

**AURAIT-ON PU PRÉVOIR L'ALLONGEMENT
DES DISTANCES DES DÉPLACEMENTS URBAINS
OBSERVÉ CES VINGT DERNIÈRES ANNÉES
AVEC LE MODÈLE DE DISTRIBUTION GRAVITAIRE ?**

JORGE E. CABRERA DELGADO, PATRICK BONNEL

LABORATOIRE D'ÉCONOMIE DES TRANSPORTS
UNIVERSITÉ DE LYON, ENTPE, CNRS

1. INTRODUCTION

Dans la pratique de la planification urbaine, le modèle gravitaire de distribution des déplacements continue de jouer un rôle important dans les exercices de prospective, à l'intérieur de la séquence classique de modélisation à quatre étapes (ORTÚZAR, WILLUMSEN, 2001 ; BONNEL, 2004). Celle-ci demeure l'outil principal dès lors que l'on cherche, par exemple, à prévoir le trafic d'une nouvelle infrastructure (dimensionnement ou évaluation) ou à simuler l'évolution des déplacements sur un territoire dans le cadre d'études de planification (PDU, DVA, SCOT. . .)¹ à moyen ou long terme, soit environ dix à vingt ans (TRB, 2007). Elle a également pris une place importante en

¹ PDU : Plan de Déplacements Urbains ; DVA : Dossier Voirie de l'Agglomération ; SCOT : Schéma de Cohérence Territoriale.

tant que partie constituante des efforts de modélisation de l'interaction transport urbanisme², caractérisés notamment par des ambitions de simulation à des horizons de plus long terme (au-delà de vingt ans).

Le modèle gravitaire, tout comme les autres composants de la modélisation à quatre étapes, a conservé une actualité pratique en grande partie en raison de sa compatibilité avec les données usuellement disponibles. S'agissant de modèles statiques, le calage fait appel à des données en coupe instantanée³ : données de la dernière enquête déplacements disponible, considérée comme étant la plus représentative des comportements des individus. La reconstitution des réseaux, étape particulièrement coûteuse, n'est conduite que pour la situation de référence qui correspond souvent à l'année de réalisation de la dernière enquête déplacements ou tout au moins à l'année pour laquelle le plus de données est disponible. Les exercices de prévision sont réalisés en supposant que les paramètres des différents modules du modèle à quatre étapes sont stables dans le temps.

Compte tenu des horizons de prévision, l'hypothèse de stabilité est loin d'aller de soi. D'une part, les erreurs de prévision restent monnaie courante. FLYVBJERG et alii (2005) mettent en évidence des sous-estimations de demande pour les projets routiers de 9,5 % en moyenne et de surestimations de 51,4 % en moyenne pour les projets ferroviaires. Parmi les causes des erreurs de prévision évoquées par les acteurs interrogés dans cette étude, les erreurs faites pendant la phase de distribution apparaissent en troisième position pour les projets routiers et en première position pour les projets ferroviaires⁴. D'autre part, l'analyse des données d'enquêtes déplacements met en évidence des évolutions dans les comportements de déplacement aux horizons de prévision usuels. Au niveau de la distribution des déplacements, on constate un allongement régulier des distances sans pouvoir établir avec précision s'il résulte exclusivement de l'accès plus fréquent aux modes motorisés et de l'accroissement des vitesses moyennes de déplacements, de la périphérisation des flux ou s'il n'y a pas en plus un besoin (souhaité ou subi selon les interprétations) des individus d'élargir leur aire de pratique des activités quotidiennes.

² Voir par exemple les modèles SIMAURIF (NGUYEN-LUONG, 2011) et SIMBAD (NICOLAS, 2011) en France ou l'utilisation de la plate-forme Urbansim combinée avec le modèle transport du Wasatch Front Regional Council, aux États-Unis (WADDELL et alii, 2007).

³ Dans l'idéal, une modélisation dynamique prenant appui sur des données en séries temporelles, serait la meilleure manière de prendre en compte l'impact de l'évolution des comportements dans les prévisions. Ainsi, par exemple, dans le transport de longue distance, l'accès à certaines séries chronologiques a permis de mettre en place des modèles dynamiques, permettant de prendre en compte des évolutions temporelles (CABANNE, 2005). La contrainte des données joue, en urbain, un rôle déterminant.

