

Rapport d'études

Modèles de trafic routier

**Influence des composantes du niveau de service (temps,
péage, confort, ...) sur le choix d'itinéraire**



Ministère
de l'écologie,
du Développement
durable
et de l'énergie

Page laissée blanche intentionnellement

Sommaire

Introduction.....	4
Chapitre I : Définition du problème.....	5
1.1 - D'Ariane à TransCAD.....	6
1.2 - Valeurs du temps : définitions, références et méthodes.....	11
1.3 - Objectifs de l'étude.....	14
1.4 - Principes de calibrage des modèles.....	15
Chapitre 2 : Données utilisées.....	17
2.1 - Base nationale (données nationales du SETRA – 2001).....	18
2.2 - Base « CETE NC » (2 coupures sélectionnées en Haute Normandie).....	21
2.3 - Enquête de préférences déclarées sur les traversées vosgiennes.....	23
Chapitre 3 : Analyse des résultats des modèles déterministes.....	28
3.1 - Application des valeurs de l'instruction.....	29
3.2 - Malus vs valeurs du temps.....	29
3.3 - Choix des malus et formulation de la fonction de coût.....	31
3.4 - Influence de la distance et des motifs sur la valeur du temps.....	33
3.5 - Calibrage des valeur du temps.....	36
3.6 - Influence des vitesses retenues.....	38
3.7 - Vérification des valeurs de malus initialement retenues.....	40
3.8 - De l'utilisation de valeurs de référence sur un cas particulier.....	41
Chapitre 4 : Analyse des résultats des modèles stochastiques.....	42
Chapitre 5 : Synthèse des résultats et recommandations.....	47
5.1 - Déterministe vs Stochastique.....	48
5.2 - Synthèse sur les malus et les valeurs du temps.....	49
5.3 - Éléments de recommandations pour l'utilisation de modèles de choix d'itinéraire et l'introduction de valeurs du temps.....	50
5.4 - Limites de l'exercice et suites à donner.....	51
Bibliographie.....	53

Introduction

Deux modèles d'affectation distincts sont principalement utilisés pour les études de trafic et d'évaluation socio-économique des projet d'infrastructures routières interurbaines : la loi dite d'Abraham implantée dans le logiciel Ariane utilisé jusqu'en 2007 par le Séttra, et la loi déterministe prix/temps utilisée par exemple dans les modules SETRA de TransCAD. Dans les deux cas, le choix d'itinéraire s'effectue en fonction d'une analyse fondée sur le temps, le coût du parcours (péage, carburant, utilisation du véhicule, et malus d'inconfort), et la valeur du temps, définie ici comme le consentement à payer des usagers pour gagner du temps.

Il existe principalement trois degrés de liberté pour ajuster ces modèles, c'est à dire pour reproduire les choix d'itinéraires réellement observés :

- la description des niveaux de services
- le choix des malus
- le choix de la distribution des valeurs du temps

Le premier point, qui renvoie à la connaissance des conditions de circulation (habituellement modélisé par l'utilisation de courbes débits/vitesses), ne sera pas abordé ici. Les deux autres points constituent le cœur de cette étude : en effet, même si le cadre conceptuel de ces deux modèles d'affectation est bien connu (de nombreuses expériences de l'utilisation de ces modèles ont été réalisées, notamment dans le cadre du Schéma National des Infrastructures de Transport), les modélisateurs manquent toujours de références pour calibrer leurs modèles de choix d'itinéraires en choisissant des valeurs du temps et des malus adaptés. Par ailleurs, ces deux méthodes (Abraham, et prix/temps) ou logiciels (Ariane et TransCAD) ont certes déjà fait l'objet de multiples comparaisons d'un point de vue théorique et pratique, mais aucune évaluation de la capacité de ces modèles à reproduire des choix d'itinéraire effectivement observés n'a encore été réalisée de manière systématique sur une base de donnée très diversifiée.

L'enjeu de cette étude est donc de discuter de l'application de ces deux modèles de choix d'itinéraires, et de donner un cadre de travail aux chargés d'études en proposant des recommandations générales pour l'utilisation de modèles de choix d'itinéraires et l'introduction de malus et de valeurs du temps (et en particulier lors du paramétrage des modules SETRA de TransCAD).

L'étude s'articule donc autour de deux éléments partiellement complémentaires :

- d'une part les méthodes de reconstitution de choix d'itinéraire pour l'évaluation de projets routiers
- d'autre part les notions de valeur du temps (dans une conception recentrée dans le domaine du transport sur l'effet d'un gain de temps sur les choix des usagers au vu d'un éventuel surcoût résultant) et de malus.

Les éléments présentés ici constituent une exploration des comportements constatés. Ce type d'analyse ne saurait avoir une vocation de normalisation générale en l'absence d'autres travaux contradictoires. A notre connaissance, peu d'études ont été publiées en France sur le sujet depuis les analyses sur le cas du tunnel du Prado-Carénage en 1996.

Chapitre I : Définition du problème

1.1 - D'Ariane à TransCAD

Ariane est l'outil « historique » qui a été utilisé jusqu'en 2007 par le ministère pour la réalisation d'études de trafic et d'évaluation socio-économique en milieu interurbain ou périurbain. L'affectation des trafics y était effectuée à l'aide d'un modèle logit logarithmique appelé « loi d'Abraham » [1].

La loi d'Abraham est un modèle probabiliste : l'utilisateur opère son choix d'itinéraire selon l'utilité de chaque itinéraire, celle-ci étant définie par des critères connus, dit « déterministes » (temps, coût du péage, carburant, inconfort...), mais aussi par d'autres critères inconnus du modélisateur. Afin de raisonner de manière probabiliste, on définit cette partie inconnue par une variable aléatoire que l'on somme à la partie déterministe, et on suppose que l'utilisateur cherche à maximiser son utilité parmi tous les itinéraires possibles.

La formulation mathématique de la répartition entre les itinéraires dépend de l'hypothèse émise sur la distribution de la variable aléatoire. Dans le cas de la loi d'Abraham, elle peut s'écrire sous une forme très simple :

$$CG_1^{10} \times F_1 = CG_2^{10} \times F_2 = \dots = CG_n^{10} \times F_n$$

où :

- CG_i est le coût généralisé de l'itinéraire i
- F_i le flux empruntant l'itinéraire i
- $1, 2, \dots, n$ l'ensemble des itinéraires empruntés par le trafic de l'O/D considérée

Pour deux itinéraires, la probabilité de choisir l'itinéraire 1 en fonction du rapport CG_2/CG_1 , se traduit ainsi de la manière suivante :

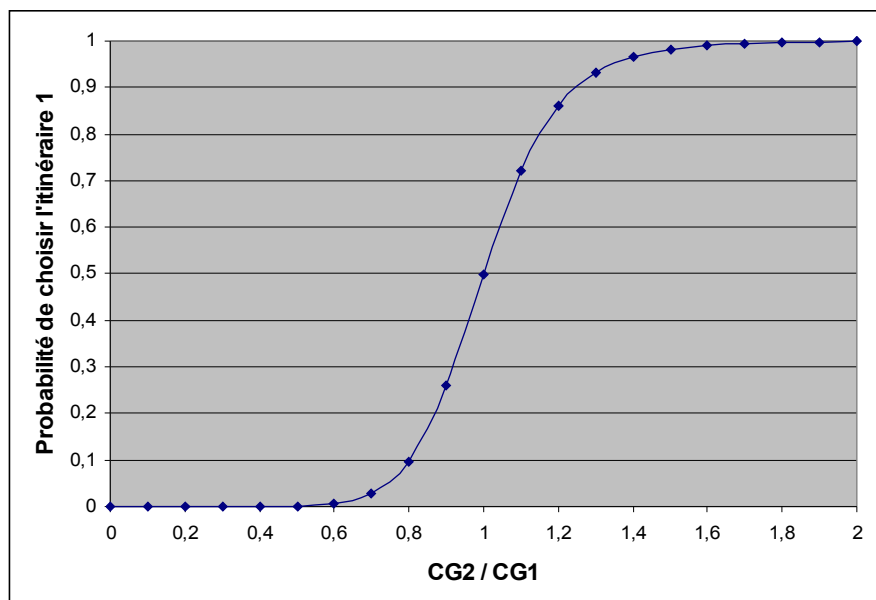


Illustration 1 : Probabilité de choix d'un itinéraire en fonction du rapport des coûts généralisés, pour la loi d'Abraham

Dans Ariane, le paramétrage de la valeur du temps servant à calculer ce coût généralisé était impossible, la valeur du temps VL étant automatiquement fixée à 14€₂₀₀₀/h.

En 2003, le SETRA (Service d'Étude sur les Transports, les Routes et leurs Aménagements) lança un appel d'offre pour le développement d'un nouvel outil afin de pallier aux difficultés d'Ariane (incompatibilité avec les systèmes d'information géographique, sélection des itinéraires à la main) et pour se mettre en conformité avec l'instruction cadre d'évaluation [2]. Un groupement Setec / Caliper a été retenu, et des modules spécifiques (appelés modules SETRA) intégrés au logiciel TransCAD ont ainsi été développés et livrés en 2006. La nouvelle loi d'affectation utilisée dans les modules SETRA de TransCAD est le prix-temps (algorithme spécifié par F. Leurent [3]).

Dans la règle prix-temps, le coût généralisé d'un itinéraire est composé de deux termes distincts :

- le temps (T)
- le prix, qui englobe les péages, l'inconfort, les dépenses de carburant, d'usure et d'entretien du véhicule (P)

On suppose que chaque automobiliste a sa propre valeur du temps, et qu'il raisonne de façon déterministe, c'est à dire qu'il choisit l'itinéraire qui lui permet de minimiser son coût généralisé.

Ainsi, la diversité des choix d'itinéraire dépend de la diversité des valeurs du temps au sein de la population, comme le montre l'illustration suivante :

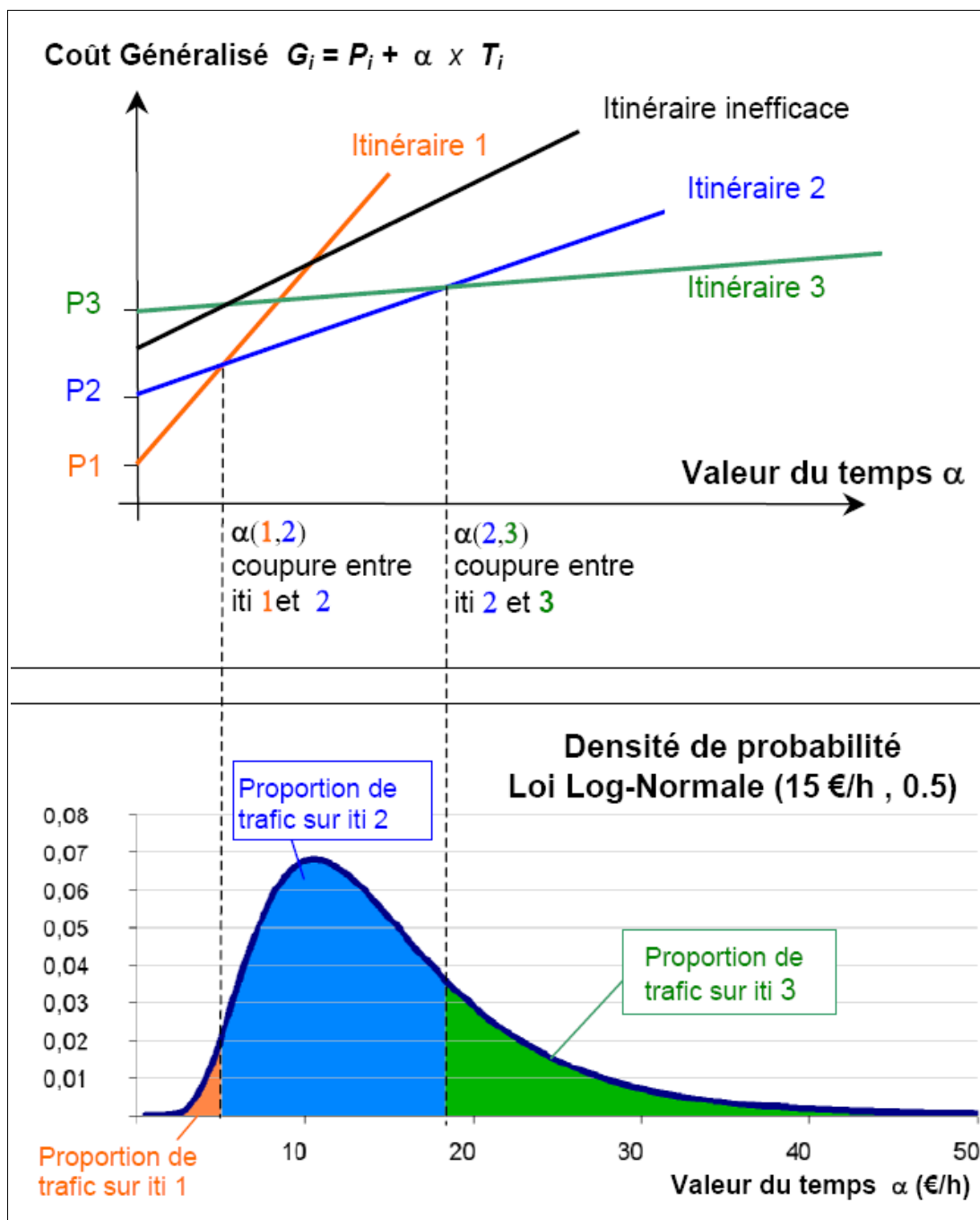


Illustration 2 : Principe de fonctionnement de la loi d'affectation prix/temps

Dans les modules SETRA de TransCAD, on suppose que la valeur du temps est distribuée selon une loi log-normale (même forme que la distribution des revenus au sein de la population française). La forme de cette loi de distribution est donc fixe, et sa densité de probabilité est donnée par la formule suivante :

$$f(vdt; \mu, \sigma) = \frac{e^{-0,5((\ln(vdt) - \mu)/\sigma)^2}}{vdt \times \sigma \sqrt{2\pi}}$$

Les paramètres (mu et sigma) sont à définir par l'utilisateur.

Il est important de préciser que ce paramètre μ n'est pas égal à la moyenne (l'espérance) de cette loi log-normale. En effet, la moyenne est donnée par la relation :

$$moyenne = e^{\mu + \frac{\sigma^2}{2}}$$

La médiane, quant à elle, est donnée par la relation :

$$médiane = e^{\mu}$$

En pratique, il est plus aisé de définir cette loi log-normale à partir de sa médiane. En effet, une relation bijective directe existant entre μ et la médiane, il est tout aussi pertinent et plus facile pour l'analyse de raisonner directement à partir de cette dernière. Par ailleurs, la distribution log-normale est étalée vers des valeurs très importantes de valeur du temps, ce qui peut conduire à des moyennes très élevées, qui rendent moins bien compte de la distribution réelle que la médiane. A l'aide de la médiane, la densité de probabilité s'exprime sous la forme :

$$f(vdt ; médiane, \sigma) = \frac{e^{-0,5((\ln(vdt) - \ln(médiane))/\sigma)^2}}{vdt \times \sigma \sqrt{2\pi}}$$

Les deux illustrations suivantes présentent la forme de cette loi distribution suivant plusieurs valeurs de la médiane et de sigma :

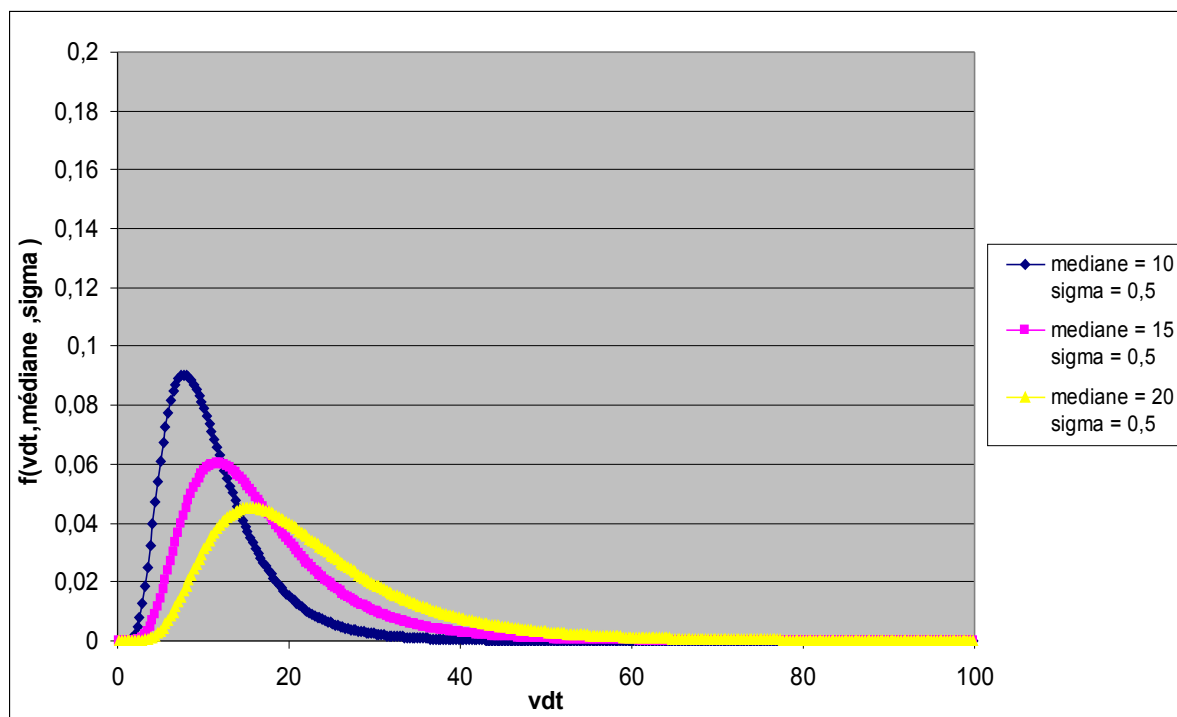


Illustration 3 : Forme de la loi log-normale pour plusieurs valeurs de médiane

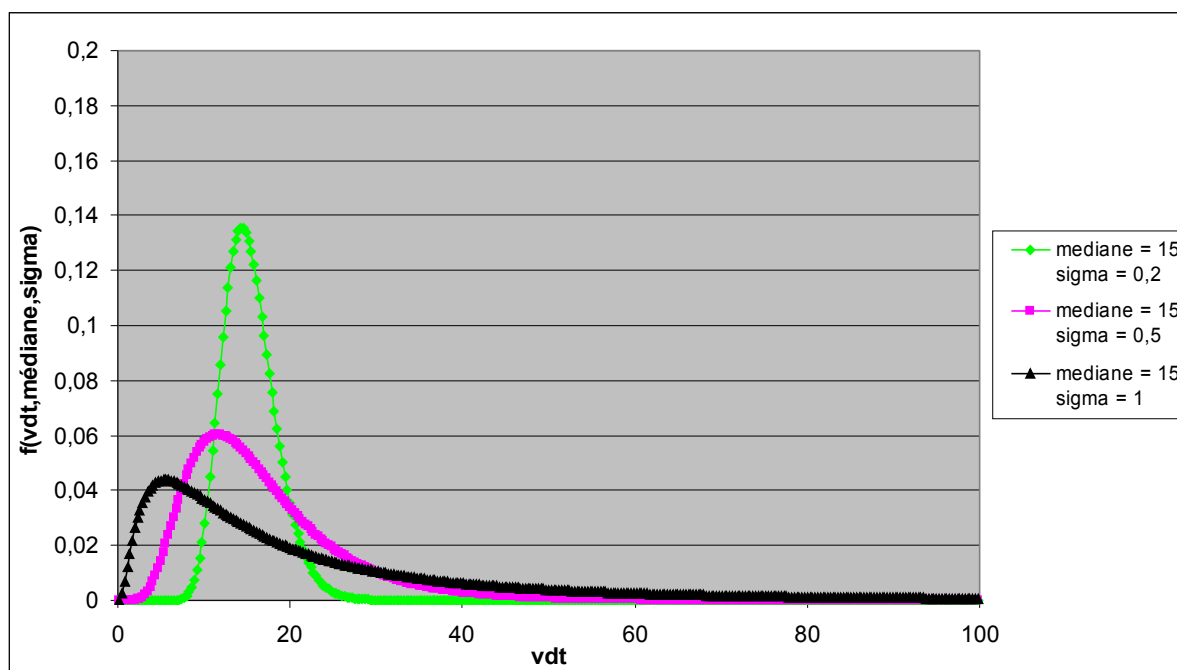


Illustration 4 : Forme de la loi log-normale pour plusieurs valeurs de sigma

Aujourd'hui, le cadre conceptuel de l'algorithme prix-temps est bien connu, et de nombreuses expériences de l'utilisation des modules SETRA ont été réalisées (notamment dans le cadre de l'évaluation des projets du Schéma National des Infrastructures de Transport). Cependant, les modélisateurs manquent toujours de références pour choisir les valeurs du temps initiales en entrée de leur modèle d'affectation – cf. paragraphe 1.2. L'objectif de cette étude est donc de donner un cadre de travail aux chargés d'études utilisant les modules

