseils pour la SNCF.

Il existe bien en Ile-de-France un modèle désagrégé de choix modal, IMPACT, développé par HCG pour la RATP et basé sur l'EGT91. Mais ce modèle sert principalement à des études d'évaluation de politiques globales à court terme. Il n'est pas utilisé dans le modèle global de la RATP (voir document [11]). Le STP dispose aussi d'un modèle entièrement désagrégé, appelé ANTONIN. Ce modèle présente des aspects très innovants mais n'a pas encore été utilisé à l'heure actuelle dans des études.

Rappelons que l'un des problèmes majeurs posé par l'étude d'une nouvelle ligne de rocade de transport en commun en Ile-de-France (Tangentielle ou Orbitale) consiste à estimer la clientèle potentielle. Celle-ci est constituée de trois parts :

- une part provient d'un changement d'itinéraire de transport en commun ;
- une part provient d'une clientèle purement « induite », c'est-à-dire celle qui, sans la nouvelle ligne, n'effectuerait pas de déplacement (autrement dit, la demande latente);
- une part provient d'un report modal de la voiture particulière sur les transports en commun.

Comment sont calculées ces trois parts ? La première part est déduite d'une affectation avec puis sans projet. En l'état actuel des modèles en milieu urbain, la deuxième part est incalculable. Elle suppose en effet que l'étape de génération des déplacements intègre la possibilité de prendre en compte l'incidence d'une nouvelle offre sur les taux de mobilité, ce qui n'existe pas aujourd'hui. Cette part est donc toujours ajoutée au final de façon forfaitaire (par exemple 10% de trafic en plus). Reste la troisième part, celle qui concerne notre étude. Seul un modèle de choix modal fiable et de qualité permet de calculer le report modal de la voiture particulière sur une nouvelle liaison de transports en commun.

historiques (l'IAURIF était dans les années 60 plus orienté vers la recherche que l'IAURIF et la DREIF aujourd'hui et donc pouvait consacrer plus de temps), des raisons culturelles (pourquoi remettre à plat la seule étude qui fournisse des résultats chiffrés prêts à utiliser, et qui sont indiscutables car élaborés par des sommités de l'époque) et des raisons organisationnelles (aujourd'hui les équipes chargées du développement et de la maintenance des modèles dans chaque organisme comportent plutôt des praticiens que des théoriciens).

ANNEXE 2

**POINTS ESSENTIELS DE LA PARTIE "*PREVISIONS DE TRAFIC A LA MISE EN SERVICE*" (§ III.1) DE L'EXPERTISE DU CGPC SUR LA TANGENTIELLE NORD
(3 décembre 1999)**

Le Conseil général des ponts et chaussées a comparé les résultats de 2 études de trafic, l'une réalisée par la SNCF, l'autre par la DREIF, sur la tangentielle nord. Il n'a pas cherché à élucider les mécanismes de fonctionnement des deux modèles utilisés, ce qui aurait peut-être permis de mieux comprendre les différences importantes de résultats. Il se contente de faire quelques observations sur les résultats de la ligne Sartrouville - Noisy-le-Sec :

- 1) **Le trafic total à l'heure de pointe du matin** prévu par la DREIF en 2005 est pratiquement le double de celui de la SNCF (38000 voyageurs contre 19400). Le CGPC explique cette différence par le nombre important d'usagers qui ne gagneraient qu'un temps minime (1 ou 2 minutes) dans le modèle de la DREIF, celui-ci affectant la demande de déplacement pour une O/D donnée sur un chemin unique en "tout-ou-rien". Si on retire ces usagers qui gagnent un temps négligeable, le CGPC affirme que l'on peut considérer les trafics totaux des deux modèles comme concordants (environ 19000 voyageurs).

- 2) Sur l'évaluation du **nombre d'usagers reportés de la voiture particulière**, la discordance entre les deux modèles est encore plus importante. Les deux modèles partent de la même matrice de déplacements tous motorisés (celle de la DREIF), puis chacun applique son propre modèle de répartition modale. C'est donc à ce stade qu'apparaît la différence d'évaluation. La SNCF trouve deux fois plus de reportés (7348, soit 38% du total) que la DREIF (3700 soit 10% du total). Les deux modèles de choix modal utilisés utilisent chacun des courbes de partage modal, mais très peu d'explications sont données sur la façon dont elles ont été élaborées et calées. Sur la base de comparaisons avec des lignes de rocade existantes (ligne La Verrière-La Défense et ligne St Denis-Bobigny), le CGPC pense que le chiffre de 38% de transférés de la VP est surestimé. Il propose donc de retenir l'évaluation de la DREIF.