⁴ Il convient de noter que, dans l'étude de FLYVBJERG et alii, les causes de la mauvaise estimation de la demande dans les projets routiers sont d'ordre technique (différents aspects de la modélisation), alors que pour les projets ferroviaires des enjeux exogènes semblent l'emporter.

La littérature empirique relative aux « aspects temporels »⁵ du modèle gravitaire, dont la stabilité paramétrique, est relativement réduite au regard de celle consacrée aux autres modules de la séquence à quatre étapes (ORTÚZAR, WILLUMSEN, 2001 ; FOX, HESS, 2010). L'objectif de cet article est de mettre à l'épreuve la capacité du modèle de distribution gravitaire à simuler l'effet des évolutions du système urbain⁶ sur les distances de déplacement jouant dans les dix ou vingt ans que l'on cherche à simuler. On s'intéresse à deux motifs de déplacement : Travail et Loisirs, qui sont ceux pour lesquels l'accroissement des distances est le plus visible. Accessoirement, il s'agit également des motifs de déplacement considérés comme le plus et le moins contraints, respectivement. Le terrain d'application retenu est l'agglomération lyonnaise, choisi essentiellement pour des raisons de disponibilité de données. À partir des enquêtes ménages déplacements (CERTU, 2008) de 1985, 1995 et 2006 et de données de temps issues de réseaux codifiés pour ces années, on calibre des modèles gravitaires pour les trois dates. Les résultats des calages permettent d'étudier la capacité d'un modèle (1985) à reproduire les distributions des distances observées à une date ultérieure (1995 et 2006).

Un tel exercice nécessite que l'on se dote d'indicateurs permettant de comparer des simulations à plusieurs dates. Comme SMITH et HUTCHINSON (1981) le faisaient remarquer, les critères de calibration et les mesures de qualité de l'ajustement peuvent être classés en deux approches fondamentales : (i) la similitude de certaines macro-caractéristiques, en particulier la distribution de longueur des trajets, et (ii) la comparaison cellule par cellule des matrices observées et simulées à partir du modèle. Cette dernière a été privilégiée dans les quelques études existantes sur la validation temporelle du modèle gravitaire. Dans la mesure où l'on est directement intéressé par l'étude de l'allongement des distances de déplacement, on combine les deux types d'approches pour la construction d'indicateurs.

La présentation se structure de la manière suivante. La section 2 propose un panorama de l'état des connaissances relatives à la stabilité temporelle des paramètres du modèle gravitaire. La section 3 définit les principes méthodologiques de notre démarche (notamment l'unité d'étude de la mobilité, les spécifications des modèles à tester et la technique de calibration). La section 4 décrit les données utilisées. La section 5 analyse les principaux résultats obtenus, avant de conclure dans la section 6.

2. REVUE BIBLIOGRAPHIQUE

Une revue rapide de la littérature consacrée au modèle de distribution

⁵ Définis plus en détail dans ce qui suit (section 2).

⁶ Nous faisons référence ici aux travaux de BONNAFOUS et PUEL (1983) pour lesquels le système urbain est composé de trois sous-systèmes : localisation des activités, réseaux de transport et interactions sociales.

gravitaire révèle une abondance d'études portant sur des considérations d'ordre théorique, comme les fondements du modèle. On peut citer, entre autres, WILSON (1967) pour l'approche par maximisation de l'entropie, SNICKARS et WEIBULL (1977) pour l'approche par minimisation de la divergence d'information ou encore COCHRANE (1975) et ANAS (1983) pour la dérivation du modèle gravitaire à partir de la maximisation de l'utilité. Les contributions méthodologiques, comme les techniques de calibration (HYMAN, 1969 ; EVANS, 1971 ; BATTY, MACKIE, 1972 ; KIRBY, 1974 ; WILLIAMS, 1976) ou la forme de la fonction de résistance (EVANS, KIRBY, 1974), sont au moins aussi nombreuses. Ces recherches ont été conduites pendant les années 1970, le sujet étant moins d'actualité aujourd'hui.