Par ailleurs, on abordera également ici la question du calibrage des malus d'itinéraire, qui sont directement liés aux valeurs du temps. En effet, les malus d'itinéraires sont habituellement utilisés pour exprimer l'inconfort ressenti par l'utilisateur sur certains itinéraires : les voies autoroutières sont supposées avoir un malus nul, et les bidirectionnelles ont un malus positif de telle sorte que leur coût généralisé est pénalisé. On comprend donc facilement que les valeurs des malus sont intimement liées à celles des valeurs du temps, c'est la raison pour laquelle une tentative de calibrage des malus sera également menée dans cette approche.

Enfin, ces deux méthodes (Abraham, et prix/temps) ou logiciels (Ariane et TransCAD) ont certes déjà fait l'objet de multiples comparaisons d'un point de vue théorique [4] et pratique [5], mais aucune évaluation de la capacité de ces modèles à reproduire des choix d'itinéraire effectivement observés n'a encore été réalisée de manière systématique sur une base de données très diversifiée. Il s'agira donc ici de mettre en évidence les avantages et inconvénients de chaque méthode et de pointer les limites d'application de chacun de ces outils.

1.2 - Valeurs du temps : définitions, références et méthodes

Le concept de valeur du temps est utilisé très largement, entre autres du fait de l'essor des modèles de choix discret depuis les années 1980, et couvre un large spectre de significations. Comme il sera l'élément central de ce travail, il s'agit donc de définir précisément ce qu'on entend par valeur du temps dans la suite du document. Il est par ailleurs exclu de procéder ici à une comparaison exhaustive de résultats obtenus de calibrage des valeurs du temps, mais la présentation de quelques résultats pourra être utile pour appuyer notre propos. Nous avons utilisé pour ce faire un document relativement ancien, le mémoire de thèse de MSc de M Pfeffer datant de Juillet 2001 [6], mais couvrant exactement le champs du présent exposé, et effectuant une revue de littérature suffisamment vaste pour nous fournir un bon aperçu de la question.

Parmi les résultats sur la valeur du temps antérieurs à 2001 qui sont listés par Pfeffer sur différents pays européens, le constat d'une extrême hétérogénéité ressort (les moyennes valant entre 0,75 €/h et 28 €/h pour les voyageurs selon les pays et les motifs). Les conclusions qui ont pu en être tirées sont les suivantes :

- la valeur du temps suit une distribution log-normale ;
- la médiane des valeurs du temps est régionale et est corrélée avec la médiane des revenus ;
- a contrario, le paramètre sigma de la distribution log-normale obtenu par différentes études pour les déplacements de voyageurs est compris entre 0,6 et 1,1 : il est donc plus élevé que le paramètre sigma de la distribution des revenus (qui était compris en 2000 pour les pays concernés entre 0,4 et 0,6). Pfeffer recommande néanmoins d'utiliser cet indicateur pour transférer les valeurs du temps empiriques obtenues dans un contexte particulier.

Par ailleurs, d'après Pfeffer, trois méthodes sont proposées pour déterminer la valeur du temps :

- la distribution constatée des salaires ;
- les enquêtes de préférences révélées ;
- les enquêtes de préférences déclarées.

La première méthode est basée sur une hypothèse similaire à celle de Boiteux II (voir plus loin) quant aux choix des individus en tant qu'acteurs rationnels. Elle nécessite par ailleurs d'émettre des hypothèses supplémentaires en fonction des différents motifs sur la part du salaire horaire que représente la valeur du temps. Enfin, elle ne se base pas sur des éléments spécifiques de choix d'itinéraire.

Les préférences déclarées sont selon Pfeffer l'instrumentation la plus courante pour estimer les valeurs du temps. Elles présentent l'avantage de permettre de contrôler la plage de valeurs du temps de coupure pour lesquelles les choix des usagers peuvent être examinés. A contrario, les choix effectués par les usagers dans ces enquêtes sont

toujours sujet à nuances multiples (caractère fictif du choix, mauvaise interprétation possible des questions,...). En outre, il est très difficile d'intégrer des attributs de confort dans de tels questionnaires : ceux-ci seront seulement mesurés indirectement par les choix effectués et donc englobés à l'intérieur des attributs de préférence d'alternative.

Enfin, les enquêtes de préférences révélées représentent la méthode se rapprochant le plus d'une mesure de la valeur du temps. Cependant, elles se situent dans un cadre précis (plage de valeurs du temps de coupure restreinte) et leurs résultats ne sont d'aucune utilité pour un utilisateur amené à étudier des situations très différentes de ce qui est actuellement observé (par exemple un péage kilométrique très élevé).

En France, les documents de référence actuels sur la valeur du temps sont d'une part le rapport Boiteux II [7] et d'autre part le projet d'instruction sur l'évaluation économique des projets d'investissements routiers interurbains de mai 2007 :

- dans le rapport Boiteux II, le concept de valeur du temps fait surtout référence à l'évaluation économique, qui est fondée sur le principe selon lequel les gains de temps vont pouvoir être convertis en avantages monétarisables (par exemple, un gain de une heure de trajet pour un poids lourd va diminuer le coût pour le transporteur, ce qui lui permet d'augmenter ses profits, sa productivité, ou encore d'être plus compétitif). Cette notion ne reflète cependant pas nécessairement tous les comportements de long terme [8], où les réactions des acteurs peuvent être multiples. Ainsi, de nombreuses études théoriques ou empiriques (Duprez [9], Offner[10] ou Bazin[11] montrent qu'il serait illusoire de rechercher des effets structurants systématiques d'une infrastructure de transport, effets qui iraient au-delà de ce qui est déjà pris en compte dans la valeur des gains de temps.

Nonobstant ces limites de l'évaluation économique, le rapport Boiteux II postule que « l'évaluation des avantages pour l'usager par des procédures cohérentes avec le modèle de trafic utilisé est [...] satisfaisante sur le plan de la logique économique », ce qui justifierait l'utilisation d'une valeur du temps identique pour l'évaluation et la modélisation des choix d'itinéraires.

Or utiliser les valeurs du temps de l'évaluation pour modéliser le choix d'itinéraire signifierait que les usagers sont des acteurs rationnels, cherchant à effectuer des arbitrages « entre leur consommation et leur activité en fonction de leur temps disponible ». Cependant, cette approche très restrictive ne semble pas correspondre à l'intuition dans différentes situations de choix d'itinéraire :

- les acteurs à très faibles revenus effectuent un arbitrage uniquement en fonction du coût, le temps de parcours n'intervient que dans le choix ou non d'effectuer le déplacement ;
- les acteurs peuvent également avoir une aversion à s'acquitter d'un péage ;
- il faut également tenir compte d'une certaine méconnaissance des conditions de circulation par les usagers. Par exemple, les vacanciers descendant vers la Méditerranée, connaissant mal le réseau routier, empruntent massivement tous les étés les autoroutes à péage, acceptant de perdre des heures dans les bouchons ;
- pour les déplacements professionnels remboursés par l'employeur, sur des sections de coût du péage « moyen » pratiqué actuellement, le surcoût est faible dans l'activité économique et joue peu ou pas de rôle dans leur choix d'itinéraire. Ces usagers ne s'orientent que sur le temps de parcours : il en résulte notamment qu'ils représentent de 40 à 50 % de l'usage de certaines autoroutes concédées (A4, A13 par exemple) un jour normal de semaine.

Ces situations peuvent conduire à des valeurs du temps très faibles ou très élevées, qui n'auraient pas de sens pour une évaluation économique. Cependant, ces limites du concept de la valeur du temps ne remettent pas en cause le fait qu'il faille pondérer le coût, le temps et d'autres attributs du confort des itinéraires pour modéliser les choix. La valeur du temps de la modélisation est donc par définition la pondération permettant de mettre en relation le coût et le temps de déplacement sur le choix d'itinéraire.

- dans le projet d'instruction de mai 2007 [2], les valeurs de proposées pour l'évaluation socio-économiques (qui sont proches des valeurs proposées dans le rapport Boiteux II [7]), sont les suivantes (valeurs exprimées en euros 2000) :
 - $d < 20\text{km}$: 9,88 €/h
 - $20\text{km} < d < 50\text{km}$: 13,41 €/h
 - $50\text{km} < d < 400\text{km}$: $0,0304 \times d + 15,39$
 - $d > 400$: 34,36

Il est toutefois précisé (annexe 5, page 6) : « Les valeurs unitaires à prendre en compte pour les calculs des coûts de circulation sont définies en annexe 7, à l'exception des valeurs du temps qui, dans les modèles d'affectation, sont celles qui représentent le mieux le comportement des usagers en situation de choix. Elles sont généralement différentes des valeurs normalisées pour le calcul économique ». En l'absence de calibrage spécifique, la valeur du temps recommandée pour le modèle d'affectation prix-temps est considérée suivre une distribution log normale au sein de la population, de paramètre sigma variant entre 0,5 et 0,7 et de moyenne variant entre 12 et 15 €₂₀₀₀/h.

En définitive, même si le fait d'utiliser les valeurs du temps de l'évaluation pour modéliser le choix d'itinéraire a été remis en cause dans l'instruction, nous chercherons tout de même à observer la pertinence de l'utilisation des valeurs du temps de l'évaluation pour modéliser le choix d'itinéraire. Les valeurs de l'évaluation étant de simples moyennes, et la démarche proposée ici faisant appel à des distributions, la valeur du paramètre sigma (0,5) proposée dans l'instruction pour l'affectation sera également utilisée. Ce qui sera appelé « valeurs du temps de l'instruction » dans toute la suite fera ainsi référence à la distribution des valeurs du temps construite à partir d'une loi log-normale de moyenne égale aux valeurs proposées pour l'évaluation (données ci-dessus) et de paramètre sigma égal à la valeur proposée pour l'affectation (0,5). Les valeurs du temps qui résulteront du calibrage sont à interpréter avec précaution : elles ne sont utilisables et transférables que dans des contextes de concurrences entre itinéraires similaires à celles présentes dans la base de données servant au calibrage (voir partie 2 pour la description de ces itinéraires). En particulier, les différences de coût de déplacement disponibles dans cette base de donnée sont inférieures à 45 € et 14c€₂₀₀₀/km.

Ceci signifie que le modèle obtenu n'est pas utilisable pour analyser la concurrence entre deux itinéraires dont la différence de coût dépasse fortement ces bornes supérieures.

La limite de la présente étude peut également être envisagée sous l'angle de la valeur du temps de coupure. Il s'agit de la valeur du temps pour laquelle les coûts généralisés de deux itinéraires comparés sont égaux (voir schéma sur l'algorithme prix-temps au 1.1). Les situations de concurrence examinées dans cette étude concernent, pour 95 % des trafics représentés, des valeurs du temps de coupure de moins de 20 €/h. Obtenir une médiane des valeurs du temps de 319,2 €/h et un paramètre sigma de 2,66 (comme c'est le cas des résultats obtenus pour les déplacements de plus de 400 km, cf conclusions de l'étude partie 5.2) n'a pas dans ce contexte une signification absolue : il convient plutôt plutôt d'interpréter qu'environ 15 % des usagers n'acceptent pas de payer 20 € pour gagner une heure de trajet – voir illustration 5.

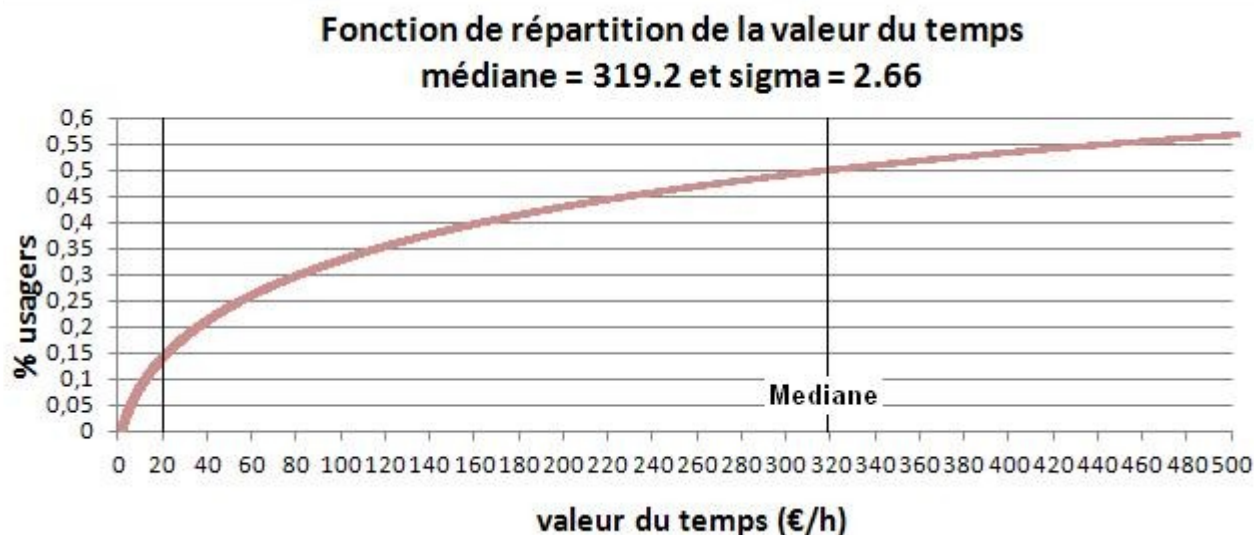


Illustration 5 : Fonction de répartition de la valeur du temps pour une médiane de 319,2 €/h et $\sigma=2,66$

1.3 - Objectifs de l'étude

L'étude est centrée sur le calibrage des méthodes de reproduction des choix d'itinéraires interurbains pour les déplacements de voyageurs, à partir d'enquêtes O/D constituant des préférences révélées et d'une enquête de préférences déclarées. L'objet principal de ces travaux est de déterminer l'influence du niveau de service d'une route sur les décisions des usagers.

L'étude s'articule autour de deux éléments partiellement complémentaires :

1. d'une part les méthodes de reconstitution de choix d'itinéraire pour l'évaluation de projets routiers ;
2. d'autre part la notion de valeur du temps, dans une conception recentrée dans le domaine du transport sur l'effet d'un gain de temps sur les choix des usagers au vu d'un éventuel surcoût résultant.

Au vu de l'ampleur du travail pour cerner précisément ces deux sujets, de la connaissance actuelle des niveaux de service sur routes et autoroutes et de la disponibilité des bases de données, l'étude servira principalement à donner de premières orientations et à tester puis valider des méthodes.

Les causes de la variabilité des comportements ne sont examinées que pour obtenir une reproduction des données suffisamment satisfaisante. Dans ce cadre, les facteurs examinés ont été la distance et le motif du déplacement, tandis que le taux d'occupation des véhicules n'a pas été retenu. Il résulte de cette démarche que l'unité d'analyse est le véhicule léger (VL), et donc en particulier que les valeurs du temps seront exprimées en euros/VL et non en euros/voyageur.

Par ailleurs, les points suivants ne seront pas abordés :

- les choix d'itinéraires pour le transport de marchandises ;
- le fonctionnement des déplacements de voyageurs en milieu urbain ;
- le lien entre la distribution des valeurs du temps et du salaire.

1.4 - Principes de calibrage des modèles

Pour cette étude, on dispose de données d'enquêtes O/D routières [12]. Ces enquêtes peuvent être considérées comme des enquêtes de « préférences révélées », car elles donnent une information sur le choix effectivement réalisé par l'utilisateur. On travaille ici à partir de données désagrégées uniquement, et c'est dans ce cadre bien précis que l'étude sera limitée. Les deux modèles, déterministe et stochastique, permettent pour chaque individu de calculer la probabilité de choix d'itinéraire qui a effectivement été choisi par l'utilisateur : le produit de ces probabilités sur l'ensemble des individus enquêtés permet de définir une fonction appelée fonction de vraisemblance. C'est cette fonction de vraisemblance que l'on cherche à maximiser lors du calibrage du modèle.

Fonction de vraisemblance :

$$L = \prod_n P(n_i)$$

Où n_i caractérise le choix constaté pour l'individu n de l'itinéraire i , et $P(n_i)$ la probabilité (calculée par le modèle) que l'individu choisisse cet itinéraire.