- 3) La **part de trafic induit** (demande latente) n'est estimée par aucun des deux modèles. Le CGPC propose de prendre 500 voyageurs à l'heure de pointe du matin.

En définitive, le CGPC propose de retenir les valeurs suivantes :

- le total calculé par la SNCF, soit 19 400 voyageurs
- le report modal de la DREIF, auquel s'ajoutent les induits : 4 200 voyageurs.

ANNEXE 3

- **ZONAGE DU MODELE GLOBAL DE L'IAURIF**
- **DECOUPAGE MORPHOLOGIQUE REGIONAL**

ANNEXE 4

QUELQUES MODELES DE CHOIX MODAL AUX ETATS-UNIS

Le modèle désagrégé de choix modal est devenu dans la pratique la référence aux Etats-Unis (voir [1] chapitre 10). En particulier, une variante est actuellement très en vogue : le modèle logit emboîté. Le terme « emboîté » est employé pour décrire un ensemble de choix où les alternatives sont classées de façon hiérarchique et où le choix d'une alternative implique que toutes les alternatives classées plus bas ont été choisies aussi. Le modèle simule donc un processus de choix multidimensionnel où une hiérarchie logique existe dans le processus de décision, utilisant des probabilités conditionnelles.

Nous allons d'abord décrire deux modèles de façon détaillée puis faire une revue de la littérature décrivant la pratique de la modélisation du choix modal dans quelques métropoles américaines.