À l'opposé, la littérature empirique sur le modèle gravitaire est beaucoup plus rare, particulièrement en ce qui concerne la validité de celui-ci à plusieurs dates différentes (pour plus de détails sur la validation, voir CAMBRIDGE SYSTEMATICS, INC, 2010 ou encore CERTU, 2002). Cette question est désignée par le terme *transférabilité* temporelle. La transférabilité d'un modèle peut être définie de la manière suivante (KOPPELMAN, WILMOT, 1982) : le transfert est l'application d'un modèle, information ou théorie sur le comportement, développé pour un contexte particulier, dans l'objectif de décrire le comportement correspondant dans un autre contexte. La transférabilité est la valeur, l'utilité du modèle, information ou théorie transféré dans le nouveau contexte. On distingue généralement deux types de transférabilité : temporelle, qui correspond au cas dans lequel on s'intéresse à un territoire à des moments différents dans le temps, et spatiale, qui correspond au cas où l'on s'intéresse à plusieurs territoires au même moment dans le temps. KARASMAA (2008) distingue deux types d'approches de la transférabilité. La première étudie l'hypothèse d'égalité des paramètres par des tests statistiques (biais de transfert). KARASMAA note que cette approche a l'inconvénient d'être trop restrictive dans la mesure où aucun modèle n'est parfaitement spécifié et qu'il y a toujours une certaine dépendance au contexte (l'hypothèse de stabilité serait quasi-systématiquement rejetée). La deuxième approche consiste à reconnaître dès le départ qu'il y a eu des changements et à étudier les moyens par lesquels les écarts et les biais qui en résultent peuvent être pris en compte sans avoir à ré-estimer un nouveau modèle en entier ni recueillir une grande quantité de données nouvelles. En matière de prévision, la non-corroboration de l'hypothèse de transférabilité temporelle du modèle suppose une incertitude accrue sur la qualité des prévisions. On indique, dans ce qui suit, les principales contributions portant sur la validation temporelle des modèles gravitaires.

OPENSHAW est la référence incontournable en ce qui concerne les études empiriques en rapport avec le modèle gravitaire. En 1976, il proposa une première étude sur la capacité de neuf variantes du modèle gravitaire à reproduire les interactions spatiales observées. Il conclut que les modèles fondés sur la

théorie, comme ceux ayant une origine empirique, ont du mal à atteindre une performance descriptive satisfaisante. Plus tard, en 1979, il s'intéressa à la performance du modèle gravitaire en prévision en étudiant la région de Durham, au nord de l'Angleterre, pour laquelle des temps-réseaux précis sont disponibles pour deux dates, 1966 et 1971. Si la performance descriptive du modèle est moyenne, celle en prévision est pratiquement la même que celle obtenue lors de la calibration, à condition que des estimations correctes des variables d'entrée soient disponibles. Les erreurs liées à l'hypothèse de stabilité des paramètres de calage sont relativement moins importantes que celles liées aux erreurs de prévision des variables exogènes. L'amélioration de la performance descriptive est possible, mais au détriment de la qualité des prévisions, en raison d'une sensibilité accrue aux changements de paramètres. SOUTHWORTH (1979) étudie l'impact de la structure spatiale (segmentation de la population par niveau de revenu et lieu de résidence) sur les paramètres du modèle gravitaire en calibrant plusieurs modèles contraints à la destination, en faisant une distinction par motif de déplacement, à partir de données de la *London traffic survey* de 1962. Il observe une grande variabilité des paramètres, et une amélioration de la qualité des calages apportée par la segmentation. Dans une deuxième étude (SOUTHWORTH, 1983), il poursuit son travail sur les modèles segmentés en introduisant la composante temporelle, à partir de données du Grand Londres pour 1962 et 1971. Il cherche à comparer l'impact des changements dans les temps de parcours et ceux de désagréments non-temporelles sur les paramètres du modèle gravitaire. Ses résultats confirment ce qui avait déjà été avancé par OPENSHAW : la segmentation n'améliore pas la qualité des prévisions. Les prévisions sont plus sensibles aux écarts de temps de parcours qu'aux changements paramétriques. Compte tenu de ces résultats, tant OPENSHAW que SOUTHWORTH recommandent de limiter l'usage du modèle gravitaire à des études descriptives et d'éviter son utilisation en prospective.