En pratique, on ne cherchera pas à maximiser la vraisemblance mais la log-vraisemblance :

$$\ln(L) = \sum_n \ln(P(n_i))$$

Par ailleurs, puisque l'on dispose de données d'enquêtes O/D routière, les individus n'ont pas tous le même poids (coefficients de redressement). Ainsi, On pondérera chaque individu par son poids respectif, ce qui implique qu'en définitive on cherche à maximiser la fonction suivante :

$$\ln(L) = \sum_n Poids_n \times \ln(P(n_i))$$

En pratique, pour le calibrage des **modèles stochastiques**, on utilisera le logiciel Biogème, qui fournit tous les indicateurs statistiques pour s'assurer de la fiabilité des paramètres estimés.

Pour évaluer la qualité du modèle, on observera d'abord la valeur de la log-vraisemblance ($\log V$) obtenue à l'issue de processus de calibrage. Cette log-vraisemblance pourra être comparée avec la log-vraisemblance initiale ($\log 0$), qui est définie dans cette étude comme la valeur de la log-vraisemblance obtenue pour un modèle où toutes les alternatives sont équiprobables : l'amélioration entre le modèle évalué et ce modèle initial doit être significative, ce que l'on considère comme réalisé lorsque la valeur de l'expression $2 \times (\log V - \log 0)$ est supérieure à la valeur prise par la loi du χ^2 à k degrés de liberté (k étant le nombre de variables du modèle à évaluer).

Par ailleurs, la qualité du modèle sera également évaluée en comparant à un niveau agrégé les prévisions du modèle aux données observées. Pour ce faire, on fournira le coefficient de régression et le coefficient de détermination obtenu par régression linéaire entre les choix d'itinéraires agrégés observés et les résultats du modèle, comme le montre l'exemple suivant :

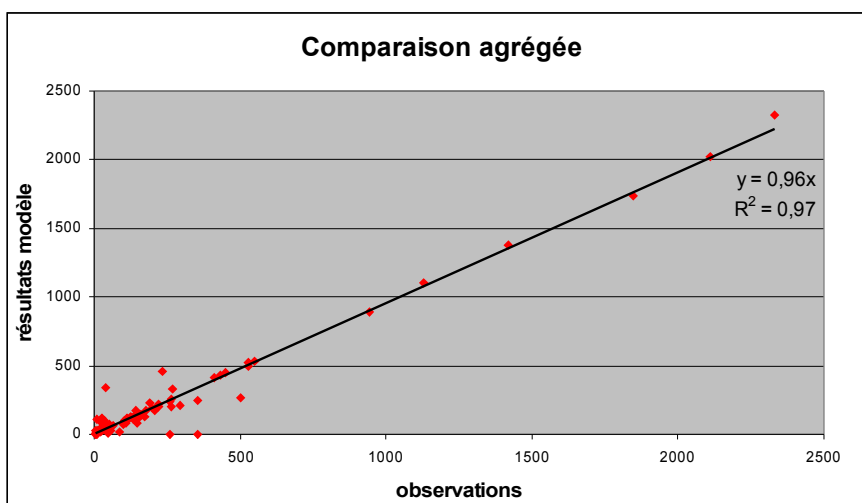


Illustration 6 : Exemple de résultat d'une comparaison sur les choix d'itinéraires agrégés

Pour le **modèle déterministe**, la répartition du trafic sur chaque itinéraire suppose de faire appel à la fonction de distribution de la valeur du temps, ce qui ne peut se faire avec le logiciel Biogème (celui-ci est adapté pour les modèles d'utilité aléatoire uniquement). Ainsi, la résolution du problème s'est effectuée sur excel à l'aide du solveur.

Les garanties en terme de fiabilité des paramètres sont toutefois moins bonnes qu'avec Biogème (pas de test de Student par exemple), mais il est possible, de la même manière qu'avec les modèles stochastiques, d'évaluer la qualité des différents modèles à l'aide de la valeur de la log-vraisemblance obtenue, et de chercher à les valider à l'aide de comparaisons effectuées sur des choix d'itinéraires agrégés.

Une différence importante est néanmoins présente entre les modèles déterministes et les modèles stochastiques. Si certains usagers choisissent des itinéraires non efficaces, la log-vraisemblance ne peut plus être calculée (elle est infinie). Ceci impose donc de travailler uniquement sur la sélection des O/D pour lesquelles tous les itinéraires choisis sont efficaces. La comparaison des log vraisemblances entre modèles stochastiques et déterministes est donc hasardeuse.

Les calculs sont tous effectués en euros 2000. Par contre, les valeurs du paramètre sont à indiquer à une date précise, dépendant de la date de réalisation de l'enquête. Ceci est lié au fait que les choix des usagers traduisent des arbitrages économiques et des perceptions temporaires, qui peuvent évoluer dans le temps. Nous utiliserons trois bases de données différentes (voir partie 2). Par convention, les résultats de la base nationale, s'appuyant sur des données d'enquête de 1995 à 2000, sont supposées être des valeurs 2000. Les deux autres bases (base « CETE NC » et base de l'enquête de préférences déclarées dans les Vosges) sont exprimées en valeurs 2010. Étant donné que les paramètres de choix d'itinéraire sont classiquement surtout reliés aux arbitrages des ménages, l'indice de consommation finale des ménages par tête est généralement retenu pour les faire évoluer. La comparaison des valeurs 2000 et 2010 devra donc tenir compte d'une augmentation de 7,5 % depuis 2000 (calcul basé sur une augmentation de la consommation finale des ménages de 40 %, de la population française de 7 % et de l'indice des prix à la consommation de 20%, source : www.insee.fr).

Chapitre 2 : Données utilisées

2.1 - Base nationale (données nationales du SETRA – 2001)

Cette base de type « préférences révélées » a été construite par l'ensemble des CETE (Centres d'Études Techniques de l'Équipement) à partir des enquêtes O/D réalisées auprès des conducteurs de VL entre 1995 et 2000. Les O/D ont été sélectionnées selon trois principes :

- tous les itinéraires ont été enquêtés la même année, à la même période, le même type de jour
- le volume de trafic doit être suffisamment important pour que les flux par motif et par itinéraire aient un sens
- il n'y a pas eu de modification importante du niveau de service depuis la date d'enquête, ni de grands changements dans les temps de parcours

Les résultats fournis par les CETE pour chaque itinéraire et chaque O/D sont le trafic global, sa décomposition par motifs (selon cinq catégories : Travail fixe, Affaires personnelles, Loisirs et Affaires professionnelles, distinguant pour 60% des O/D le fait que le coût du déplacement soit à la charge ou non de l'employeur – cette distinction n'a pas été analysée par la suite afin de conserver une base de données suffisamment conséquente) et la description du trajet – longueur par type de route (autoroute concédée, autoroute non concédée, RN 2x2 ou 2x3 voies, RN à 7m, RD, traversées d'agglomérations), temps de parcours, et prix du péage.

Dans cette base de donnée, le calcul des temps et distances de parcours a été effectué à l'aide du calculateur d'itinéraires MICHELIN 2003, en faisant l'hypothèse implicite que les distances et temps à l'année de l'enquête sont les mêmes qu'en 2003 (il a été soigneusement vérifié que les routes empruntées lors de ce calcul existaient au moment de l'enquête). Les temps de parcours renseignés ici sont donc des temps de parcours « moyens », intermédiaires entre les temps de parcours à vide (par exemple ceux utilisés en entrée des modèles d'affectation), et les temps de parcours observés en situation congestionnée.

Le calcul des coûts, incluant la valeur des malus, est effectué en francs 2000. La base permet donc de calibrer des valeurs du temps 2000.

Au final, 157 relations O/D sont disponibles, dont 118 ne disposent que de deux itinéraires, 31 de trois itinéraires, 7 de quatre itinéraires et 1 de cinq itinéraires, soit un total de 362 itinéraires. Ces relations O/D représentent un trafic total de près de 185 000 véhicules/jour. Les itinéraires sont répartis de manière relativement homogène entre des distances de 20 km et de 800 km : la base nous permettra donc de traiter tous les types d'O/D hormis les relations de très courte distance. Il s'agit par contre uniquement d'O/D nationales.

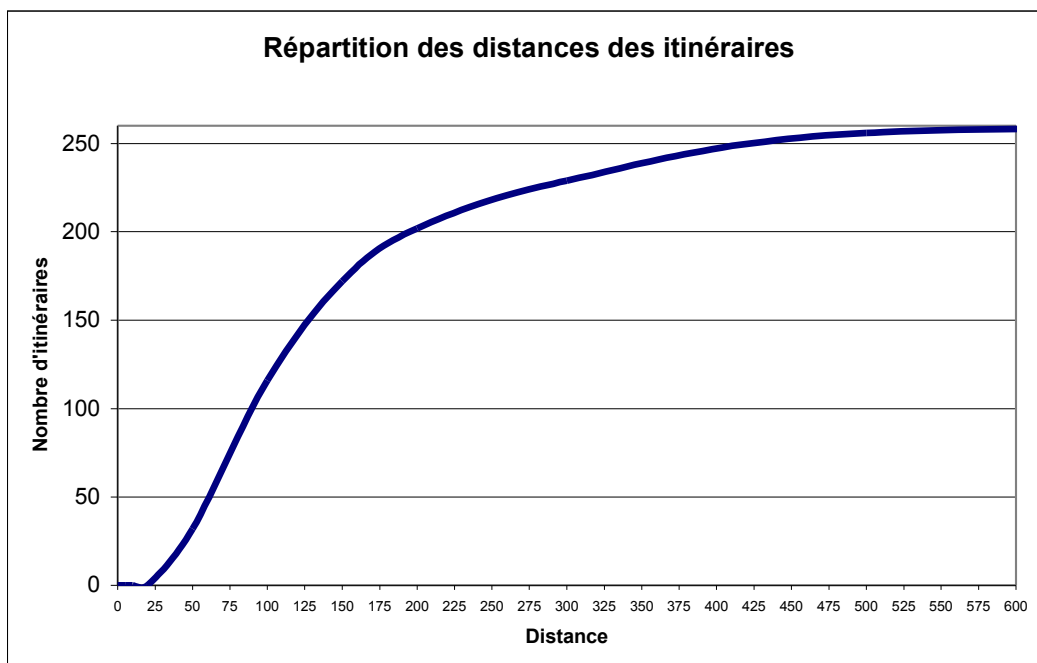


Illustration 7 : Répartition des distances des itinéraires de la base nationale

Les trajets sont décrits selon 6 types de voies. Le tableau 1 décrit le nombre d'itinéraires (sur les 362) selon la proportion de chacun de ces types de route dans le trajet total. On constate que seuls deux types de voies, à savoir les autoroutes concédées et les RN à 7 m comprennent des proportions majoritaires (plus de 50 %) des trajets. Pour 248 itinéraires, les RD constituent moins de 5 % du trajet : l'analyse spécifique du malus qu'elles provoquent risque donc d'être difficile à analyser. Ceci est également le cas, dans une moindre mesure, des autoroutes non concédées (243 trajets) et des RN à 2x2 ou 2x3 voies (211 trajets). Enfin, les longueurs de traversées d'agglomération sont variables dans un spectre de 0 à 20 % du trajet : le malus devrait pouvoir y être correctement calibré.

% du trajet total	Longueur autoroute concédée	Longueur autoroute non concédée	Longueur RN à 2x2 ou 2x3 voies	Longueur autres RN à 7 m	Longueur RD	Longueur traversées d'agglomérations
0 %	191	187	159	42	154	7
De 0 à 5 %	0	46	39	63	94	113
De 5 à 10 %	2	43	33	24	26	116
De 10 à 20 %	2	37	63	26	31	104
De 20 à 30 %	10	31	34	36	14	16
De 30 à 40 %	15	9	16	27	13	4
De 40 à 50 %	15	3	7	17	9	1
De 50 à 60 %	22	3	8	29	7	1
De 60 à 70 %	26	3	3	25	7	0
De 70 à 80 %	32	0	0	39	4	0
De 80 à 90 %	33	0	0	20	2	0
De 90 à 100 %	14	0	0	14	1	0
Total	362	362	362	362	362	362

Tableau 1 : Répartition des longueurs de chaque type de route au sein des itinéraires de la base nationale

Sur le tableau suivant, les 157 O/D ont été ventilées en fonction de la part des 4 motifs. On note une dispersion relative : si le motif loisirs se situe dans la plage 0 % à 10 % et le motif affaires personnelles dans la plage de 10 à 30 %, le motif travail varie entre 0 et 40 % et le motif affaires professionnelles entre 20 et 70 %. La variabilité

se situe donc principalement sur ces deux derniers motifs. Or, malgré une forte dispersion, la part de ces deux motifs est fortement liée à la distance, ainsi qu'on peut le constater sur l'illustration 8. Si la valeur du temps est variable en fonction de la distance, il ne faudra donc pas s'attendre à une forte amélioration des modèles par la prise en compte des motifs.

% du trajet total	Loisirs	Affaires personnelles	Travail fixe	Affaires professionnelles
0 %	11	2	11	0
De 0 à 10 %	95	7	32	3
De 10 à 20 %	33	43	43	11
De 20 à 30 %	7	66	32	17
De 30 à 40 %	7	25	22	28
De 40 à 50 %	2	11	15	33
De 50 à 60 %	0	1	2	32
De 60 à 70 %	0	1	0	25
De 70 à 80 %	0	1	0	5
De 80 à 90 %	2	0	0	3
De 90 à 100 %	0	0	0	0
Total	157	157	157	157

Tableau 2 : Répartition des motifs de déplacement au sein de la base nationale

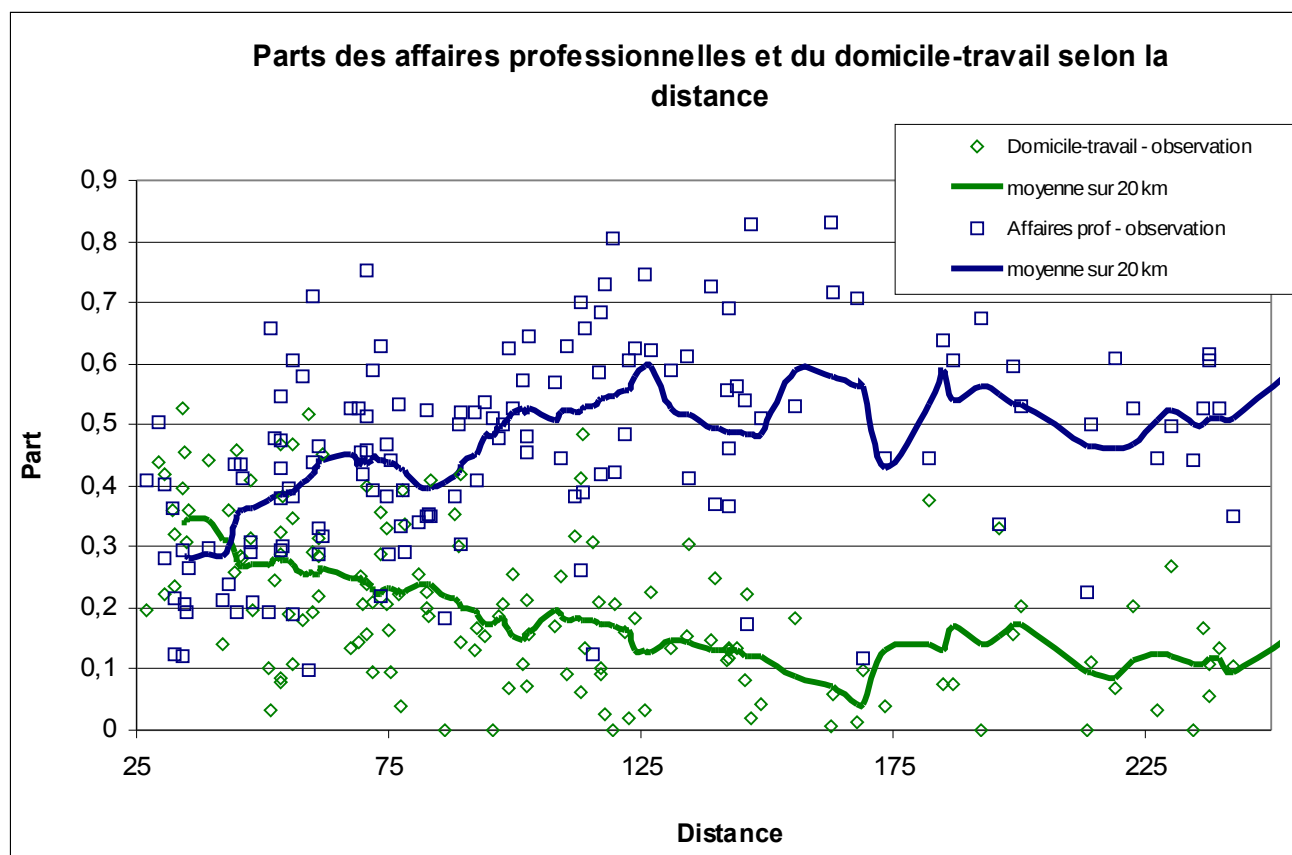


Illustration 8 : Parts des affaires professionnelles et du domicile-travail selon la distance au sein de la base nationale

2.2 - Base « CETE NC » (2 coupures sélectionnées en Haute Normandie)

Une seconde base de données de type « préférences révélées » a également été constituée, à l'aide de données plus récentes sur la zone d'action du CETE Normandie-Centre. 5 enquêtes ont ainsi été retenues, de manière à constituer deux coupures distinctes sur le réseau, et composées d'itinéraires présentant des caractéristiques très distinctes et diversifiées (7m, route express, autoroute concédée...) :

- coupure Est : A13 et RD6014, entre Paris et Rouen (6566 véhicules enquêtés)
- coupure Sud : A13, RN154, A28; RD438 (10616 véhicules enquêtés)

Ces coupures sont représentées sur la carte ci dessous :