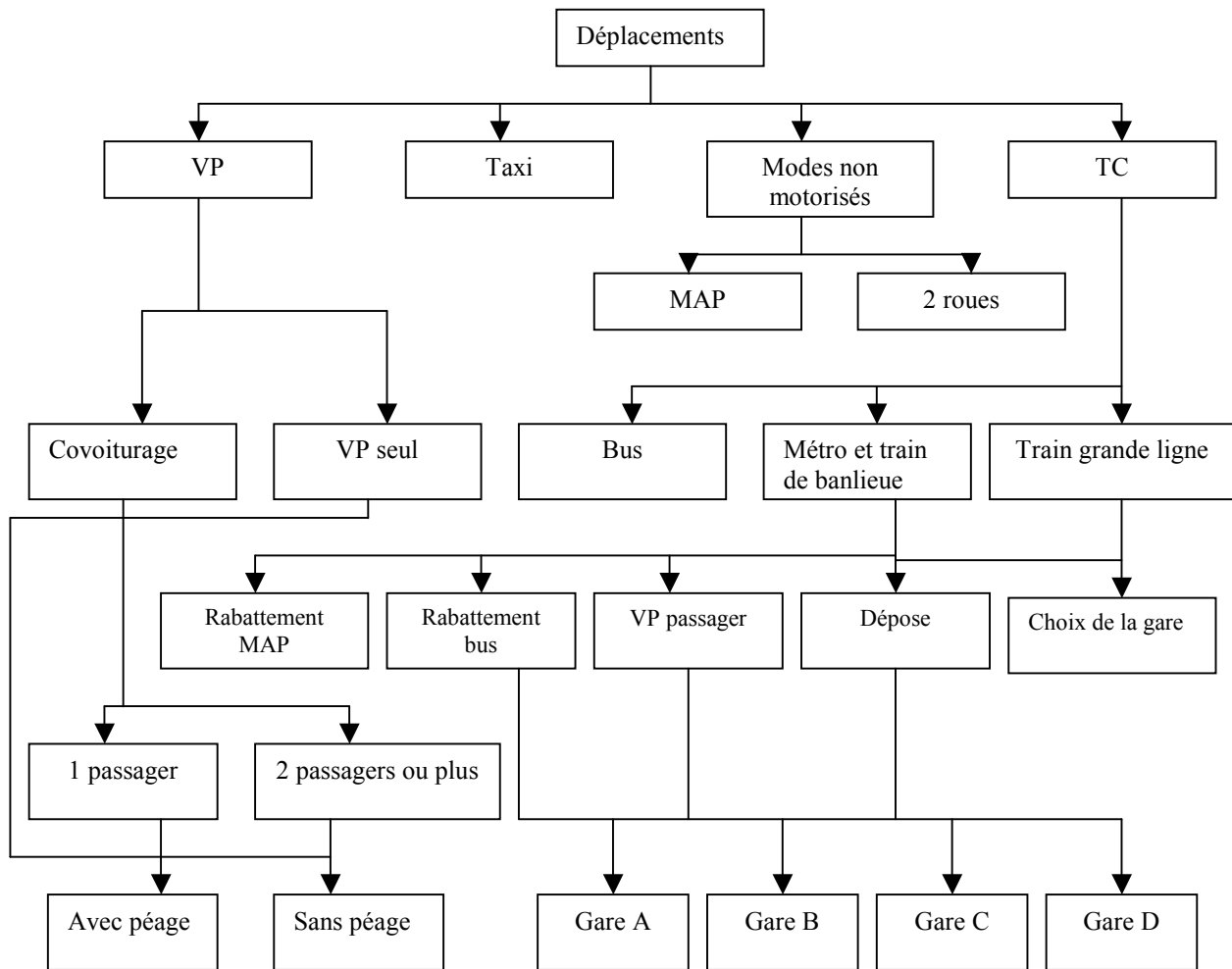
1. New York

La région comprend la ville de New York (qui elle-même est constituée de 5 comtés) et cinq comtés : Nassau, Putnam, Suffolk, Rockland, Westchester. La surface de la région est de 6100 km² pour 11,3 millions habitants et 5 millions d'emplois (données 1997).

Le modèle de trafic de la Région de New York est développé par le bureau d'études régional **New York Metropolitan Transportation Council (NYMTC)**. L'étape de choix modal est un modèle logit emboîté.

Il estime les quatre modes primaires : TC, VP, Taxi et modes non motorisés. Sous le mode TC, le modèle distinguera les déplacements en métro/train de banlieue, les déplacements par train grande ligne et les déplacements en bus. Les modes ferrés auront un troisième niveau, qui est le mode de rabattement aux gares (MAP, bus, accompagnement, dépose). Pour les trois derniers modes, il y a encore une « boîte » qui représente le modèle de choix de la gare, parmi quatre possibilités. Ainsi pour un déplacement donné, il y a 65 possibilités de modes en transport en commun. Quant au mode VP, il aura un niveau secondaire séparant les déplacements en VP seul et en déplacements de covoiturage. Ces derniers auront eux-mêmes deux alternatives : un passager ou deux passagers par voiture. Pour chacun des modes VP, il y aura une « boîte » qui sépare les déplacements en fonction des itinéraires avec et sans péage. Le taxi quant à lui n'a aucune alternative. Les modes non motorisés auront un niveau secondaire de décision séparant la MAP et le vélo.

La figure ci-dessous montre la structure hiérarchique du processus de choix modal :



Il y a un modèle de choix modal pour chaque motif.

La probabilité qu'un déplacement d'un individu du groupe g soit effectué en mode m est exprimée par le modèle logit suivant :

$$P_{g,m} = \frac{\exp [U_{g,m}(x_{g,m})]}{\text{SUM}_i \{ \exp [U_{g,i}(x_{g,i})] \}}$$

Avec :

$x_{g,m}$: attributs du mode m décrivant son utilité pour les individus du groupe g

$U_{g,i}(x_{g,i})$: utilité du mode i pour les individus du groupe g

Typiquement, la fonction d'utilité pour chaque alternative prend la forme suivante :

$$U_{g,i}(x_{g,i}) = a_i + b_i \text{NDS}_i + c_{g,i} \text{SE}_g + d_i \text{DEP}$$

Avec :

- NDS_i : variables décrivant le niveau de service assuré par le mode i
- SE_g : variables décrivant les caractéristiques socio-économiques du groupe g
- DEP : variables caractéristiques du déplacement
- a_i : constante spécifique du mode i qui prend en compte tous les facteurs non modélisés (confort, sécurité, ...)
- b_i : vecteur de coefficients mesurant l'importance de la variable NDS_i
- $c_{g,i}$: vecteur de coefficients mesurant l'importance de chaque SE_g
- d_i : vecteur de coefficients mesurant l'importance de chaque DEP

Parmi les caractéristiques de la variable DEP , il y a le coût de stationnement pour le mode VP. Le modèle estime le coût pour chaque zone en fonction de la densité d'occupation du sol (tel que le nombre d'emploi par km^2). Pour les TC, une attention particulière est portée sur la modélisation des différents modes de rabattement des centroides sur les gares à l'origine et à la destination : rabattement MAP de courte distance, rabattement MAP de longue distance, rabattement en voiture (à l'origine seulement). La combinaison deux à deux donne donc six modes de rabattement différents pour un déplacement en TC donné. Les impédances utilisées dans les utilités dépendent du motif de déplacement. Pour les motifs obligés (travail et école), c'est l'impédance à l'heure de pointe qui est utilisée. Pour les autres motifs, c'est l'heure creuse.