VOLET et HUTCHINSON (1986) ont testé le pouvoir descriptif d'un modèle gravitaire comparé à celui d'un modèle de type « facteurs de croissance » (Fratat) sur les flux interzonaux domicile-travail de la ville de Toronto, pour 1971 et 1981. Ils ont également testé la capacité du modèle gravitaire calibré en 1971 à prévoir les flux observés en 1981. Leurs résultats suggèrent que le modèle très simple de type « facteurs de croissance » donne de meilleurs résultats en termes de prévision que le modèle gravitaire. Plus généralement, les deux types de modèles ont des difficultés à imiter les modifications dans la structure des trajets domicile-travail qui sont dus à des changements des modèles de vie et de travail des différentes catégories de la population.

DUFFUS et alii (1987) ont étudié la performance prospective d'un modèle gravitaire à fonction de résistance « par paliers » et l'impact de l'inclusion de K-facteurs (correctifs spécifiques à certaines O-D) en utilisant des données d'enquête déplacements sur la ville de Winnipeg pour 1962, 1971, 1976 et

1981. Ils observent que l'allongement de l'horizon de prévision se traduit par une baisse de la qualité des prévisions réalisées par le modèle. L'inclusion des K-facteurs permet d'améliorer le pouvoir descriptif des modèles, mais se traduit également par une baisse de leur pouvoir prospectif. Ceci est sans doute lié à l'absence de fondements comportementaux permettant de déterminer les K-facteurs.

Plus récemment, ELMI et alii (1999) ont entrepris une étude sur la transférabilité temporelle du modèle gravitaire pour les déplacements liés au motif Travail. Cette étude est très proche de celle de VOLET et HUTCHINSON (1986) et la complète sur plusieurs aspects. L'aire d'étude est identique : Toronto, bien que ELMI et alii (1999) utilisent des données de trois enquêtes déplacements (la *Metropolitan Toronto and Region Transport Study* de 1964 et les *Transportation Tomorrow Survey* de 1986 et 1996) sur une période de 30 ans. Le modèle testé est le même, c'est-à-dire le modèle gravitaire doublement contraint avec une fonction de résistance exponentielle. ELMI et alii (1999) enrichissent leur étude de quatre segmentations du modèle : en fonction de la zone principale d'activité (relation centre-périphérie), du sexe, de la disponibilité d'un véhicule particulier et du type d'activité des agents. Les résultats de l'étude montrent une variabilité du paramètre de la fonction de résistance dans le temps, qui suggère une évolution des comportements et une inclinaison à consacrer plus de temps aux déplacements. Globalement, les modèles segmentés ont une meilleure performance descriptive, notamment le modèle qui utilise la segmentation par type d'occupation. La performance prospective des modèles segmentés est également meilleure, ce qui semble aller à l'encontre des études que nous avons présentées jusqu'ici. Néanmoins, les auteurs tempèrent leurs propos en suggérant que, d'un point de vue pratique (outil d'aide à la décision), le gain apporté par la segmentation ne semble pas suffisant pour recommander son introduction dans les études de terrain. Ainsi, les auteurs concluent que le modèle gravitaire doublement contraint usuel apparaît comme étant satisfaisant pour l'usage en modélisation.

Dans un cas un peu à part, MIKKONEN et LUOMA (1999) s'intéressent à l'évolution des paramètres d'un modèle d'interaction spatiale par paliers (à ne pas confondre avec une fonction de résistance par paliers) sur les villes de Vaasa, Seinäjoki, Kokkola et Pietarsaari en 1971, 1980-81 et 1991. Leur étude est un peu différente des autres études présentées ici car le modèle utilise la distance et non le temps de déplacement comme variable en entrée. Ils trouvent une évolution du paramètre de la fonction de résistance mais cette évolution est non monotone (on se serait attendu normalement à une réduction continue du paramètre). Les auteurs trouvent une explication à ce phénomène dans la localisation des gros centres commerciaux.