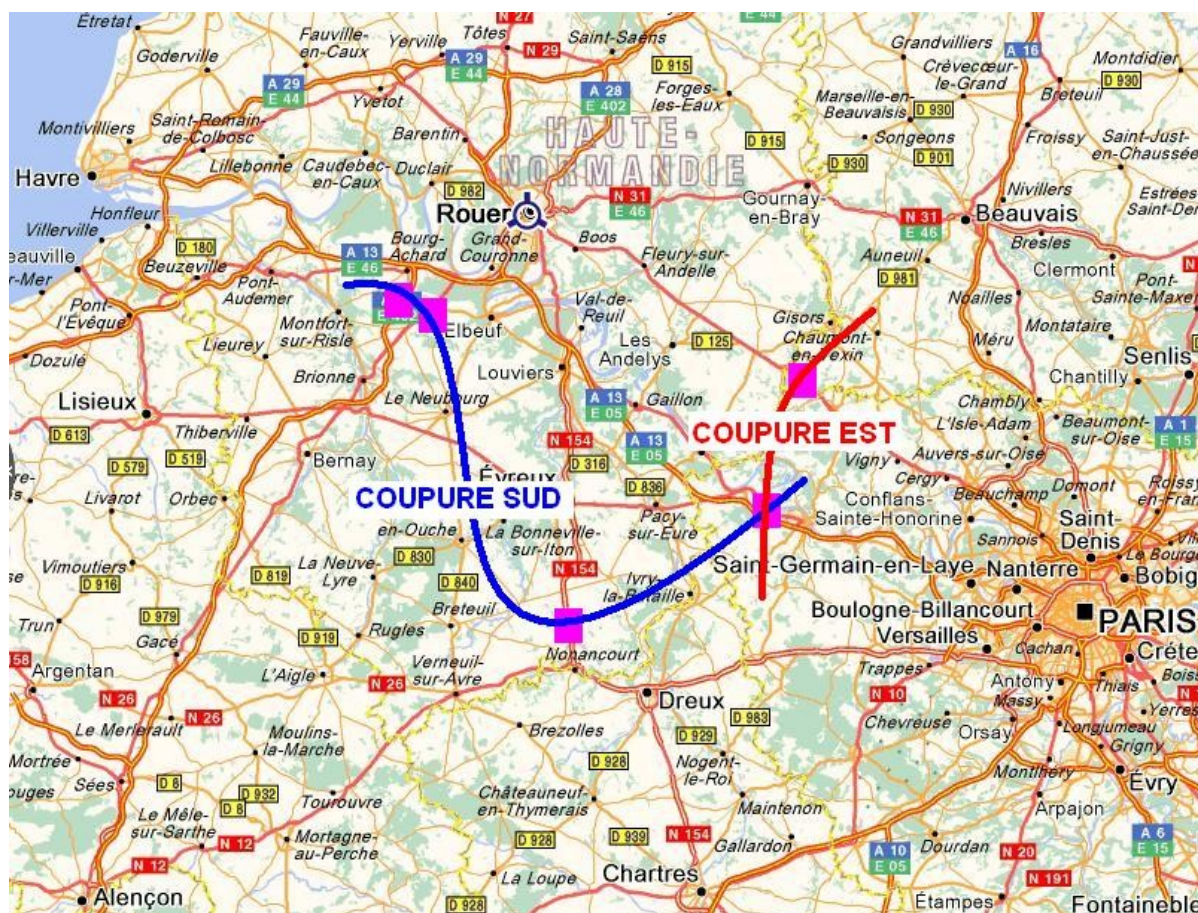


Illustration 9 : Présentation des deux coupures retenues pour la base « CETE NC »

La répartition des distances des déplacements interceptés sur ces deux coupures est la suivante :

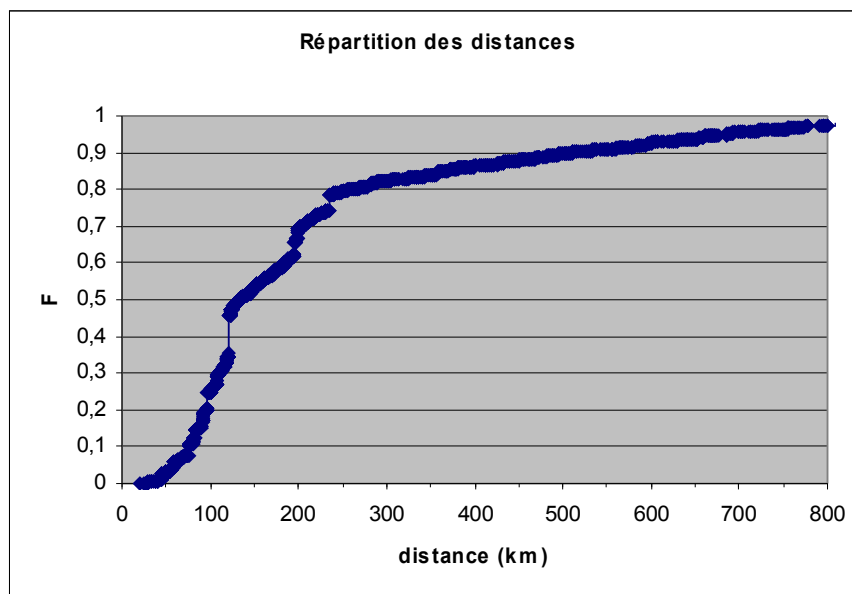


Illustration 10: Répartition des distances des itinéraires de la base CETE NC

Quelques O/D très importantes se distinguent, notamment Paris/Rouen, Paris/Le Havre, et Mantes/Rouen. Leur importance relative dans la base imposera un certain nombre de précautions dans les interprétations. De plus, vu la taille des échantillons, cette base n'est pas pertinente sur les courtes distances (inférieures à 50 km).

L'offre de transport se base sur la description du réseau routier national 30 000 Arcs, entretenue par le SETRA. Pour les besoins spécifiques de l'étude, les longueurs, les types d'arcs et les coûts de parcours ont été revus sur les axes ayant fait l'objet d'enquêtes O/D. Les temps renseignés sont les temps de parcours à vide obtenus à partir de la table de typologie du SETRA. Les calculs de niveau de service par itinéraire sont basés sur la méthode suivante :

- calcul du plus court chemin (en temps de parcours) de la commune origine à l'emplacement de l'enquête et détermination des temps et coût sur ce trajet ;
- calcul du plus court chemin (en temps de parcours) de l'emplacement de l'enquête à la commune destination et détermination des temps et coût sur ce trajet ;
- sommation des deux résultats.

Cette recombinaison des itinéraires permet donc de calibrer les modèles de choix d'itinéraire sur des trajets 2010, avec des valeurs monétaires exprimées en euros 2000. Elle introduit un risque d'incohérence interne dans les modèles :

- dans le cas du modèle prix-temps, le calcul devrait se baser sur l'intégralité des itinéraires efficaces (voir paragraphe 1.1). Or rien ne garantit que notre méthode de calcul des itinéraires permette de les avoir complètement identifiés. Cependant, une hypothèse sur la valeur du temps devrait être émise pour effectuer une recherche plus complète, fondée sur les principes du prix-temps, ce qui imposerait la réalisation du calibrage de manière itérative.
- dans le cas des modèles stochastiques utilisés, nous nous basons sur l'hypothèse d'indépendance entre les alternatives (modèles de type Logit). Or les itinéraires trouvés peuvent être identiques sur une portion importante du trajet (par exemple, pour une O/D Rouen – Lyon les itinéraires A13-A6 et RD 6014-A6 sont communs sur 450 des 600 km), ce qui invalide l'hypothèse pour une portion des O/D (surtout sur les longues distances). L'incohérence provient également du fait que la donnée collectée consiste en l'observation de points de passages, alors que nous cherchons à reproduire des choix d'itinéraire. Pour construire un modèle cohérent, il faudrait donc parvenir à lister l'ensemble des itinéraires passant par

l'emplacement de chaque enquête. Ceci nécessite la disponibilité d'un outil de recherche de chemins stochastique qui n'était pas disponible au moment de l'étude. Les obstacles pour assurer la cohérence interne des deux modèles nous ont conduit à retenir l'approche simple exposée, qui correspond à l'intuition. Elle n'en comporte pas moins un risque d'erreur.

Pour la base nationale, l'approche manuelle de sélection des itinéraires permet de le limiter : la connaissance du territoire par les chargés d'étude apporte une certaine garantie quant à la sélection de l'intégralité des itinéraires efficaces ; les O/D pour lesquelles les choix d'itinéraire sont imbriqués de manière trop complexe ne sont simplement pas retenues.

Pour la base CETE NC, ce risque est plus complexe à limiter. Une première étape a consisté en la suppression des O/D pour lesquelles seul un itinéraire était efficace (ce qui pose également des problèmes pour le calcul de la log vraisemblance dans le cas du modèle prix-temps). La deuxième étape est constituée par un traitement spécifique des deux coupures :

- en ce qui concerne la coupure est, seules deux enquêtes sont disponibles. Le filtre effectué sur les O/D revient donc à ne conserver que les O/D pour lesquelles les deux itinéraires par A13 et la RD 6014 sont efficaces. Or, les O/D de la RD 6014 sont très majoritairement entre la Haute-Normandie et l'Île-de-France, sur des courtes et moyennes distances. Les portions de trajet communes aux deux itinéraires s'en trouvent automatiquement limitées ; sur ces O/D mettant en concurrence A13 et RD 6014, il n'y a pas d'autre alternative réellement efficace.
- en ce qui concerne la coupure sud, seules les enquêtes sur la RD 438 et l'A28 peuvent conduire à une remise en cause de l'hypothèse d'indépendance entre les alternatives. Il peut en effet y avoir un phénomène d'évitement de péage via la RD 438 : les usagers en provenance de Rouen iraient emprunter l'A28 au sud de l'enquête O/D, en passant par l'emplacement de l'enquête sur la RD 438. Ce type de comportement reste relativement circonscrit à une faible portion des O/D conservées (10% environ). Par ailleurs, le problème de non sélection de l'intégralité des itinéraires pourrait également concerner ces deux enquêtes. Pour des trajets de type Est-Ouest comme Rouen-Nantes, le passage par A13 via Caen puis A84 constitue ainsi une alternative au passage par A28, non prise en compte dans la coupure. Ces O/D ont ainsi été retirées de l'échantillon.

2.3 - Enquête de préférences déclarées sur les traversées vosgiennes

Cette enquête, réalisée dans le cadre de l'étude d'opportunité sur le doublement du tunnel Maurice Lemaire, a fait l'objet d'un rapport détaillé, intitulé Enquêtes sur les déplacements du massif vosgien - Résultats du questionnaire internet – Août 2010. Les principaux résultats en sont rappelés ici, car ils présentent un double intérêt :

- d'une part il s'agit de données très récentes ;
- d'autre part il s'agit cette fois de préférences déclarées et non révélées (ce qui est le cas de toutes les autres bases de données utilisées dans cette étude). Les préférences déclarées sont notamment pertinentes lorsqu'on veut tester la réaction des usagers à une nouvelle alternative.

Dans le cas présent, il s'agit d'anticiper comment les usagers réagiraient à une variation du coût du péage d'un tunnel permettant d'éviter de devoir franchir des cols pour traverser les Vosges. Etant donné que l'analyse de préférences révélées (constituée par des enquêtes O/D) n'aurait été basée que sur la valeur actuelle du péage, il aurait été difficile d'être certain que la sensibilité au péage soit correctement captée. L'enquête de préférences déclarées, via l'introduction de scénarios avec différentes valeurs du péage, permet de remédier à cette lacune.

Des enquêtes O/D routières ont été effectuées dans le massif vosgien en juin 2010.

Elles ont été représentées sur la carte suivante :



Illustration 11 : Localisation des enquêtes de la base de préférence déclarées

Lors de ces enquêtes, un document a été distribué aux personnes interviewées, exposant les raisons de la réalisation de l'enquête et leur proposant de se rendre sur un site internet pour répondre à quelques questions complémentaires.

Cette enquête internet reposait sur un enchaînement de 3 questions :

- une première série de questions sur les caractéristiques de la personne : activité, revenus, âge, sexe et lieu de résidence ;
- une deuxième série de questions sur les choix d'itinéraire : la personne se voit proposer successivement 5 déplacements, pour lesquels différents itinéraires alternatifs sont possibles. Chaque itinéraire est décrit selon :
 - le motif de déplacement : Domicile – Travail/études (à un rythme quotidien ou hebdomadaire), affaires professionnelles (en distinguant si les frais sont pris en charge ou non par l'employeur), affaires personnelles ou loisir ;
 - l'origine et la destination du déplacement parmi les 32 couples O/D indiqués dans le tableau suivant :

Origine / destination		Origine / destination	
Saint-Dié-des-Vosges	Strasbourg	Epinal	Mulhouse
Saint-Dié-des-Vosges	Sélestat	Epinal	Obernai
Saint-Dié-des-Vosges	Obernai	Epinal	Sélestat
Saint-Dié-des-Vosges	Mulhouse	Epinal	Strasbourg
Saint-Dié-des-Vosges	Colmar	Metz	Mulhouse
Gérardmer	Obernai	Metz	Obernai
Gérardmer	Sélestat	Metz	Sélestat
Gérardmer	Strasbourg	Nancy	Mulhouse
Lunéville	Mulhouse	Nancy	Obernai
Lunéville	Obernai	Nancy	Sélestat
Lunéville	Sélestat	Paris	Colmar
Lunéville	Colmar	Paris	Mulhouse
Colmar	Epinal	Paris	Sélestat
Colmar	Metz	Paris	Obernai
Colmar	Nancy		
Remiremont	Obernai		
Remiremont	Sélestat		
Remiremont	Strasbourg		

Tableau 3 : O/D retenues dans le questionnaire

– la longueur du trajet, le temps de parcours, le coût du péage et le coût du carburant : parmi ces éléments, les trois derniers sont variables en fonction des évolutions possibles du réseau routier et autoroutier, de la politique de tarification du tunnel Maurice Lemaire et des fluctuations du cours du baril de pétrole. Les valeurs proposées sont contrôlées par la génération de nombres aléatoires à chaque prise de décision. Les valeurs actuelles de temps et de longueur des parcours ont été déterminées à l'aide des informations du modèle de trafic MODESTE (décrit dans le document [14]) du CETE de l'Est.

- la dernière question consiste à mettre à disposition des usagers un espace d'expression, leur permettant d'effectuer des remarques sur les conditions de déplacement dans le massif vosgien.

Après apurement des données inutilisables et des entrées multiples (personnes ayant répondu plusieurs fois au questionnaire du fait de difficultés d'utilisation du site internet), 248 personnes ont répondu au questionnaire internet dont 118 ont laissé un commentaire. Le taux de réponse est donc particulièrement faible (moins de 2 %), étant donné que l'adresse du site a été distribuée à près de 14 000 personnes. Cependant, comme 5 choix étaient proposés à chaque personne, 1326 choix ont été au total proposés, dont 1028 ont fait l'objet d'une réponse. Sur les 248 personnes ayant répondu au questionnaire, 213 personnes ont renseigné en moyenne 4,8 choix d'itinéraire (27 personnes ont renseigné moins de 5 itinéraires) : globalement, les personnes ayant décidé de répondre au questionnaire se sont donc correctement appropriées l'enquête.

En ce qui concerne les caractéristiques des individus, l'échantillon est globalement très varié et une analyse catégorielle des résultats est donc possible. Le déséquilibre le plus marquant concerne la répartition hommes / femmes valant 70 % / 30 %.

Les choix proposés se répartissent de manière équilibrée sur toutes les O/D, à l'exception notable des déplacements longue distance, en lien avec Paris.

L'illustration 12 présente les motifs recueillis pour chaque O/D : la répartition est équilibrée entre le Domicile – Travail (25 %), les loisirs (25 %), les affaires personnelles (31 %) et les affaires professionnelles (19 %).

La répartition est également satisfaisante : hormis pour les flux parisiens, tous les motifs sont présents pour chacune des O/D.

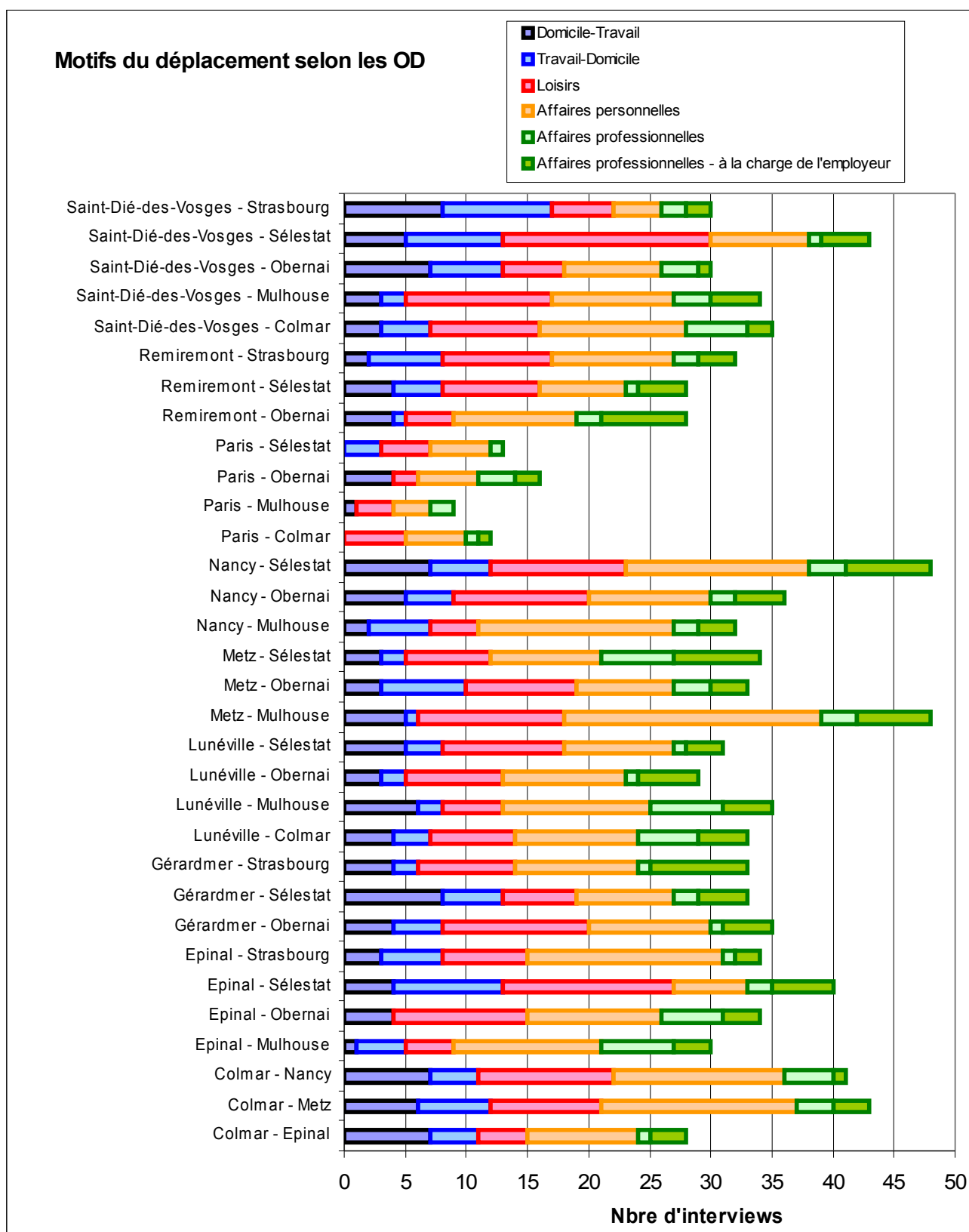


Illustration 12 : Motifs du déplacement par O/D

Dans le tableau 4, pour chacune des choix effectués, nous présentons comment s'établit la concurrence le tunnel Maurice Lemaire et l'itinéraire alternatif de coût généralisé le moins élevé (en se basant sur l'instruction relative aux méthodes d'évaluation économique des projets routiers interurbains de mai 2007 [2] pour ce calcul), selon deux axes : la différence de temps et la différence de coût. La quantité de réponses pour chaque classe de différence de temps et de différence de coût est représentée, ainsi que la part d'utilisateurs déclarant prendre le tunnel dans ces conditions.