2. Chicago

Située au nord est de l'Illinois, la région comprend la ville de Chicago, six juridictions dans le comté de Cook et cinq autres comtés (DuPage, Kane, Lake, McHenry et Will). En 1997, la population était de 7,5 millions d'habitants, le nombre d'emplois de 4 millions sur une surface de 10000 km^2

Le modèle régional est développé et maintenu par le bureau d'études régional, le **Chicago Area Transportation Study Policy Committee (CATS)**.

Le modèle de choix modal est un logit binomial. Les variables prises en compte sont des variables caractérisant l'offre. Le temps de parcours par le mode m pour chaque O-D est décomposé en :

- T^m_1 : temps de roulement
- T^m_2 : temps d'attente (en TC seulement)
- T^m_3 : temps de correspondance (en TC seulement)
- T^m_4 : temps de rabattement à la gare ou temps de recherche de parking

Pour caractériser l'impédance, on ajoute le coût monétaire T^m_5

La probabilité qu'un déplacement soit effectué en TC vaut :

$$P^{TC} = \frac{\exp[\sum \alpha_i (T^{VP}_i - T^{TC}_i) + \beta]}{1 + \exp[\sum \alpha_i (T^{VP}_i - T^{TC}_i) + \beta]} \quad (i = 1 \dots 5)$$

Les coefficients α_i et β dépendent du motif de déplacement et de la destination.

Il y a exactement trois vecteurs de coefficients :

Variable	Coefficients pour D-T		Coefficients pour autres motifs
	Destination hors centre-ville	Destination centre-ville	
Temps roulement	0,0186	0,0159	0,0114
Temps attente	0,0811	0,0173	0,0610
Temps correspond.	0,0399	0,0290	0,0589
Temps rabattement	0,0584	0,0468	0,0663
Coût	0,0072	0,0085	0,0329
Coefficient modal	-0,4983	-0,6059	-0,2726 (D-autre)
			-0,8072 (autre-autre)

Ces coefficients ont été calés à partir de valeurs du temps par composante de temps de parcours.

Exemple d'application pour une O-D avec destination centre-ville et motif D-T :

Soit une O-D avec les données de temps de parcours suivantes :

Variable	VP	TC	Différence	Coefficient	Produit
Temps roulement	25	45	-20	0,0159	-0,318
Temps attente	0	5	-5	0,0173	-0,0865
Temps correspond.	0	10	-10	0,0290	-0,290
Temps rabattement	5	7	-2	0,0468	-0,0936
Coût	200	100	100	0,0085	0,850
Coefficient modal				-0,6059	-0,605
Σ					-0,544
$\exp(\Sigma)$					0,580
$1+\exp(\Sigma)$					1,580
Probabilité TC					0,367

En ce qui concerne les données, le bureau d'études CATS a mis en lumière le problème de l'utilisation de valeurs moyennes. En effet, une source importante d'imprécision dans le modèle de choix modal est l'utilisation de valeurs moyennes. La solution adoptée a alors été d'identifier les zones qui sont affectées par des valeurs moyennes et de « convertir » ces valeurs moyennes en valeurs individuelles. Cette approche fait appel à la simulation par la méthode de Monte Carlo. Celle-ci permet de sélectionner de façon aléatoire une valeur pour une variable donnée connaissant la distribution des valeurs.

La méthode de Monte Carlo est utilisée ici pour estimer quatre variables :

- les temps de rabattement à l'origine et à la destination pour le mode TC,
- le revenu des usagers,
- le coût de parking.

3. Autres régions métropolitaines

Les descriptions suivantes concernent le motif domicile-travail.

3.1. San Francisco

Cette région, appelée Baie de San Francisco, comprend neuf comtés : Sonoma, Napa, Solano, Marin, Contra Costa, Alameda, San Mateo, Santa Clara et San Francisco.

La surface de la région est de 18700 km² pour 6,7 millions habitants et 3,3 millions d'emplois (données 1997).