Les quelques études de caractère empirique sur la transférabilité du modèle gravitaire semblent converger sur l'idée que les affinements nécessaires pour

reproduire une situation de calage de manière plus satisfaisante rendent hasardeuse l'utilisation du modèle en prospective. La recherche d'une alternative s'est révélée difficile. Les deux dernières décennies ont vu le développement de l'application des techniques du champ de l'intelligence computationnelle (logique floue, réseaux de neurones...) aux problèmes d'optimisation dans le domaine des transports. Ainsi, à partir du milieu des années 1990, on a assisté à l'apparition d'une série d'études sur l'utilisation d'une architecture de type « réseaux de neurones » pour l'étape de distribution (voir notamment BLACK, 1995 ; OPENSHAW, 1998). Elles suggèrent que cette approche serait plus performante que les modèles de type gravitaire. Cependant, MOZOLIN et alii (2000) avancent que la supposée meilleure performance des modèles en réseaux de neurones est uniquement descriptive et non prédictive. Ils montrent, en utilisant des données de recensement relatives aux déplacements domicile-travail de 1980 et 1990 sur l'aire métropolitaine d'Atlanta, qu'un modèle gravitaire simple, avec une fonction de résistance exponentielle, prédit mieux une matrice O-D future qu'un modèle de type réseaux de neurones. On retrouve donc un cadre de figure similaire à celui dépeint par OPENSHAW et SOUTHWORTH : les nouvelles approches sont plus performantes en matière descriptive -elles permettent de gérer significativement mieux les situations où l'information est rare- mais elles sont moins performantes que l'approche classique en prévision.

3. MÉTHODOLOGIE

L'objectif de cette étude est d'évaluer la capacité du modèle gravitaire à faire des prévisions sur un horizon temporel de moyen-long terme (10 ans et plus), et de voir s'il est capable de rendre compte de l'allongement des distances observé ces vingt dernières années. Pour ce faire, nous étudions le cas de l'agglomération lyonnaise à partir des données des enquêtes ménages déplacements (EMD) de 1985, 1995 et 2006 et d'une codification des réseaux correspondant à ces trois dates et permettant d'avoir des temps de parcours comparables dans le temps. Ces données permettent de tester le pouvoir prédictif du modèle *a posteriori*, autrement dit de vérifier si les paramètres de calibration d'une année permettent de reproduire raisonnablement bien les données observées à une autre année. Cette section commence par une description de l'unité d'analyse de la mobilité retenue (3.1.), puis quantifie l'allongement des distances observé (3.2.). Elle donne également les spécifications des modèles étudiés (3.3.) et la méthode de calibration ainsi que les indicateurs permettant d'analyser les résultats (3.4.).

3.1. UNE APPROCHE EN TERMES DE SORTIES DU DOMICILE

L'unité de description de la mobilité retenue est la sortie du domicile (AXHAUSEN, 2000 ; BONNEL, 2004), c'est-à-dire la séquence de déplacements en lien avec les différentes activités réalisées par les résidents de l'aire

d'étude à partir du moment où ceux-ci sortent de leur domicile jusqu'au moment où ils y reviennent. Le choix de la sortie plutôt que du déplacement comme unité d'analyse repose sur :

- . un désir de comparabilité entre les données des différentes enquêtes : la définition du déplacement est restée stable dans la méthode « standard Certu » (CERTU, 2008). En revanche, selon la qualité des enquêteurs et du contrôle, il est possible que l'application de la définition ait pu changer d'une enquête à une autre, notamment pour les arrêts courts comme acheter du pain ou prendre de l'essence qui sont censés être identifiés au cours de l'enquête. Le choix de la sortie comme unité d'analyse permet de minimiser l'impact de ce risque d'inconsistance dans les données ;
- . une symétrisation des flux des matrices O-D par motif à la journée, en cohérence avec les données produites par le modèle de distribution gravitaire ;
- . une meilleure intégration avec les étapes de génération et répartition modale. Au niveau de la génération, la définition en termes de déplacements peut conduire à perdre le motif principal de la sortie du domicile pour les sorties complexes (plus de 2 déplacements), réduisant d'autant la pertinence des lois de génération. Au niveau de la répartition modale, les sorties comportent une forte homogénéité modale, au moins pour les modes motorisés. Dans une approche en termes de sorties, le motif principal justifiant la sortie peut être conservé et la cohérence modale interne à la sortie peut être assurée (BONNEL, 2004).