Différence de temps	Données	Différence de coût										Total
		de - 1 à + 0 €	de + 0 à + 1 €	de + 1 à + 2 €	de + 2 à + 3 €	de + 3 à + 4 €	de + 4 à + 5 €	de + 5 à + 6 €	de + 6 à + 7 €	de + 7 à + 8 €	de + 8 à + 9 €	
Gain de 20 à 25 min	Nombre réponses	21	21	19	29	23	28	37	14	13	6	211
	% TML	48%	57%	58%	34%	43%	39%	32%	29%	23%	17%	40%
Gain de 15 à 20 min	Nombre réponses		2	12	14	14	9	8	17	19	10	105
	% TML		50%	50%	7%	50%	44%	13%	24%	47%	20%	33%
Gain de 10 à 15 min	Nombre réponses	26	22	41	49	43	35	39	38	21	19	333
	% TML	65%	45%	59%	47%	30%	34%	31%	32%	19%	21%	39%
Gain de 5 à 10 min	Nombre réponses		6	10	14	18	14	13	22	15	19	131
	% TML		67%	70%	36%	39%	57%	38%	23%	13%	16%	35%
Gain de 0 à 5 min	Nombre réponses		5	6	17	16	12	18	16	14	29	133
	% TML		80%	50%	24%	25%	8%	17%	25%	7%	10%	20%
Total	Nombre réponses	47	54	76	109	100	89	107	90	63	73	808
	% TML	57%	57%	67%	39%	41%	40%	31%	32%	30%	18%	34%

Tableau 4 : Utilisation du TML en fonction de la différence de coût et de temps avec l'itinéraire alternatif de coût généralisé minimum

Ce tableau nous permet de comprendre comment la base de données est structurée :

- 808 réponses, soit 80 %, sont dans des situations de concurrence inscrites dans le domaine [-0 min , -25 min] x [-1 € , +9 €] ;
- à l'intérieur de ce domaine, les situations sont très diverses, ce qui permet de bien capter toute la variété des comportements. Il semble cependant qu'il y a un sur-échantillonnage autour d'un gain de temps de 10 à 15 minutes ;
- non représentées ici, on trouve également 58 situations pour lesquelles le tunnel est moins rapide (entre + 0 et +10 minutes).
- il peut y avoir des comportements semblant non efficaces : ainsi, pour les concurrences où le tunnel est à la fois moins cher et plus rapide, il devrait être sélectionné par tous les usagers. Les modèles stochastiques, tenant compte de ce type de problème, auront donc un avantage intrinsèque sur les modèles déterministes.
- Il semble y avoir un seuil autour de 2 € qui permettrait au tunnel de capter une majorité des usagers dès qu'il fait gagner du temps.

Enfin, les commentaires ont été analysés et comparés, ce qui a permis de définir 5 catégories d'usagers :

- 24 personnes réclament principalement une baisse des prix du péage
- 25 personnes mettent l'accent sur les problématiques d'état de la route et de sécurité routière
- 33 personnes se sentent fortement pénalisées par les bouchons aux heures de pointe
- 36 personnes considèrent que le trafic PL est trop élevé dans les cols et induit tout un ensemble de problèmes : suivi de véhicules lents, sentiment d'insécurité, blocages de la route en hiver,...
- 130 personnes n'ont pas laissé de commentaires.

Chapitre 3 : Analyse des résultats des modèles déterministes

3.1 - Application des valeurs de l'instruction

Avant même de chercher à calibrer les valeurs du temps et les malus, il nous a paru intéressant, comme point de départ, d'observer la qualité des modèles constitués à partir des valeurs du temps et de malus de l'instruction.

On rappelle que les valeurs de l'instruction sont les suivantes (attention : dans l'instruction les valeurs sont exprimées avec la moyenne, ici nous donnons la médiane) :

Valeurs du temps		Malus
Médiane	Sigma	
$d < 20 \text{ km} : 8,7$ $d < 50 \text{ km} : 11,8$ $d < 400 \text{ km} : 13,6 + 0,027 \times d$ $d > 400 \text{ km} : 30,3$	0,5	routes express : 0,7 c€ / km bidirectionnelle (7m) : 5,4 c€ / km traversée d'agglomération : 5,4 c€/km

Tableau 5 : Valeurs du temps et malus de l'instruction

Les résultats obtenus, sur les 3 bases de données, sont les suivants :

	Base nationale	Base CETE NC	Préférences déclarées
Log0	-149817	-50932	-934
LogV	-151450	-56356	-1748
Coefficient de régression	0,9556	0,5367	1,0175
Coefficient de détermination	0,6505	0,3109	0,6062

Tableau 6 : Résultats obtenus sur les trois bases avec les valeurs du temps et malus de l'instruction

Sur la base des log-vraisemblances, le modèle fondé sur les « valeurs de l'instruction » est moins efficace qu'un modèle simple supposant l'équiprobabilité des alternatives (caractérisé par le log0), quelle que soit la base de données utilisée. La reproduction des données agrégées est de qualité variable selon les bases de données, mais toujours insatisfaisante au vu des coefficients de détermination, traduisant tous une forte dispersion des résultats. La valeur du coefficient de régression sur les données de la base nationale semble néanmoins indiquer un comportement suffisamment réaliste « en moyenne » et sur les flux à fort trafic.

3.2 - Malus vs valeurs du temps

Selon les principes présentés au chapitre 1, il n'est pas possible avec les outils à disposition de calibrer le malus dans les modèles déterministes lorsque les O/D ont plus de deux itinéraires : comme les malus interagissent avec l'efficacité des itinéraires et leur ordre de priorité, il faudrait que le solveur soit en mesure de réordonner les itinéraires selon leur coût à chaque itération, ce qui n'est pas le cas dans l'outil utilisé. Les malus seront donc calibrés à l'aide d'une sous-base constituée des O/D à deux itinéraires.

Dans un premier temps, l'objectif est d'identifier s'il est plus efficace de calibrer les malus ou les valeurs du temps. On part donc des valeurs du temps de l'instruction, et on calibre les malus. En parallèle, on part des malus de l'instruction, et on calibre les valeurs du temps. Puis, on compare les résultats obtenus par ces deux méthodes. Les résultats obtenus lorsque l'on calibre les deux paramètres en même temps ont également été estimés (voir 3.3).

			VDT instruction Calage malus	Calage VDT Malus instruction
BASE NATIONALE	valeur du temps	Médiane	d < 20 km : 8,7 d < 50 km : 11,8 d < 400 km : 13,6 + 0,027 x d d > 400 km : 30,3	d < 50 km : 4,8 d < 400 km : 4,8 + 0,176 x d d > 400 km : 75,0
		σ	0,5	1,89
	Malus	autoroute concédée	0	0
		autoroutes non concédées	0	0,7
		route express	0	0,7
		bidirectionnelle (7m)	2,4	5,4
		traversée d'agglomération	13,5	5,4
	Résultats	Log0	-85903	-85903
		LogV	-78453	-69555
		Coefficient de régression	1,0041	0,9458
		Coefficient de détermination	0,884	0,898
BASE CETE NC	valeur du temps	Médiane	d < 20 km : 8,7 d < 50 km : 11,8 d < 400 km : 13,6 + 0,027 x d d > 400 km : 30,3	d < 50 km : 0+1,603 x d d < 400 km : 80,1+ 0 x d d > 400 km : 80,1
		σ	0,5	1,56
	Malus	autoroute concédée	0	0
		autoroutes non concédées	0	0,7
		route express	0	0,7
		bidirectionnelle (7m)	7,9	5,4
		traversée d'agglomération	13,1	5,4
	Résultats	Log0	-31492	-31492
		LogV	-23330	-17446
		Coefficient de régression	0,54	0,83
		Coefficient de détermination	0,43	0,93

Tableau 7 : Comparaison de l'efficacité du calibrage des malus et des valeurs du temps de manière séparée

On s'aperçoit dans les deux cas que le modèle obtenu par calibrage des valeurs du temps est bien meilleur que celui obtenu par calibrage des malus, et ceci s'observe aussi bien au niveau désagrégé (log-vraisemblance supérieure) qu'au niveau agrégé (coefficient de corrélation supérieur).

Par ailleurs, les valeurs du temps obtenues diffèrent sensiblement des valeurs du temps recommandées :

- le paramètre sigma est 3 à 4 fois plus élevé, ce qui traduit une dispersion plus forte des valeurs du temps (résultat similaire sur les deux bases de données) ;
- les médianes ont tendance à évoluer plus fortement avec les distances parcourues.

Ce résultat pose question sur les temps de parcours utilisés : une hypothèse pour expliquer cette divergence avec les valeurs usuelles serait de supposer que des temps de parcours trop faibles sur les bidirectionnelles conduisent à devoir utiliser des valeurs du temps élevées pour reproduire la différence réelle entre les itinéraires sur les trajets longue distance. Deux éléments complémentaires sont présentés ici pour illustrer la crédibilité et les limites de nos résultats :

- d'une part, cette analyse était basée sur l'hypothèse, correspondant aux pratiques habituelles, que le malus est perçu comme un coût supplémentaire. Cependant, il peut également correspondre à une incertitude sur les temps de parcours (suivi de véhicule lent sur route bidirectionnelle par exemple). Les modèles testés sur la base nationale n'ont cependant pas donné de résultats cohérents, ce qui semble indiquer qu'il n'y a pas un biais global par type de voie et que l'estimation des temps de parcours est donc suffisamment pertinente pour reproduire les choix d'itinéraire ;
- d'autre part, les vitesses moyennes de parcours par type de voie utilisées pour la base nationale sont indiquées sur le tableau 8. Ces éléments conduisent à rejeter clairement l'hypothèse de temps de parcours trop faibles (notamment la médiane de 70 km/h sur les bidirectionnelles).

Type de voie	Distribution des vitesses (km/h)		
	1er décile	médiane	9ème décile
autoroute concédée	110	115	120
autoroutes non concédées	90	110	120
route express	75	90	105
bidirectionnelle (7m)	60	70	80
traversée d'agglomération	30	30	40

Tableau 8 : Distribution des vitesses sur la base nationale

En conclusion, il ressort de ce premier examen des données que calibrer les valeurs du temps est beaucoup plus efficace que calibrer les malus, ce qui rend finalement un peu secondaire le choix de ces derniers et peut également conduire à des distributions sensiblement différentes de celles couramment utilisées.

3.3 - Choix des malus et formulation de la fonction de coût

Comme on l'a vu précédemment, le choix des valeurs de malus est de second ordre par rapport aux choix de valeur du temps. Il n'en reste pas moins qu'il est nécessaire de choisir des valeurs pour ces malus avant d'aller plus loin dans notre étude.

Au paragraphe précédent, nous avons évoqué la difficulté posée par le calibrage des malus sur une base composée de plus de deux itinéraires. En conséquence, nous choisissons de retenir des valeurs de malus à partir des bases à deux itinéraires uniquement.

Par ailleurs, utiliser des bases « locales » (CETE NC et préférences déclarées) pour retenir des valeurs de malus au niveau national est ici moins pertinent, c'est la raison pour laquelle nous nous limitons ici à la base nationale.

Ainsi, nous choisissons ici de calibrer le malus en laissant également libre la valeur du temps. Les résultats obtenus sont les suivants :

			Calage VDT Calage malus
BASE NATIONALE	Valeur du temps	Médiane	d < 20 km : 4.0 d < 50 km : -0.1+0,204 x d d < 400 km : -9.6 + 0.394 x d d > 400 km : 148.1
		σ	1,76
	Malus	autoroute concédée	0
		autoroutes non concédées	0
		route express	0
		bidirectionnelle (7m)	2,7
		traversée d'agglomération	10,7
	Résultats	Log0	-85903
		LogV	-69690
		Coefficient de régression	0,9419
		Coefficient de détermination	0,9037

Tableau 9 : Résultats du calibrage des malus sur la base nationale à deux itinéraires

Les valeurs trouvées pour les malus (en italique) seront retenues dans toute la suite de l'étude. Cette hypothèse implique un biais dans le calibrage des valeurs du temps qui sera effectué ensuite, puisque la valeur du temps et celle du malus sont dépendantes l'une de l'autre. Ainsi, pour s'assurer de la validité des résultats sur la valeur du temps, il sera nécessaire de vérifier qu'un nouveau calibrage des malus en partant des valeurs du temps trouvées au final fournira des valeurs relativement proches de celles retenues ici. Cette vérification devra donc permettre de valider les valeurs de malus et de valeur du temps retenues.

Elle sera effectuée au paragraphe 3.7.

Au niveau des valeurs du temps trouvées (même s'il ne s'agit que la base à deux itinéraires, on peut d'ores et déjà observer quelques tendances), on observe ici un résultat similaire à celui du 3.3, à savoir que la médiane des valeurs du temps est plus faible que celle de l'instruction pour les faibles distance (inférieures à 50 km), et plus importante ensuite. Par ailleurs, le paramètre sigma est nettement plus important que celui de l'instruction, ce qui dénote une grande hétérogénéité dans les comportements de choix d'itinéraire.

La question de la formulation du coût doit être également posée à ce stade et sur la base de données nationale – pour les mêmes raisons de contraintes logicielles. De fait, les calculs sont effectués avec l'hypothèse implicite que l'utilisateur pondère de manière identique les différents éléments du coût (péage, carburant et usure du véhicule) puis leur ajoute un malus en fonction du type de route. Or, la difficulté qu'ont les usagers à estimer correctement les coûts liés à l'usage du véhicule est souvent mis en avant. Le test de pondération distinctes de ces trois paramètres a été effectué ici, et donne des résultats peu concluants, le carburant et l'usure étant fortement corrélés : la pondération du carburant serait négative, ce qui est contre-intuitif.

En dehors du péage, dont la prise en compte semble la plus naturelle (à nuancer toutefois pour les usagers occasionnels qui pouvaient ignorer le coût du trajet à cette époque – 1995 à 2000 – ce qui peut sembler moins évident aujourd'hui), la prise en compte de l'impact des coûts d'usure, et dans un moindre mesure des coûts du carburant, sur le choix d'itinéraire peut être remise en question. Ceci a conduit à tester un modèle où le seul coût considéré est le péage, dont les résultats sont présentés sur le tableau 10. Le test a également été effectué avec un coût égal à la somme du péage et du carburant, mais les résultats sont moins concluants.

			Calage VDT Calage malus Coût=péage
BASE NATIONALE	Valeur du temps	Médiane	$d < 20 \text{ km} : 1,6$ $d < 50 \text{ km} : -0.2 + 0,089 \times d$ $d < 400 \text{ km} : -15,9 + 0.404 \times d$ $d > 400 \text{ km} : 145,6$
		σ	1,93
	Malus	autoroute concédée	0
		autoroutes non concédées	0
		route express	0
		bidirectionnelle (7m)	2,8
		traversée d'agglomération	2,8
	Résultats	Log0	-85903
		LogV	-69176
		Coefficient de régression	0,9422
		Coefficient de détermination	0,914

Tableau 10 : Résultats du calibrage des malus sur la base nationale à deux itinéraires

Trois enseignements ressortent de ce tableau :

- la limitation du coût au seul péage donne des résultats légèrement meilleurs qu'avec une formulation complète du coût, sans que cela représente un gain décisif ;
- dans cette configuration du modèle, le surcoût de malus lié aux traversées d'agglomération disparaît : il est identique aux bidirectionnelles ;
- les résultats sur la valeur du temps sont très similaires à ceux du modèle précédent.

La faible différence de qualité entre les deux modèles, l'absence de différenciation du coût dans les traversées d'agglomération dans le modèle ne tenant compte que du péage, et la possible évolution des comportements vis à vis du carburant depuis la période 1995 - 2000 (augmentation du coût du SP95 et SP98 de plus de 40% en euros courants et de plus de 65 % pour le gazole entre 1999 et 2010) incitent néanmoins à conserver la formulation complète du coût dans la suite du document. Ce résultat conforte par contre également l'importance du calibrage des valeurs du temps, au vu de la remarquable stabilité des valeurs de paramètres obtenus.

3.4 - Influence de la distance et des motifs sur la valeur du temps

3.4.1 - Effet de la distance

L'objectif est d'examiner comment se comporte la distribution vis à vis de la distance : notamment, y-a-t-il une relation linéaire entre la médiane et la distance, ainsi que supposé dans le modèle « instruction » ?

Voici le résultat de deux tests effectués à l'aide de la base nationale, en définissant 6 et 9 classes de distance.