Le modèle de trafic de la Région est développé par le bureau d'études régional **Metropolitan Transportation Council (MTC)**.

1965-1969 : courbes de partage modal

Le premier modèle de choix modal utilisé dans la région de San Francisco (Baie de San Francisco) a été développé au milieu des années 60. La méthode des courbes de partage modal était typique de la pratique entre 1950 et 1970. Le choix modal ne comportait que deux alternatives : VP ou TC. Les courbes, estimées « à la main », étaient basées sur le ratio du temps de parcours TC sur le temps de parcours VP. Plus ce ratio était élevé, moins forte était la probabilité d'utiliser les TC.

Il y avait en tout 6 courbes, fonction de la densité de population du lieu de résidence (3 classes de densité) et de la densité d'emplois du lieu de destination (2 classes de densité).

Le calage de ces courbes a été effectué à partir de l'enquête-ménages de 1965 portant sur 20500 ménages.

La fin des années 60 a vu l'émergence d'un nouveau modèle de choix modal : le modèle de choix désagrégé.

1972-1974 : modèle logit

MTC a été le premier aux Etats-Unis à innover avec un modèle de choix discret. Un modèle logit binomial a été utilisé, séparant les déplacements VP et TC. La fonction d'utilité du mode TC comprend le temps de roulement et le temps hors roulement pondéré par un facteur 2.5 , ainsi que le coût TC en équivalent temps. La fonction d'utilité du mode VP comprend le temps de roulement, le temps de recherche de parking, le péage et le coût du parking. Il semble que le modèle dispose d'un module qui sépare les déplacements TC suivant le mode de rabattement (MAP ou VP). Le calage a été aussi effectué avec l'enquête-ménage de 1965. Le logiciel TRANPLAN a été utilisé pour calculer les temps de parcours VP et TC sur des graphes.

1974-1979 : grand projet de modélisation initié par l'Université de Berkeley

UTDFP (Urban Travel Demand Forecasting Project) était un projet de recherche s'étalant sur plusieurs années, financé par la Fondation Nationale de Science, le Département fédéral de Transport et la fondation Sloan. Le Directeur du projet était le Pr McFadden. L'un des objectifs principaux de UTDFP était d'appliquer le modèle désagrégé pour tester le projet de transport en commun express BART dans une aire métropolitaine. Plusieurs enquêtes-ménages ont été effectuées : sur 213 ménages en 1972, sur 1724 ménages en 1973, sur 258 individus en 1975 (enquête téléphonique).

Le modèle de choix modal était un modèle logit multinomial. Il prenait en compte les modes VP, MAP+TC, VP+TC. Le mode VP est ensuite séparé en VP seul et en VP avec passagers. Les variables prises en compte sont les temps de parcours et les coûts, des variables socio-économiques et d'occupation du sol (revenu du ménage, taille, nombre de conducteurs, taux de motorisation, densité de population).

Il est intéressant de noter qu'à l'époque, les enquêtes-ménages utilisées par le UTDFP portaient sur des échantillons très faibles. MTC a développé un programme de génération de population synthétique.

1974-1978 : projet de développement d'un nouveau modèle de trafic (TMDP)

En 1974, MTC a lancé un appel d'offres pour développer un nouveau modèle de trafic pour la région de San Francisco. Une équipe de consultants a été retenue, comprenant en particulier le Pr Moshe Ben Akiva. Le développement a pris 3 ans.

Le nouveau modèle marque un tournant dans l'histoire de la modélisation de trafic. Basé sur la thèse de Ben Akiva, le modèle de MTC a été le premier à incorporer la théorie du choix de mode désagrégé et un modèle logit emboîté dans un système opérationnel.

Le modèle logit de choix modal comporte un modèle d'estimation de taux de motorisation. Le nouveau modèle, appelé MTCFCAST, est extrêmement documenté. Il a fait l'objet d'innombrables articles de recherche pendant plusieurs années.

1980-1988 : programme de mise à jour du modèle

En 1981, MTC a mené la plus grande enquête-ménage depuis 1965. 6200 ménages ont été interrogés. Les résultats de cette enquête ont permis de recalibrer le modèle de choix modal,

entièrement en interne, entre 1984 et 1987. Le nouveau modèle de choix modal comporte quatre modes principaux : VP seul, VP+1 passager, VP+2 passagers ou plus, TC. Pour la première fois, un modèle du commerce est utilisé : UTPS (Urban Transportation Planning System).