Dans une approche en termes de déplacements, la distribution conduit à relier les émissions et attractions entre elles. Elle conduit donc à relier l'origine et la destination des déplacements. Dans une sortie, l'origine et la destination sont au domicile. Il est donc nécessaire d'adapter la méthode pour identifier l'origine et la destination de chacun des déplacements inclus dans la sortie. Dans la pratique, il s'agit de déterminer le lieu de réalisation de chacune des activités incluses dans la sortie. L'analyse est rapidement complexe dans la mesure où l'observation des sorties met en évidence la variabilité du nombre d'activités présentes au sein des sorties et surtout la très grande diversité des combinaisons d'activités. Il est donc difficile de définir des typologies simples. Nous avons défini huit motifs de déplacement correspondant aux différents types d'activités, en les classant hiérarchiquement du plus contraint au moins contraint : Travail > Enseignement (primaire, secondaire et supérieur) > Achats-Services > Accompagnement > Loisirs > Autres. Le motif principal de la sortie est défini en retenant le plus élevé des motifs présents dans la sortie. Le lieu de résidence et le lieu de réalisation de l'activité principale correspondent à l'origine et à la destination de la sortie. D'un point de vue comptable, au niveau de la matrice O-D, il y a une demi-sortie aller (domicile vers lieu de réalisation de

l'activité principale) et une demi-sortie retour (lieu de réalisation de l'activité principale vers domicile). Cette manière de procéder repose sur le fait qu'une analyse préalable des sorties complexes, à partir des données de l'enquête ménages déplacements de 2006, montre que la distance supplémentaire générée par les activités autres que celle correspondant au motif principal de la sortie est très faible par rapport à la distance aller-retour entre le domicile et le lieu du motif principal de la sortie.

3.2. ÉVALUATION DE L'ALLONGEMENT DES DISTANCES OBSERVÉES

Les modèles de distribution gravitaire sont le plus souvent appliqués par motif car les distributions de distances sont généralement très différentes selon les motifs comme le montre le Tableau 1 qui donne l'évolution des distances moyennes à vol d'oiseau d'une demi-sortie du domicile pour les motifs Travail et Loisirs quel que soit le mode de transport utilisé. Une distinction est faite par périmètre de résidence des ménages (Cf. Figure 2). Les deux premières lignes du tableau, pour chaque motif, donnent des évolutions de distance à périmètre de résidence constant (périmètre couvert par les EMD de 1985 et 1995). La dernière ligne donne les distances moyennes calculées à partir de l'ensemble des ménages enquêtés (en ne retenant que les sorties réalisées au sein du périmètre). L'augmentation des distances de déplacement peut se voir clairement pour les motifs Travail et Loisirs, sur tous les périmètres. Les distances moyennes augmentent avec le temps mais également avec le périmètre choisi : les distances parcourues par les résidents des zones les plus éloignées du centre sont les plus grandes. Les distances moyennes entre l'origine et la destination pour le motif Travail sont supérieures à celles pour le motif Loisirs, en accord avec l'intuition selon laquelle on est prêt à aller plus loin lorsque le motif de la sortie est plus « contraint ».