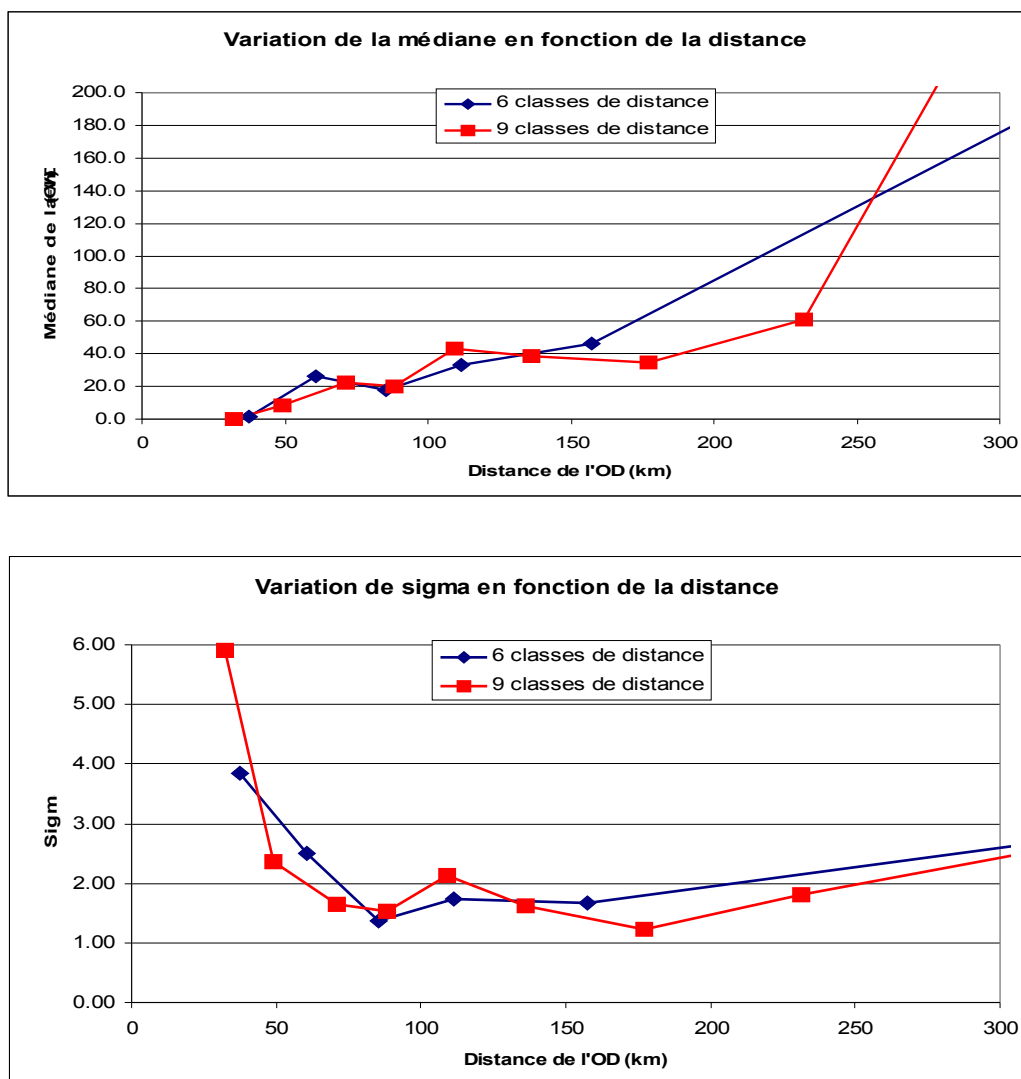


Illustration 13 : Variation de la médiane et du paramètre sigma de la distribution de la valeur du temps en fonction de la distance

Que le modèle comporte 6 ou 9 classes, on observe un comportement similaire, à savoir que la valeur du temps augmente de façon quasi-linéaire selon la distance, ce qui confirme l'hypothèse de l'instruction. Par ailleurs, un seuil autour de 40 km semble exister en dessous duquel les usagers ne semblent pas accepter un surcoût pour leur trajet : la médiane de la valeur du temps est nulle. Enfin, au delà de 250 km, la médiane croît très fortement : les usagers effectuent des choix comme s'ils n'étaient quasiment plus sensibles au péage. Ceci trouve une explication dans la structure des motifs sur ces trajets longue distance : on a notamment vu que les affaires professionnelles y étaient prépondérantes.

En ce qui concerne le paramètre sigma, il semble que celui-ci ait également un comportement identique quelle que soit la segmentation choisie : celui-ci commence par diminuer très fortement sur la (resp. les deux) première(s) classe(s), pour un seuil compris entre 50 et 75 km puis est stable voire en légère augmentation ensuite. Quelle que soit la classe de distance, les valeurs trouvées pour le paramètre sigma sont nettement supérieures à celles de l'instruction. Étant donné que les écarts type naviguent tous dans des valeurs comprises entre 1,5 et 2,5 à l'exception des distances inférieures à 50 km, et au vu de la complexification du modèle qui suivrait d'une valeur de sigma variable en fonction de la distance, nous n'avons retenu et présenté qu'une unique valeur de sigma pour le calibrage final du modèle déterministe (cf. 3.5). Les tests effectués ont par ailleurs mis en évidence une faible amélioration du modèle lorsqu'on faisait cette hypothèse.

3.4.2 - Différenciation selon les motifs

Nous avons également cherché, à partir de la base nationale, à calibrer des modèles intégrant la prise en compte des motifs. Ces modèles sont intéressants à double titre :

- est-ce-que les modèles génériques tenant compte d'une seule distribution de la valeur du temps selon la distance sont aussi performants pour reproduire des données observées que les modèles séparant les motifs ?
- est-ce-que le comportement des usagers permet de mettre en évidence des différences de valeur du temps selon les motifs ?

Les résultats obtenus sont les suivants :

	Avec Motifs	Sans Motifs
Log0	-149817	-149817
LogV	-85687	-88281
Coefficient de régression	0.9613	0.9584
Coefficient de détermination	0.8816	0.869

Tableau 11 : Comparaison des résultats obtenus avec et sans différenciation selon le motifs

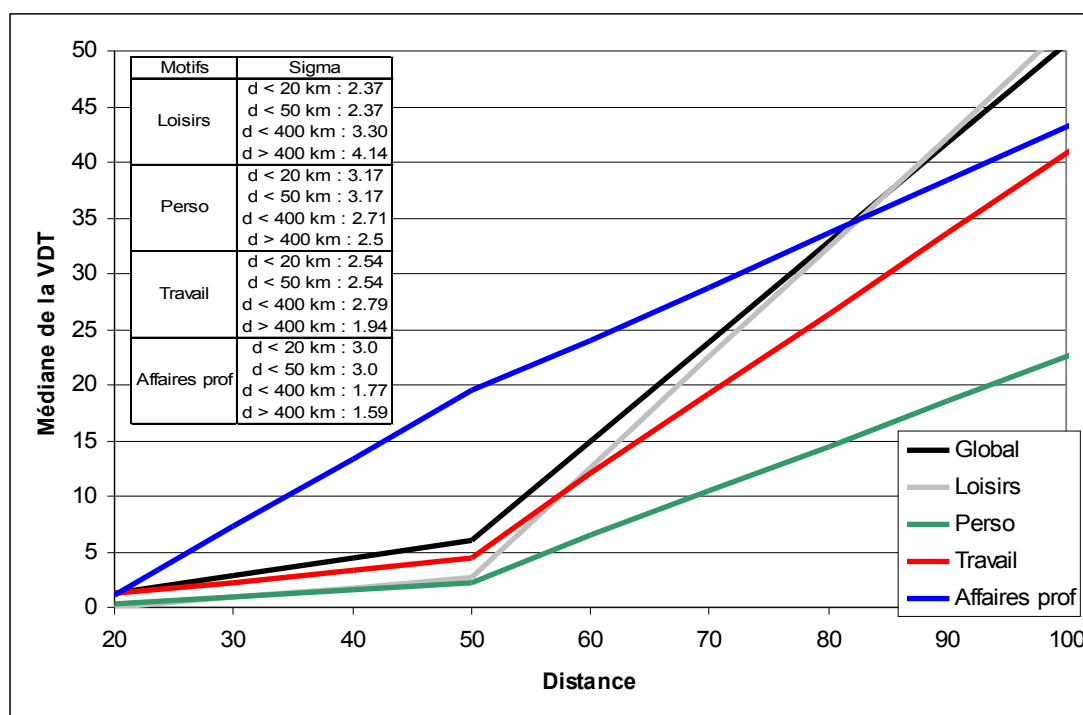


Illustration 14 : Evolution de la valeur du temps (médiane) selon le motif et la distance

Au niveau agrégé, on ne constate pas d'améliorations nettes par rapport au modèle précédemment calibré : sur ce jeu de données, la prise en compte des motifs n'est donc pas essentielle pour caler un modèle d'affectation.

En revanche, la log-vraisemblance augmente nettement, malgré l'augmentation du nombre de paramètres.

Des valeurs du temps distinctes par motifs peuvent donc être clairement identifiées, et les conclusions sont les suivantes pour les médianes :

- pour les distances de 20 à 50 km, les trois motifs hors affaires professionnelles ont des médianes de valeur du temps évoluant entre 0 et 5 €/h ;
- au delà du seuil de 50 km, les médianes de ces trois motifs croissent de manière très différentes : les loisirs croissent le plus fortement pour atteindre 35 €/h pour une distance de 80 km, les domicile-travail atteignent une valeur du temps supérieure à 25 €/h et les affaires personnelles 15 €/h ;
- les affaires professionnelles ont des valeurs du temps supérieures à celles des autres motifs pour les distances de moins de 80 km ; elles sont dépassées par le motif loisirs au delà et convergent avec les flux domicile-travail.

Ces comportements semblent relativement cohérents, à l'exception éventuelle de la croissance de la valeur du temps pour les flux domicile-travail, qui semble très soutenue. Cependant, la part des aller-retours hebdomadaires augmente avec la distance et peut sans doute être un facteur explicatif. De même, la croissance des valeurs du temps des affaires professionnelles pourrait paraître faible par rapport au motif loisirs. La fréquence de déplacement pourrait être également jouer un rôle dans ce phénomène.

En ce qui concerne le paramètre sigma, on note :

- des valeurs similaires pour tous les motifs sur les déplacements de moins de 50 km, autour de la fourchette [2,5 : 3] ;
- une décroissance pour le motif affaires professionnelles avec la distance, marquant une homogénéisation des comportements ;
- a contrario une croissance pour le motif loisirs ;
- une évolution moins marquée et certainement peu significative (faiblesse des échantillons pour les distances de plus de 400 km) pour les motifs affaires personnelles et domicile-travail.

Les évolutions des motifs affaires professionnelles et loisirs semblent cohérentes. Elles peuvent être interprétées par des notions de poids de la rationalisation du budget pour les entreprises avec des coûts croissants dans le premier cas et par les extrêmes disparités de budget disponibles pour les loisirs dans le second cas.

3.5 - Calibrage des valeur du temps

Cette partie de l'étude cherche à définir des valeur du temps qui permettront d'aider les modélisateurs lors de la construction de leurs modèles. Ainsi, nous décidons d'abandonner la décomposition par motif pour le moment, car malgré son intérêt pour améliorer la reconstitution des choix d'itinéraires observés (voir paragraphe 3.4.2), elle serait peu voire pas du tout utilisée en pratique puisqu'il faudrait disposer d'un modèle dont la demande est décomposée selon les motifs, ce qui d'une part est rarement le cas en milieu interurbain ou périurbain et d'autre part représente un effort de modélisation rédhibitoire au vu des pratiques actuelles de constitution des matrices O/D [14]

En revanche, on choisit de conserver une décomposition des valeurs du temps par classe de distance et de conserver les mêmes classes que l'instruction.

A partir des valeur de malus retenues au paragraphe 3.3, les valeurs du temps ont été calibrées sur chacune des trois bases de données. Pour les bases nationale et CETE NC, le résultat obtenu si l'on applique directement le modèle calibré sur la base nationale restreinte à deux itinéraires (modèle obtenu au paragraphe 3.3) est également fourni à titre de comparaison.

			Résultat avec valeurs du temps obtenues au 3.3	Résultat avec calibrage des valeurs du temps	
BASE NATIONALE	Valeur du temps	Médiane	d < 20 km : 4.0 d < 50 km : -0.1+0,204 x d d < 400 km : -9.6 + 0.394 x d d > 400 km : 148.1	d < 20 km : 1.25 d < 50 km : -1.89+0,157 x d d < 400 km : -38.8 + 0,895 x d d > 400 km : 319.2	
		σ	1,76	2,66	
	Malus	autoroute concédée	0	0	
		autoroutes non concédées	0	0	
		route express	0	0	
		bidirectionnelle (7m)	2,7	2,7	
		traversée d'agglomération	10,7	10,7	
	Résultats	Log0	-149617	-149617	
		LogV	-89141	-88281	
		Coefficient de régression	0,9522	0,9584	
		Coefficient de détermination	0,8482	0,869	
BASE CETE NC	Valeur du temps	Médiane	d < 20 km : 4.0 d < 50 km : -0.1+0,204 x d d < 400 km : -9.6 + 0.394 x d d > 400 km : 148.1	d < 20 km : 32,5 d < 50 km : 1,62 x d d < 400 km : 81,2 d > 400 km : 81,2	
		σ	1,76	1,62	
	Malus	autoroute concédée	0	0	
		autoroutes non concédées	0	0	
		route express	0	0	
		bidirectionnelle (7m)	2,7	2,7	
		traversée d'agglomération	10,7	10,7	
	Résultats	Log0	-50932	-50932	
		LogV	-21682	-18679	
		Coefficient de régression	0,76	0,88	
		Coefficient de détermination	0,82	0,95	
BASE préférences déclarées	Valeur du temps	Médiane		13,1	
		σ		2,39	
	Malus	Calibré par axe			
	Résultats	Log0			-1614
		LogV	-700		
		Coefficient de régression	1,007		
		Coefficient de détermination	0,7		

Tableau 12 : Résultats du calibrage des valeurs du temps

Le calibrage sur la base de données nationale complète donne des valeurs du temps assez différents de celles de la base restreinte à deux itinéraires. Pour la base nationale, les médianes des valeurs du temps sont encore plus faibles pour les courtes distance et encore plus fortes pour les longues distances. Le paramètre sigma a encore augmenté, ce qui était prévisible puisque la base est plus importante et donc les comportements encore plus hétérogènes. Le résultat final a bien été amélioré, mais de manière relativement faible.

Pour la base CETE NC, on obtient des valeurs du temps qui évoluent moins avec la distance que celles de la base nationale, et un Sigma un peu plus faible. Le résultat au niveau agrégé a été nettement amélioré. Toutefois, la particularité de cette base (quelques O/D seulement déterminent l'intégralité des résultats) pousse à ne pas trop tirer de conclusions à partir des valeurs obtenues, l'important pour cette étude est qu'elle permet bien de confirmer les tendances mises en évidence par la base nationale, à savoir le fait que les valeurs du temps médianes sont supérieures à celles de l'instruction pour les longues distance (la base CETE NC n'est pas

pertinente pour les courtes distances), et que le paramètre sigma est plus élevé que celui de l'instruction.

Pour la base de préférences déclarées, on obtient également un Sigma élevé (proche de celui de la base nationale), et une valeur du temps médiane unique proche de 13 €/h. La relation avec la distance sur cette base de données n'a pu être établie, faute de suffisamment de disparités et également du fait de la faible taille de l'échantillon d'individus interrogés (248). La distance moyenne des déplacements étant de 140 km, cela correspondrait à une médiane de 20 €/h selon le modèle de la base nationale. En retenant cette valeur, et 2,66 pour sigma, la log-vraisemblance serait de -716, la reproduction agrégée étant par contre nettement moins satisfaisante, avec un coefficient de détermination de 0,58.

3.6 - Influence des vitesses retenues

Les valeurs du temps obtenues dans les paragraphes précédents correspondent à un choix particulier sur les vitesses de chaque arc pour la base nationale - voir description au 3.2. Outre le fait qu'il n'y ait pas une vitesse unique par type d'arc sans vraiment savoir si la détermination de la valeur s'est faite à dire d'expert ou en s'appuyant sur des données, les médianes de ces valeurs contrastent avec les vitesses à vide utilisées dans les modèles d'affectation du trafic – cf. tableau 13. Notamment, on note qu'il faut des niveaux de trafic très importants pour atteindre les vitesses médianes de la base nationale sur les bidirectionnelles. Par contre, le rapport entre les vitesses de chaque type et la vitesse des autoroutes concédées est relativement identique (plus faible de 4 à 9 points sur la base nationale).

Type de voie	Vitesse médiane (base nationale)		Vitesse à vide Sétra		Trafic nécessaire pour que la vitesse en charge Sétra soit égale à la vitesse médiane en véh/jour
	valeur	% de la vitesse sur autoroute concédée	valeur	% de la vitesse sur autoroute concédée	
autoroute concédée	115		131		40 000
autoroute non concédée	110	96%	131	100%	46 000
route express	90	78%	110	84%	54 000
bidirectionnelle (7m)	70	61%	91	69%	32 000
traversée d'agglomération	30	26%	46	35%	16 000
* calculs de vitesses en charge effectués sur la base de 2x2 voies pour les routes à chaussées séparées, avec 10 % de PL					

Tableau 13 : Comparaison des vitesses de la base nationale avec les recommandations SETRA

Afin d'examiner dans quelle mesure les éventuelles erreurs sur la description des vitesses pourraient avoir un impact sur le calibrage du modèle, deux modèles fondés sur les vitesses à vide du Sétra et sur une description mixte des vitesses (en retenant les vitesses à vide sur autoroute et les vitesses médianes de la base nationale sur bidirectionnelles et traversées d'agglomérations) ont été testés – cf. tableau 14. Le deuxième modèle a notamment pour objectif d'examiner si la prise en compte du malus dans le temps de parcours, au travers de la vitesse (avec des vitesses élevées sur autoroutes et plutôt faibles sur bidirectionnelles et traversées urbaines), est plus pertinente que celle fondée sur un malus kilométrique intégré dans le calcul du coût.

			Calage VDT Vitesses à vide	Calage VDT Vitesses mixtes
BASE NATIONALE	Valeur du temps	Médiane	d < 20 km : 0,7 d < 50 km : -8,5 + 0,4619 x d d < 400 km : -72,0 + 1,731 x d d > 400 km : 620,3	d < 20 km : 0,7 d < 50 km : -2,5 + 0,157 x d d < 400 km : -36,3 + 0,833 x d d > 400 km : 296,9
		σ	2,91	2,88
	Malus / vitesse	autoroute concédée	0 / 131	0 / 131
		autoroutes non concédées	0 / 131	0,7 / 131
		route express	0 / 110	0,7 / 110
		bidirectionnelle (7m)	2,4 / 91	5,4 / 70
		traversée d'agglomération	13,5 / 46	5,4 / 30
	Résultats	Log0	-149617	-149617
		LogV	-82673	-90010
		Coefficient de régression	0,9795	0,9829
		Coefficient de détermination	0,8647	0,833

Tableau 14 : Influence de la vitesse sur les valeurs du temps

L'utilisation de vitesses mixte donne des résultats nettement moins satisfaisants que le modèle de base (cf. tableau 12) : la log vraisemblance est plus élevée et le coefficient de détermination nettement moins bon. Dans ce cas également, l'introduction de malus particulier par type de route ne semble pas être une solution efficace.

A contrario, le modèle fondé sur les vitesses à vide du SETRA est d'aussi bonne qualité, voire légèrement meilleur, que le modèle initial. Par contre, les vitesses plus élevées ont un effet conséquent sur les valeurs du temps :

- le paramètre sigma augmente encore (2,91 contre 2,66),
- les valeurs extrêmes des médianes de valeur du temps sont encore respectivement plus faibles (0,7 contre 1,25) et plus élevées (620 contre 319).

Le rapport entre les vitesses pratiquées sur les différents types de route semblent donc cohérent, menant à un modèle plus satisfaisant que dans le cas où on supposerait des écarts plus importants entre bidirectionnelles et routes à chaussées séparées. Par contre, les choix effectués sur ces vitesses conditionnent la valeur du temps obtenue. Il semble préférable de conserver les vitesses de la base nationale, qui semblent refléter plus correctement les vitesses moyennes pratiquées que les vitesses à vide.