1990-1997 : nouveau programme de mise à jour

En 1990, MTC a conduit une nouvelle enquête-ménage portant sur 9359 ménages. Entre 1994 et 1997, MTC s'est engagé dans un processus d'adaptation de son modèle à partir de l'enquête de 1990 et du logiciel MINUTP.

Le modèle de choix modal est un modèle de type logit emboîté avec sept alternatives: VP seul, VP+1 passager, VP+2 passagers ou plus, MAP+TC, VP+TC, vélo, MAP. Au niveau supérieur, les déplacements sont séparés en mode motorisé, MAP et vélo. Au niveau intermédiaire, les déplacements motorisés sont séparés en : VP seul, VP+1 passager, VP+2 passagers ou plus, TC. La boîte inférieure sépare les déplacements TC en MAP+TC et VP+TC.

Le modèle de choix est appliqué à chaque motif, à part le motif domicile-école primaire qui est un logit multinomial. Les variables incluent le revenu du ménage, la taille du ménage, la distribution des individus par tranche d'âge et la densité de population.

Une particularité importante du modèle de choix modal de BAYCAST est qu'il traite aussi bien la période de pointe que la période creuse. Dans les versions précédentes, seule l'heure de pointe était prise en compte. Cette amélioration du système permet de rendre sensible le choix du mode à la période de la journée.

Tous les modèles de choix de mode incorporent les modes non motorisés : vélo et MAP. Les temps de parcours en vélo ou à pied sont calculés sur un réseau spécifique non routier.

3.2. Portland (Oregon)

Le bureau d'études régional **METRO** de la région de Portland est un précurseur en matière d'innovations dans la modélisation de trafic en milieu urbain. Le modèle de choix modal de METRO est le premier à prendre en compte les modes non motorisés. En effet, une étape préliminaire du modèle consiste à séparer par un modèle logit binomial les déplacements motorisés des déplacements marche à pied et vélo.⁹ La deuxième étape est un modèle logit multinomial avec quatre choix, appliqué aux ménages motorisés : VP seul, VP+passagers, MAP+TC, VP+TC. Pour les ménages non motorisés, une simple répartition est appliquée :

22.4% pour le mode VP+passagers, 77.6% pour le mode TC. Les variables prises en compte sont : le temps de roulement, le temps hors roulement, le coût, la distance, la densité d'emplois à la destination, la densité de population à l'origine, le type de zone à la destination (centre-ville, autres).

3.3. Los Angeles (Californie)

Le dernier modèle de choix modal du bureau d'études régional **SCAG (Southern California Association of Governments)** a été entièrement sous-traité en 1993. Les enquêtes utilisées sont l'enquête-ménage de 1991, ainsi que des enquêtes clientèles spécifiques réalisées par les opérateurs de transport en commun.

La base d'observations servant au calage comporte 10946 déplacements.

Il y a 12 modes pour chaque motif : MAP, vélo, passager de voiture, VP seule, VP+1 passager, VP+2 passagers ou plus, bus direct, bus omnibus, MAP+TC, bus+TC, VP+TC, autobus scolaire. En fait, un ou plusieurs modes sont éliminés dès le départ pour certains motifs parce que les déplacements effectués sont marginaux.

Le modèle logit emboîté est utilisé pour chaque motif (SCAG avait testé le modèle logit multinomial, mais sans résultat). Il a identifié quatre formes candidates de structures emboîtées et l'estimation a été faite avec le programme ALOGIT.

3.4. Washington DC

Le modèle, développé par le **WCOG (Washington Council of Governments)**, est constitué de quatre modèles logit, un pour chaque motif de déplacement de voyageur. Cinq modes sont traités : TC, VP seul, VP avec 1 passager, VP avec 2 passagers, VP avec plus de 2 passagers.

C'est un modèle logit multinomial par étapes. L'étape supérieure sépare les déplacements en VP seul, VP avec passagers, TC. L'étape inférieure sépare les déplacements de covoiturage par taux d'occupation : 2, 3, 4 personnes ou plus par véhicule. La demande est segmentée par le critère de taux de motorisation (0,1,2 ou plus). Ce n'est pas un modèle dit

⁹ Aux Etats-Unis, un mode non motorisé est soit la MAP, soit le vélo alors qu'en France, il n'y a que la MAP.

logit emboîté car l'utilité du taux d'occupation n'affecte pas l'utilité de covoiturage dans l'étape supérieure. Le modèle de choix modal a été calé avec l'enquête-ménages de 1994.

Les facteurs considérés dans le modèle incluent l'accessibilité au transport en commun, le taux de motorisation, la proximité de routes comportant des voies affectées au covoiturage, les dépenses et le temps de parcours. Le coût considéré ici est le coût marginal, c'est-à-dire le prix du trajet en TC, le coût de l'essence et le coût de stationnement en VP. Le temps représente le temps total de parcours, y compris par exemple le temps de recherche de parking à la destination. Tous ces déterminants de choix modal sont introduits dans une équation qui estime la probabilité pour chaque déplacement d'être effectué dans un des cinq modes.

Parmi les facteurs les plus importants dans le choix modal, il y a le coût du stationnement et les temps de rabattement. Le coût de stationnement moyen peut être estimé en fonction de la densité d'emploi de la zone.

Pour les usagers des transports en commun, de nouvelles hypothèses sont faites sur leur utilisation du bus ou du rail. Par exemple, seules les personnes situées à moins d'un km d'une station sont supposées marcher jusqu'à la gare. Le modèle suppose que le reste des usagers de transport en commun se rabattent en VP ou sont déposés à la gare.

Pour l'estimation du modèle logit, plus de 80 formes candidates ont été estimées en utilisant le programme ALOGIT. C'est un processus heuristique, empirique, bien connu des praticiens.

Une variable caractérisant l'occupation du sol est prise en compte dans les définitions des utilités des modes VP et TC :

$$V_{\text{mos}} = (\text{hhpopd} * \text{nempd}) / (\text{hhpopd} + \text{nempd})$$

Avec : hhpod : densité de population

 nempd : densité d'emploi

ANNEXE 5

PARTS MODALES DES LIGNES DE ROCADES EXISTANTES (sources : enquêtes clientèles RATP et SNCF)

1) Pour le Tramway Saint-Denis – Bobigny (T1) :

Date de mise en service : décembre 1992

Caractéristiques :

- 9 km de lignes , 21 stations, capacité de 2300 par heure par sens
- Fréquence : 4 mn à la pointe, 7 mn en heure creuse
- Vitesse commerciale : 16,5 km/h
- Correspondances avec les lignes 5,7,13, lignes de banlieue nord SNCF et une trentaine de lignes de bus.

Trafic : 63000 voyageurs quotidiens en 1996 (12000 de plus que les prévisions hautes), 74 000 aujourd'hui. Le flux d'échanges maximum a lieu chaque jour à la station Pablo-Picasso. Flux maximum enregistré en 1996 : 1900 à l'heure.

Distance moyenne parcourue : 2,5 km

Structure des déplacements :

- Cabotage (trajet simple, sans correspondance) : 36%
- Avec une seule correspondance (bus, métro, train) : 52%
- Avec au moins deux correspondances (une avant et une après) : 12%

Motifs :

- Domicile<->Travail : 38%
- Domicile<->Ecole : 20%
- Non obligés : 42%

Modes de rabattement :

- MAP : 59%
- Bus : 20%
- Autres (métro,train,VP) : 21%

Report modal :

Plus de 30% des utilisateurs actuels du tramway ont changé de comportement :

- 6% utilisaient auparavant leur VP ou un autre mode individuel
- 7% faisaient leur parcours à pied
- 13,5% n'effectuaient pas auparavant le déplacement considéré (donc augmentation de mobilité)
- 9% ont changé d'itinéraire
- sur les 42% de trajets non obligés, la part du bus était de 7.5%
- 12% déclarent que sans le tram, ils auraient renoncé à ce déplacement et 6% qu'ils l'auraient fait en voiture.

Finalement, le trafic TC sur la liaison Saint-Denis - Bobigny a connu une augmentation de 160% passant de 24000 voyageurs avant la mise en service du tramway T1 à 63000 après.

Remarques :

- L'information du tps d'attente a fidélisé la clientèle.
- L'ouverture du commercial régional Avenir à Drancy en 1995 a engendré une augmentation de 600 à 3000 voyageurs par jour à la station Avenir
- Il existe un projet de passer à 17km/h par ondes vertes.