Un autre enseignement important peut être tiré de cette comparaison : un modèle générique fondé sur des vitesses générales permet une aussi bonne reproduction des résultats qu'un modèle basé sur des vitesses particulières.

3.7 - Vérification des valeurs de malus initialement retenues

Au paragraphe 3.3, les valeurs de malus suivantes ont été retenues :

Malus	autoroute concédée	0
	autoroutes non concédées	0
	route express	0
	bidirectionnelle (7m)	2,7
	traversée d'agglomération	10,7

Tableau 15 : Rappel des malus retenus au paragraphe 3.3.

Comme évoqué au paragraphe 3.2, puisque valeurs du temps et malus sont étroitement liés et que le malus a en quelque sorte été retenu avant de choisir les valeurs du temps, il est important de vérifier que les nouvelles valeurs du temps (valeurs du tableau 12) n'impliquent pas une importante variation des malus.

Pour ce faire, on travaille à nouveau sur les bases à deux itinéraires, et on cherche à optimiser les malus (valeurs en italique) compte tenu du nouveau jeu de valeurs du temps. Les résultats sont les suivants :

	Valeur du temps	Médiane	$d < 20 \text{ km} : 1.25$ $d < 50 \text{ km} : -1.89 + 0,157 \times d$ $d < 400 \text{ km} : -38.8 + 0,895 \times d$ $d > 400 \text{ km} : 319.2$
		σ	2,66
BASE NATIONALE	Malus	autoroute concédée	0
		autoroutes non concédées	0
		route express	0
		bidirectionnelle (7m)	3,3
		traversée d'agglomération	12,8
BASE CETE NC	Malus	autoroute concédée	0
		autoroutes non concédées	0
		route express	0
		bidirectionnelle (7m)	3,7
		traversée d'agglomération	14,1

Tableau 16 : Résultat du calibrage des malus à partir des nouvelles valeurs du temps.

Les valeurs de malus trouvées sont proches de celles qui ont été retenues au 3.3, puisqu'elles se situent autour de 3 c€ / km pour les bidirectionnelles, et supérieures à 10 c€ / km pour les traversées d'agglomération. Le fait que les valeurs ne soient pas identiques n'est pas fondamentalement problématique puisque comme on l'a vu, l'optimisation des malus reste secondaire par rapport à l'optimisation des valeurs du temps. L'important est que les valeurs trouvées soient du même ordre de grandeur et surtout que les tendances par rapport à l'instruction (environ 2 fois inférieur pour les bidirectionnelles, environ 2 fois supérieur pour les traversées d'agglomération) soient bien confirmées, ce qui est le cas ici.

Il faudra donc simplement retenir que le malus pour les bidirectionnelles se situe entre 2,5 et 3,5 c€ / km, et que celui des traversées d'agglomération se situe entre 10 et 13 c€ / km (les résultats pour la base CETE NC ne peuvent être considérés comme fiables pour une préconisation générale).

Cette vérification permet en quelle que sorte de valider l'ensemble du processus de calibrage des valeurs du temps puisque celui-ci reposait complètement sur ces valeurs initiales de malus retenues.

3.8 - De l'utilisation de valeurs de référence sur un cas particulier

Le calibrage sur la base nationale (paragraphe 3.5) a permis d'obtenir des malus et des valeurs du temps qu'on pourrait être tenté de choisir comme référence. On cherche maintenant à connaître l'efficacité de ces valeurs sur une base particulière différente de celle qui a servi au calage, en l'occurrence ici la base CETE NC. Cela permet de comparer l'efficacité d'un calage spécifique par rapport à l'utilisation de valeurs « par défaut ». Le tableau suivant compare ainsi, sur la base CETE NC, les résultats obtenus par application des valeurs de références obtenues sur la base nationale et ceux obtenus après un calibrage spécifique de ces valeurs.

			Résultat avec valeurs du temps obtenues sur la base nationale	Résultat avec calibrage spécifique
BASE CETE NC	Valeur du temps	Médiane	d < 20 km : 1,25 d < 50 km : $-1,89 + 0,157 \times d$ d < 400 km : $-38,8 + 0,895 \times d$ d > 400 km : 319,2	d < 20 km : 32,5 d < 50 km : $1,62 \times d$ d < 400 km : 81,2 d > 400 km : 81,2
		σ	2,66	1,62
	Malus	autoroute concédée	0	0
		autoroutes non concédées	0	0
		route express	0	0
		bidirectionnelle (7m)	2,7	2,7
		traversée d'agglomération	10,7	10,7
	Résultats	Log0	-50932	-50932
		LogV	-21023	-18679
		Coefficient de régression	0,79	0,88
		Coefficient de détermination	0,88	0,95

Tableau 17 : Efficacité des valeurs de référence par rapport à un calibrage spécifique

Les résultats obtenus à l'aide des paramètres par défaut sont relativement bons, ce qui confirme leur efficacité et leur pertinence : le fait d'avoir un écart type important explique beaucoup ce résultat, car il permet la prise en compte d'une importante hétérogénéité dans la population, ce qui rend la distribution retenue extrêmement « prudente », et applicable sur des secteurs d'étude assez variés.

Néanmoins, on observe que le calibrage spécifique permet d'améliorer encore le modèle de manière encore significative, au niveau désagrégré comme au niveau agrégé. L'utilisation de valeurs de référence ne remplace donc pas le calibrage spécifique à le secteur d'étude. On retiendra simplement que lorsque les délais du modélisateur sont très serrés, ces valeurs de référence peuvent servir de valeurs par défaut et peuvent être conservées en première approche si le modèle est correctement calé.

Chapitre 4 : Analyse des résultats des modèles stochastiques

Pour les bases de données nationale et CETE NC, plusieurs modèles stochastiques ont été testés :

- un modèle logit simple avec les malus retenus précédemment.

Dans ce modèle logit simple, l'utilité de chaque itinéraire i (pour une O/D donnée) s'exprime comme une combinaison linéaire du temps et du coût :

$$U_{O/D}(i) = a \times Temps(i)_{O/D} + b \times Coût(i)_{O/D}$$

La probabilité de choisir l'itinéraire i parmi l'ensemble des itinéraires possibles de l'O/D s'écrit sous la forme :

$$P_{O/D}(i) = \frac{e^{U_{O/D}(i)}}{\sum_j e^{U_{O/D}(j)}}$$

- un modèle de type Abraham (logit logarithmique) avec les malus retenus précédemment

Dans ce modèle logit logarithmique, l'utilité de chaque itinéraire i (pour une O/D donnée) s'exprime de la manière suivante :

$$U_{O/D}(i) = a \times \ln(b \times Temps(i)_{O/D} + Coût(i)_{O/D})$$

La probabilité de choisir l'itinéraire i parmi l'ensemble des itinéraires possibles de l'O/D s'écrit de la même manière que pour le logit simple (la forme de l'utilité est différente, mais la forme du modèle est la même).

- un modèle de type Abraham (logit logarithmique) avec calibrage des malus

Les résultats sont les suivants :

		Modèle logit avec malus 0c€/2,7c€/10,7c€	Modèle type Abraham avec malus 0c€/2,7c€/10,7c€	Modèle type Abraham avec optimisation du malus
BASE NATIONALE	Fonction d'utilité	$U = a \times \text{Tps} + b \times \text{Cout}$	$U = a \times \ln (b \times \text{Tps} + \text{Cout})$	$U = a \times \ln (b \times \text{Tps} + \text{Cout})$
	Malus	route express = 0 c€/km bidirectionnelle = 2,7c€/km traversée d'agglomération = 10,7 c€/km	route express = 0 c€/km bidirectionnelle = 2,7c€/km traversée d'agglomération = 10,7 c€/km	route express = 0 c€/km bidirectionnelle = 6 c€/km traversée d'agglomération = 22,6 c€/km
	Paramètres obtenus pour la fonction d'utilité	a = 0 b = 0,055	a = - 10,5 b = 10,1	a = - 11,2 b = 4,82
	Log0	-149880	-149880	-149880
	LogV	-148180	-129130	-127000
BASE CETE NC	Fonction d'utilité	$U = a \times \text{Tps} + b \times \text{Cout}$	$U = a \times \ln (b \times \text{Tps} + \text{Cout})$	$U = a \times \ln (b \times \text{Tps} + \text{Cout})$
	Malus	route express = 0 c€/km bidirectionnelle = 2,7c€/km traversée d'agglomération = 10,7 c€/km	route express = 0 c€/km bidirectionnelle = 2,7c€/km traversée d'agglomération = 10,7 c€/km	route express = 0 c€/km bidirectionnelle = 0 c€/km traversée d'agglomération = 25,1 c€/km
	Paramètres obtenus pour la fonction d'utilité	a=- 0,0888 b = - 0,0784	a = - 16,8 b = 55,02	a = - 15,7 b = 48,6
	Log0	-51028	-51028	-51028
	LogV	-24261	-22609	-22118

Tableau 18 : Résultats pour différents modèles stochastiques testés sur les bases nationale et CETE NC

Ces résultats montrent que le modèle logit logarithmique (type Abraham) est beaucoup plus efficace qu'un modèle logit simple pour reproduire les comportements de choix d'itinéraires des usagers. Pour les modèles de type Abraham, le paramètre « classique » de la loi d'Abraham est situé aux alentours de 10 ou 11 pour la base nationale, et de 16 ou 17 pour la base CETE NC. Les valeurs du temps trouvées pour la base nationale sont quant à elles bien inférieures à celles trouvées pour la base CETE NC. Au niveau des malus, les valeurs trouvées confirment les tendances observées avec les modèles déterministes : le malus pour les routes express est nul, celui des bidirectionnelles est stable voire en baisse par rapport à l'instruction, et celui des traversées d'agglomération est en forte hausse.

Sur ces deux bases, il n'a pas été possible de tester des modèles plus complexes. En effet, pour pouvoir calibrer un grand nombre de paramètres, la méthode utilisée nécessite des données dont la variabilité est très importante, ce qui n'est pas suffisamment le cas ici.

En revanche, sur la base des préférences déclarées, d'autres modèles ont pu être testés :

- dans le modèle 1, l'utilité de chaque itinéraire i (pour une O/D donnée) s'exprime comme une combinaison linéaire du temps, du coût et d'une constante d'alternative:

$$U_{O/D}(i) = a \times \text{Temps}(i)_{O/D} + b \times \text{Coût}(i)_{O/D} + c_i$$

- dans le modèle 2, la formulation est identique mais le paramètre a est une variable aléatoire distribuée selon une loi lognormale

- dans le modèle 3, la formulation est identique au modèle 1, mais les paramètres a et b sont distingués en fonction de la CSP de la personne
- dans le modèle 4, 5 sous-échantillons sont tirés aléatoirement. Un modèle de type 2 est calibré sur chacun de ces échantillons. La distribution des revenus de ces échantillons a pu être estimée et est présentée sur le tableau suivant :

	Échantillon 1	Échantillon 2	Échantillon 3	Échantillon 4	Échantillon 5
Moyenne des revenus	2 471 €	2 556 €	2 991 €	2 103 €	3 743 €
Ecart-type	0,57	0,59	0,64	0,55	0,7

Les résultats sont résumés dans le tableau suivant:

Résultat	Modèle 1 : Logit simple	Modèle 2 : Logit avec distribution log-normale de la vdt	Modèle 3 : Logit avec prise en compte des CSP des individus	Modèle 4 : 5 modèles de type 2, avec 5 groupes de vdt distinctes tirés aléatoirement				
Log Vraisemblance	-911,6	-906,5	-915,0	-878,2				
				-165,7	-158,6	-170,7	-191,1	-192,1
R ²	0,106	0,109	0,096	0,123	0,124	0,151	0,083	0,05
ASC1	3,79	5,32	3,62	-3,13	8,82	5,53	-4,59	-5,48
ASC2		0,785		1,49	0,858	3,38	-0,202	0,937
ASC6	0,391	0,511		0,552	1,07	0,565	0,786	0,119
ASC7	0,517	0,684	0,481	1,09	0,964	0,386	1,92	-0,516
ASC9	-0,805	-0,684	-0,877	-0,869	-0,445	-0,883	-0,193	-0,850
Beta_cout	-0,129	-0,159		-0,198	-0,177	-0,455	-0,122	-0,105
Beta_temps	-0,034 soit 15,8 €/h	Distribution log-normale moyenne : 43,2 €/h médiane : 5,1 €/h sigma : 2,07		22,4 9,6 1,3	37,9 9,0 1,69	161943 0,8 4,92	1450 4,4 3,4	11,8 2,5 1,76
Beta_cout CSP faibles revenus			-0,12					
Beta_cout CSP revenus intermédiaires			-0,2					
Beta_cout CSP revenus élevés			-0,148					
Beta_cout CSP retraités			-0,122					
Beta_cout CSP autres			-0,093					
Beta_temps CSP faibles revenus			-0,0309					
Beta_temps CSP revenus intermédiaires			-0,0534					
Beta_temps CSP revenus élevés			-0,0470					
Beta_temps CSP retraités			-0,0156					
Beta_temps CSP autres			-0,00581					

Tableau 19 : Résultats des modèles stochastiques testés sur la base de préférences déclarées

Il faut tout d'abord noter que les modèles 3 et 4 n'ont pu être rendus complètement « propres » :

- dans le modèle 3 subsistent de nombreuses corrélations entre les paramètres de coût et de temps ; le paramètre de temps de la classe « autres CSP » n'a pu être établi à un seuil de risque satisfaisant ;
- le modèle 4 doit être pris comme une illustration de la pertinence de séparer les distributions log-normales par populations ayant des distributions de revenus différentes : une recherche complète de mise

au point d'un modèle propre n'aurait aucun sens du fait de la faible taille de l'échantillon et du caractère artificiel des populations tirées aléatoirement.

L'écart de log-vraisemblance entre le modèle 1 et le modèle 2 semble mettre en évidence que, pour cet échantillon de données, la prise en compte d'une distribution log-normale des valeurs du temps n'apporte qu'une légère amélioration. Par contre, les indicateurs de qualité du modèle 3 montrent qu'il n'est pas possible d'améliorer la qualité du modèle en utilisant les caractéristiques individuelles sans diminuer la qualité de l'estimation statistique. Enfin, pour le modèle 4, on constate que l'approche peut être pertinente (forte réduction de la log-vraisemblance), mais nécessite une forte quantité de données : les distributions des valeurs du temps des échantillons 3 et 4 paraissent incohérentes : écart-type très élevé ; moyenne et médiane non en relation avec la distribution des revenus.

Pour l'échantillon recueilli, la mise en place d'un modèle tenant compte de la distribution des préférences individuelles fonctionne donc de manière nettement plus satisfaisante que celle d'un modèle basé sur les caractéristiques individuelles.

Chapitre 5 : Synthèse des résultats et recommandations

5.1 - Déterministe vs Stochastique

On rappelle les principaux résultats obtenus aux paragraphes précédents. Pour chaque type de modèle, on ne retient que le meilleur d'entre eux du point de vue de l'estimation du maximum de vraisemblance. Les résultats sont analysés au regard du maximum de vraisemblance pour les bases nationales et CETE NC, et au niveau de la vérification agrégée pour la base de préférences déclarées (car dans le cas de cette base, le maximum de vraisemblance n'est pas comparable entre les modèles stochastiques et déterministes à cause des nombreux itinéraires inefficaces).

BASE NATIONALE	Initial	Log0 = -149 617
	Modèle déterministe	LogV = -88 281
	Modèle stochastique	LogV = -127 000
BASE CETE NC	Initial	Log0 = -50 932
	Modèle déterministe	LogV = -18 679
	Modèle stochastique	LogV = -22 118
BASE DE PREFERENCES DECLAREES	Modèle déterministe	Coefficient de la régression = 1,007 Coefficient de détermination = 0,7326
	Modèle stochastique	Coefficient de la régression = 0,9473 Coefficient de détermination = 0,8128

Tableau 20 : Comparaison des modèles stochastiques et déterministes

Le modèle déterministe semble nettement plus adapté que le modèle stochastique pour reproduire les comportements de la base nationale et de la base CETE NC. Sur la base nationale, la « faiblesse » du modèle stochastique implique de se poser la question sur la qualité des paramètres estimés. Il faut toutefois noter que ce résultat en faveur du modèle déterministe est en grande partie atteint grâce au processus de sélection des chemins efficaces.

Pour la base de préférence déclarée, la conclusion est inversée, à savoir que le modèle stochastique permet d'améliorer le modèle déterministe de façon sensible (coefficient de détermination bien meilleur). Cela est lié au fait que nombre de chemins non efficaces ont été choisis dans les faits par les usagers. Dans ce cas, le modèle stochastique est supérieur au modèle déterministe du fait de sa construction théorique.

On ne peut donc pas conclure de manière définitive à la supériorité d'un modèle par rapport à l'autre. De même, l'analyse d'un modèle mixte, semblant souhaitable, n'a pu être menée à terme, l'optimisation d'un modèle stochastique avec distribution log-normale des valeurs du temps ne s'étant pas avérée pertinente sur nos bases de données.

La principale différence entre ces deux modèles est donc liée à l'existence ou non d'itinéraires inefficaces empruntés dans les faits par les usagers, notamment lorsque la différence de coût est nulle et la différence en temps faible. Par contre, la mise en œuvre d'un modèle stochastique sur de grandes échelles implique d'effectuer au préalable un tri des itinéraires non efficaces, ce qui peut s'avérer extrêmement chronophage (un des principaux défaut d'Ariane).