2) Pour le tramway Val de Seine (TVS)

Date de mise en service : 1^{er} juillet 97

Caractéristiques :

- Fréquence : 4 mn en heure de pointe, 22 mn entre les terminus
- Capacité : 13 rames en service, 2600 places par heure par sens
- Vitesse moyenne : 32 km/h

Trafic : en octobre 1997 : 22000 voyageurs par jour . En mars 98, il atteignait 35000 et début 2000, 38000 voyageurs.

Maximum enregistré en pointe : 2200

La ligne n'est pas saturée, malgré l'attraction du pôle de la Défense. Cela s'explique en particulier par l'étalement de la pointe.

Motifs :

Les enquêtes clientèle montrent un taux important de déplacements de type achats vers la Défense

Report modal : sur les 35000 voyageurs/jour en 1998,

- 70,3% des voyageurs actuels qui n'ont pas connu de changement de situation de domicile, de lieu de travail ou d'école effectuaient déjà le déplacement avant la mise en service du TVS. Ils se répartissent comme suit :
 - 43% l'effectuait en train (RER ou SNCF). En particulier, le TVS a pris des clients de la ligne SNCF St Lazarre-Versailles.
 - 29% en bus
 - 14% en autre TC (métro, car APTR ou employeur)
 - 12% en VP
 - 2% à pied
- 12% n'auraient pas effectué le déplacement en l'absence du TVS (donc augmentation de mobilité)

- 16,2% ont connu entre la mise en service du TVS et le jour de l'enquête des changements de domicile, de lieu de travail ou d'école, donc pouvant modifier leur itinéraire.
- 1,5% ne peuvent être affectés.

Remarques :

- Le trafic devrait augmenter à l'avenir en raison des projets immobiliers (surtout à Issy) et de l'aménagement de l'Île Séguin.
- Il y a un projet de réorganisation des lignes de bus aux abords du TVS.

3) La liaison Saint-Quentin – La Défense (SQDF)

Date de mise en service : mai 1995

Caractéristiques :

- Semi-direct entre la Verrière et la Défense
- Fréquence et temps : 15 mn en heure de pointe, 35 mn en temps de parcours
- Vitesse moyenne : 60 km/h

Trafic : 38800 voyageurs par jour. Charge maximum enregistrée sur 2 heures de pointe du matin dans le sens La Verrière- La Défense : 5900

Report modal à la journée:

- 75% des voyageurs actuels qui n'ont pas connu de changement de situation de domicile, de lieu de travail ou d'école effectuaient déjà le déplacement avant la mise en service de la liaison SQDF. Ils se répartissent comme suit :
 - 8% prenaient avant leur voiture
 - 67% prenaient un mode TC
- 25% ne faisaient pas le même déplacement avant la mise en service de la nouvelle ligne. Parmi eux, 66% travaillaient ailleurs, 22% habitaient ailleurs.

Avant l'ouverture de la ligne, le temps moyen de trajet était de 73,4 mn . Après, il est de 68,8 mn (66,8 mn pour les personnes effectuant le même déplacement qu'auparavant ; 74,9 mn pour les nouveaux clients de la ligne).

Bibliographie

- [1] : "Discrete choice analysis : theory and application to travel demand" – M. Ben-Akiva and S. Lerman . MIT press. 1985.
- [2] : "Multinomial probit, the theory and its application to demand forecasting" – Carlos Daganzo. 1979.
- [3] : "The dogit model" – M. Gaudry and M. Dagenais. Dans "Transportation Research Board". 13B, 105-111. 1979.
- [4] : "Activity based disaggregate travel demand model system with daily activity schedules" – Moshe E. Ben-Akiva and John L. Bowman. Dans "Transportation Research Board". 1995.
- [5] : « Sur le choix du mode de transport par les voyageurs en milieu urbain » : Fabien Leurents - Rapport INRETS (1998)
- [6] : « La modélisation multimodale des déplacements régionaux des personnes » : MVA Consultants (1998)
- [7] : « Confort – régularité – information » : étude réalisée par un groupe de travail mené par l'INRETS , et commandée par l'instance d'évaluation de la politique des transports en Ile-de-France (1997).
- [8] : « Les modèles de choix modal en Ile-de-France » : rapport SNCF (1997).
- [9] : "Le modèle de trafic de l'IAURIF" - D.Nguyen-Luong (février 1997)
- [10] "SAS/STAT user's guide. Volume 2" - pages 1071 à 1126.
- [11] "Facteurs d'évolution du trafic des transports en commun" - Rapport RATP . G.Thibal, J.Rousseau, C.Saut (1997)