5.2 - Synthèse sur les malus et les valeurs du temps

Les résultats ci-dessus peuvent être synthétisés de la façon suivante :

Malus :

Type de voie	Instruction	Résultats de l'étude
Réseau non concédé à chaussées séparées	0,7 c€ / km	0 c€ / km
réseau routier interurbain	5,4 c€ / km	entre 2,5 et 3,5 c€ / km
traversées d'agglomération	5,4 c€ / km	entre 10 et 13 c€ / km

Tableau 21 : Valeurs obtenues pour les malus

Valeurs du temps par classe de distance :

Valeur du temps	Instruction	Résultats de l'étude
Médiane	$d < 20 \text{ km} : 8,7$ $d < 50 \text{ km} : 11,8$ $d < 400 \text{ km} : 13,6 + 0,027 \times d$ $d > 400 \text{ km} : 30,3$	$d < 20 \text{ km} : 1,25$ $d < 50 \text{ km} : -1,89 + 0,157 \times d$ $d < 400 \text{ km} : -38,8 + 0,895 \times d$ $d > 400 \text{ km} : 319,2$
Sigma	0,5	2,66

Tableau 22 : Valeurs du temps obtenues par classe de distance, pour un modèle déterministe prix/temps

Valeurs du temps par motif et par classe de distance :

Valeur du temps	Loisirs	Affaires personnelles	Travail	Affaires professionnelles
Médiane	$d < 20 \text{ km} : 0$ $d < 50 \text{ km} : -1,8 + 0,09 \times d$ $d < 400 \text{ km} : -46,8 + 0,99 \times d$ $d > 400 \text{ km} : 348,9$	$d < 20 \text{ km} : 0,3$ $d < 50 \text{ km} : -1,2 + 0,07 \times d$ $d < 400 \text{ km} : -17,5 + 0,4 \times d$ $d > 400 \text{ km} : 142,7$	$d < 20 \text{ km} : 1,3$ $d < 50 \text{ km} : -1 + 0,11 \times d$ $d < 400 \text{ km} : -31,2 + 0,72 \times d$ $d > 400 \text{ km} : 256$	$d < 20 \text{ km} : 1,1$ $d < 50 \text{ km} : -11 + 0,61 \times d$ $d < 400 \text{ km} : -4,8 + 0,48 \times d$ $d > 400 \text{ km} : 188,3$
Sigma	$d < 20 \text{ km} : 2,37$ $d < 50 \text{ km} : 2,37$ $d < 400 \text{ km} : 3,30$ $d > 400 \text{ km} : 4,14$	$d < 20 \text{ km} : 3,17$ $d < 50 \text{ km} : 3,17$ $d < 400 \text{ km} : 2,71$ $d > 400 \text{ km} : 2,5$	$d < 20 \text{ km} : 2,54$ $d < 50 \text{ km} : 2,54$ $d < 400 \text{ km} : 2,79$ $d > 400 \text{ km} : 1,94$	$d < 20 \text{ km} : 3$ $d < 50 \text{ km} : 3$ $d < 400 \text{ km} : 1,77$ $d > 400 \text{ km} : 1,59$

Tableau 23 : Valeurs du temps obtenues par motif et par classe de distance, pour un modèle déterministe prix/temps

Paramètre de la loi d'Abraham :

Le paramètre de la loi d'Abraham ne peut quant à lui pas vraiment être tiré des résultats obtenus sur la base nationale. Par ailleurs, on ne peut pas retenir une valeur « France entière » à partir des résultats (valeurs aux alentours de 16) de la seule base de données locale du CETE NC. On peut simplement retenir que la valeur du paramètre de la loi d'Abraham semble sur ces deux bases de données supérieure au 10 qui était utilisé dans Ariane.

	Instruction	résultats de l'étude
Paramètre de la loi d'Abraham	10	entre 11 à 16

Tableau 24 : Paramètre de la loi d'Abraham proposé

5.3 - Éléments de recommandations pour l'utilisation de modèles de choix d'itinéraire et l'introduction de valeurs du temps

L'analyse de données d'enquêtes Origine/Destination fournit un certain éclairage sur les arbitrages à effectuer lors de la constitution d'un modèle de choix d'itinéraire :

- une connaissance précise de l'offre de transport est incontournable. L'utilisation des paramètres de description des itinéraires recommandée (principalement vitesses à vide) actuellement par le Sétra permet de construire un modèle reproduisant les observations de manière satisfaisante, mais au prix de médianes de valeurs du temps très faibles pour les courtes distances (0,7 €/h) et très fortes pour les longues distances (jusqu'à 600 €/h). Avant tout approfondissement complémentaire de ces analyses sur le choix d'itinéraire, un travail sur les niveaux de service semble impératif.
- la priorité absolue est à donner au calage de la valeur du temps par rapport au calage du malus lors de la constitution du modèle. Ce résultat s'appuie d'une part sur la meilleure amélioration du modèle apportée par un calage des valeurs du temps et d'autre part sur la faible sensibilité de la distribution de la valeur du temps à l'introduction de malus.
- la décomposition par motifs de la demande n'apporte qu'une amélioration limitée de la reproduction des choix d'itinéraires agrégés par rapport à une distribution des valeurs du temps variable en fonction de la distance. Ceci est lié à la forte corrélation de la part des différents motifs avec la distance du déplacement.
- des principes généraux sur la forme de la distribution de la valeur du temps ressortent clairement vis à vis des valeurs actuellement retenues : un paramètre sigma nettement plus élevé (toujours supérieur à 1,5) d'une part, et des médianes très faibles voire nulles pour les distances courtes et très élevées pour les distances longues (jusqu'à plusieurs centaines d'euros/heure), avec un comportement linéaire entre les deux. Les courbes de ces deux distributions extrêmes sont tracées sur l'illustration 15, en comparaison des courbes de l'instruction.

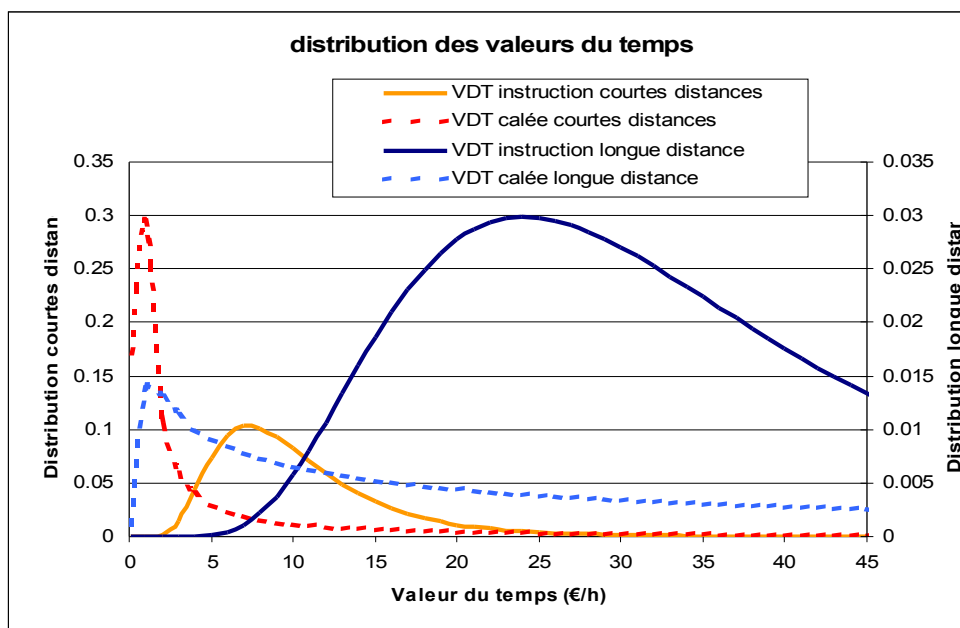


Illustration 15 : Distributions des valeurs du temps révélées et de l'instruction

- s'interroger sur la forme de la fonction de coût en cas de difficulté de reproduction des itinéraires peut s'avérer pertinent. Dans notre cas, cela aboutissait à ne retenir que le péage comme valeur réellement objectivement prise en compte par les usagers. Les données étant néanmoins relativement anciennes (1995 – 2000) et le contexte macro-économique ayant évolué depuis (prix du carburant notamment), cette conclusion ne saurait être généralisée et une analyse sur des données récentes, incluant une comparaison avec les résultats présentés ici, serait nécessaire pour approfondir le sujet ;
- la loi d'Abraham reste la formulation d'un modèle stochastique la plus pertinente ;
- le calibrage du modèle de choix d'itinéraire spécifique à un secteur est préférable à l'utilisation de valeurs génériques. Des outils simples (tableurs excel) ont été développés à cet effet, et pourront être utilisés afin de construire des modèles avec des paramètres fiables spécifiques à une zone d'étude, et non plus « à dire d'expert » comme cela peut être le cas dans un certain nombre d'études de trafic. A défaut, les valeurs du temps proposées au paragraphe 5.2 peuvent être utilisées (pour un modèle déterministe prix-temps) car elles sont assez « prudentes » du fait des Sigmas importants qui ont été retenus (prise en compte de comportements très hétérogènes).

5.4 - Limites de l'exercice et suites à donner

Le concept de valeur du temps retenu est celui d'un paramètre permettant de reproduire les choix d'itinéraire effectivement constatés, que ce soit par le biais d'enquêtes de préférences révélées ou déclarées. Il convient de rappeler ici que les distributions de valeurs du temps obtenues au 5.2 ont un sens dans un contexte spécifique, à savoir :

- le surcoût maximal de l'itinéraire le plus cher est de 45€ en valeur absolue et 14 c€/km ;
- les valeurs du temps de coupure se situent majoritairement dans une plage de 0 à 20 €/h.

La distribution des valeurs du temps est donc à interpréter par rapport à cette valeur du temps de coupure maximale : elle permet uniquement d'estimer la proportion d'usagers dont la valeur du temps est supérieure à ce seuil.

Les différentes bases de données utilisées dans cette étude présentent toutes des limites importantes :

- le problème principal concernant la base nationale est son ancienneté (1995 – 2000). Il est également regrettable qu'il s'agisse uniquement d'une sélection d'O/D très particulières (ville à ville).

- la base « CETE NC » présentait l'intérêt d'avoir une approche complète (utilisation de l'ensemble des résultats d'une enquête). Cependant, elle ne fournit des résultats que sur un secteur restreint. En outre, la sélection automatique des itinéraires imposée par la quantité d'O/D à traiter crée d'éventuels problèmes de cohérence interne qui n'ont été que partiellement résolus et examinés.
- la base « préférences déclarées » est de taille insuffisante (250 individus) pour permettre de tester des modèles complexes, notamment la sensibilité à la distance du déplacement.

Au final, les résultats s'appuient donc principalement sur la base nationale, qui semble présenter le moins de défauts. L'utilisation de valeurs du temps 2000 semblant néanmoins peu satisfaisante à terme, il conviendra donc de reprendre ces résultats avec une base équivalente constituée sur les enquêtes de 2005 à 2010, ce qui permettra d'analyser l'évolution éventuelle des comportements des usagers.

Avant de procéder à cet exercice, il sera cependant nécessaire de stabiliser la description des niveaux de service sur différents types de routes, qui conditionnent fortement le calibrage. Une fois ces résultats obtenus, deux options nous sont ouvertes pour reprendre ce travail avec des données récentes :

- mettre au point des outils de sélection automatique des itinéraires pour supprimer l'incohérence interne détectée. Dans ce cas, il serait possible de travailler avec une base très conséquente constituée de l'intégralité des résultats de toutes les enquêtes O/D. L'analyse s'en trouverait considérablement enrichie. L'effort à mener serait donc sur un plan méthodologique. Il pourrait s'inspirer de travaux universitaires récents [15].
- reconstituer une base équivalente à la base nationale, avec un travail manuel de description des itinéraires « garantissant » une base de données relativement propre, mais de taille plus réduite. L'intérêt serait également de pouvoir constituer une base contenant les mêmes O/D et ainsi avoir une comparaison directe de l'évolution des comportements.

Enfin la présente étude n'a abordé ni la question du choix d'itinéraire des PL, ni la question du lien entre la valeur du temps et les revenus. Un examen ultérieur des choix d'itinéraire devra nécessairement les traiter.

Bibliographie

- [1] Abraham C. , Coquand R. ; La répartition du trafic entre itinéraires concurrents. Revue Générale des Routes et Aéroports. N°357, pp57-76. Paris, juin 1961.
- [2] L'instruction DGR relative aux méthodes d'évaluation économique des investissements routiers interurbains (*version provisoire du 23 mai 2007*)
- [3] Leurent F. ; The theory and practice of a dual criteria assignment model with a continuously distributed value of time ; International symposium on transportation and traffic theory No13, Lyon (24/07/1996) pp. 455-477.
- [4] Leurent F. ; Calcul de deux modèles d'affectation. Convention Sétra-Inrets. 1998.
- [5] Sétra, Les outils d'évaluation des projets routiers – D'Ariane à TransCAD. Février 2010
- [6] Pfeffer M. ; Toll Demand Forecasting – a Dual-Criteria Assignment Model. Cranfield Center for Logistics & Transportation. MSc Thesis, Juillet 2001.
- [7] Commissariat Général du Plan. Transports : choix des investissements et coût des nuisances – dit "Rapport Boiteux II ". juin 2001
- [8] Lafoucade M. Thisse J.F ; New Economic Geography : a guide to Transport Analysis, Paris School of Economics, 2008
- [9] Duprez F. , Synthèse de la connaissance sur les interactions transports – économie, MEDAD, 2007
- [10] Offner, Les effets structurants du transport : mythe politique, mystification scientifique, ENPC
- [11] Bazin S. et al , L'arrivée de la LGV en Champagne-Ardenne et la nécessaire réorganisation des rapports de proximité, Cahiers Scientifiques du transport, 2006
- [12] Sétra, Enquêtes de circulation - Organisation et déroulement, guide technique, mars 2010
- [13] CETE de l'Est, Schéma National des Infrastructures de Transport – Études de trafic – A32 – modélisation des trafics, septembre 2008
- [14] Sétra, Guide Calage et validation des modèles de trafic - Techniques appliquées à l'affectation routière interurbaine , juillet 2010
- [15] Frejinger, Bierlaire et Ben Akiva, Sampling of alternatives for route choice modeling , Transportation Research Part B: Methodological, Volume 43, Issue 10, Pages 984-994 , December 2009

Index des tables

Tableau 1 : Répartition des longueurs de chaque type de route au sein des itinéraires de la base nationale.....	19
Tableau 2 : Répartition des motifs de déplacement au sein de la base nationale	20
Tableau 3 : O/D retenues dans le questionnaire	25
Tableau 4 : Utilisation du TML en fonction de la différence de coût et de temps avec l'itinéraire alternatif de coût généralisé minimum.....	27
Tableau 5 : Valeurs du temps et malus de l'instruction	29
Tableau 6 : Résultats obtenus sur les trois bases avec les valeurs du temps et malus de l'instruction	29
Tableau 7 : Comparaison de l'efficacité du calibrage des malus et des valeurs du temps de manière séparée.....	30
Tableau 8 : Distribution des vitesses sur la base nationale.....	31
Tableau 9 : Résultats du calibrage des malus sur la base nationale à deux itinéraires.....	32
Tableau 10 : Résultats du calibrage des malus sur la base nationale à deux itinéraires.....	33
Tableau 11 : Comparaison des résultats obtenus avec et sans différenciation selon le motifs.....	35
Tableau 12 : Résultats du calibrage des valeurs du temps	37
Tableau 13 : Comparaison des vitesses de la base nationale avec les recommandations SETRA.....	38
Tableau 14 : Influence de la vitesse sur les valeurs du temps.....	39
Tableau 15 : Rappel des malus retenus au paragraphe 3.3.	40
Tableau 16 : Résultat du calibrage des malus à partir des nouvelles valeurs du temps.....	40
Tableau 17 : Efficacité des valeurs de référence par rapport à un calibrage spécifique	41
Tableau 18 : Résultats pour différents modèles stochastiques testés sur les bases nationale et CETE NC	44
Tableau 19 : Résultats des modèles stochastiques testés sur la base de préférences déclarées	45
Tableau 20 : Comparaison des modèles stochastiques et déterministes	48
Tableau 21 : Valeurs obtenues pour les malus	49
Tableau 22 : Valeurs du temps obtenues par classe de distance, pour un modèle déterministe prix/temps.....	49
Tableau 23 : Valeurs du temps obtenues par motif et par classe de distance, pour un modèle déterministe prix/temps.....	49
Tableau 24 : Paramètre de la loi d'Abraham proposé.....	50

Index des illustrations

Illustration 1 : Probabilité de choix d'un itinéraire en fonction du rapport des coûts généralisés, pour la loi d'Abraham.....	6
Illustration 2 : Principe de fonctionnement de la loi d'affectation prix/temps.....	8
Illustration 3 : Forme de la loi log-normale pour plusieurs valeurs de médiane.....	10
Illustration 4 : Forme de la loi log-normale pour plusieurs valeurs de sigma.....	10
Illustration 5 : Fonction de répartition de la valeur du temps pour une médiane de 319,2 €/h et sigma=2,66.....	14
Illustration 6 : Exemple de résultat d'une comparaison sur les choix d'itinéraires agrégés.....	16
Illustration 7 : Répartition des distances des itinéraires de la base nationale	19
Illustration 8 : Parts des affaires professionnelles et du domicile-travail selon la distance au sein de la base nationale	20
Illustration 9 : Présentation des deux coupures retenues pour la base « CETE NC ».....	21
Illustration 10: Répartition des distances des itinéraires de la base CETE NC.....	22
Illustration 11 : Localisation des enquêtes de la base de préférence déclarées.....	24
Illustration 12 : Motifs du déplacement par O/D.....	26
Illustration 13 : Variation de la médiane et du paramètre sigma de la distribution de la valeur du temps en fonction de la distance	34
Illustration 14 : Evolution de la valeur du temps (médiane) selon le motif et la distance.....	35
Illustration 15 : Distributions des valeurs du temps révélées et de l'instruction.....	51

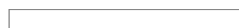
Pôle de Compétences et d'Innovation

"Méthodes, Outils et Démarches pour la Modélisation et l'Organisation des Déplacements"

Ce document a été élaboré sous le pilotage du Sétra par le PCI "Méthodes, Outils et Démarches pour la Modélisation et l'Organisation des Déplacements".

Le PCI vise à proposer des évolutions dans la manière de conduire les études de déplacements afin de pouvoir répondre aux attentes sociétales complexes relevant directement ou indirectement de ce domaine d'investigations. La modélisation des déplacements est donc la cible principale du PCI mais il a aussi vocation à proposer, spécifier ou améliorer d'autres démarches, méthodes et outils pour permettre une prise en compte plus large et plus efficace des enjeux de déplacements.

Le PCI est situé au CETE Normandie Centre .



Rédacteurs

Thomas ANSELME – CETE Normandie Centre

téléphone : 33 (0)2 35 68 88 78 – télécopie : 33 (0)2 35 68 82 19

mél : thomas.anselme@developpement-durable.gouv.fr

Mathieu JACQUOT – CETE de l'Est

téléphone : 33 (0)3 87 20 45 45 – télécopie : 33 (0)3 87 20 44 20

mél : mathieu.jacquot@developpement-durable.gouv.fr

Service d'études sur les transports, les routes et leurs aménagements
110, rue de Paris - SOURDUN – BP 124 – 77487 PROVINS Cedex – France
téléphone : 33 (0)1 60 52 31 31 – télécopie : 33 (0)1 60 52 31 69

Document consultable et téléchargeable sur les sites web du Sétra :

- Internet : <http://www.setra.developpement-durable.gouv.fr>
- Intranet (Réseau ministère) : <http://intra.setra.i2>

Ce document ne peut être vendu. La reproduction totale du document est libre de droits.
En cas de reproduction partielle, l'accord préalable du Sétra devra être demandé.

© 2012 Sétra – Référence : 1216W– ISRN : EQ-SETRA--12-ED11--FR

Le Sétra appartient
au Réseau Scientifique
et Technique
du MEDDE