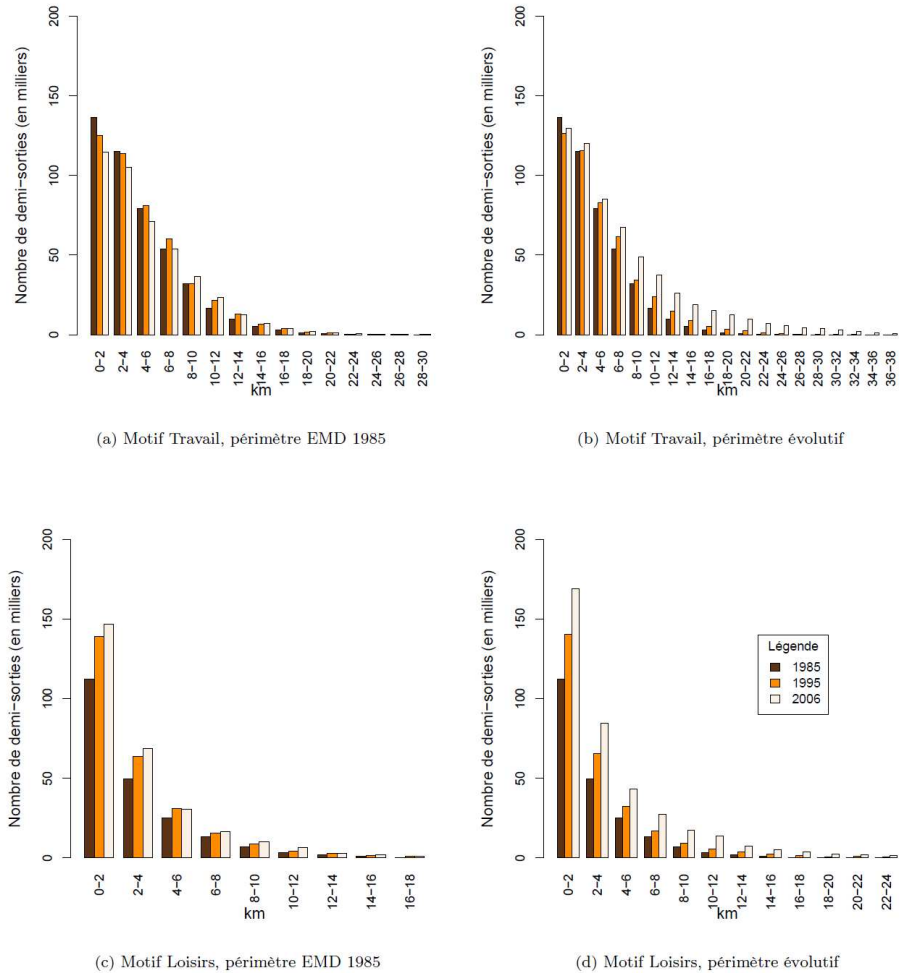
Tableau 1 : Distances moyennes à vol d'oiseau en km des demi-sorties, pour les motifs Travail et Loisirs

Date EMD	1985	1995	2006
Motif Travail			
Périmètre 1985	4,47	4,83	5,01
Périmètre 1995	-	5,17	5,47
Périmètre 2006	-	-	7,28
Motif Loisirs			
Périmètre 1985	2,88	2,97	3,03
Périmètre 1995	-	3,24	3,26
Périmètre 2006	-	-	4,18

Source : EMD 1985, 1995 et 2006. Périmètre : Cf. Figure 2.

La Figure 1 permet d'avoir le détail de la distribution des distances des demi-sorties. Pour le motif Travail, on observe que les sorties très courtes sont plus nombreuses en 1985 et, au fur et à mesure que la distance s'accroît, le nombre de sorties est plus important en 1995 et puis en 2006. Pour le motif Loisirs, on observe une augmentation généralisée du nombre de demi-sorties avec le temps. En termes relatifs elle est plus importante pour les distances plus longues, ce qui explique l'augmentation en moyenne constatée dans le Tableau 1.

Figure 1 : Évolution de la distribution des distances à vol d'oiseau des demi-sorties, pour les motifs Travail et Loisirs



Source EMD 1985,1995 et 2006. Périmètres : Cf. Figure 2.

Si on observe un allongement des distances pour les motifs Travail et Loisirs, en revanche, pour les autres motifs, on constate plutôt une stabilité entre 1985 et 2006. L'enjeu de l'analyse de la capacité du modèle de

distribution gravitaire à simuler les distances à un horizon de 10 à 20 ans est donc plus important pour les motifs Travail et Loisirs, ce qui explique que notre présentation se focalise sur ces deux motifs.

3.3. TROIS SPÉCIFICATIONS DU MODÈLE GRAVITAIRE SEGMENTÉES PAR MOTIF DE DÉPLACEMENT

Dans la famille des modèles gravitaires, la formulation qui a la meilleure performance prospective est celle du modèle doublement contraint⁷ (OPENSHAW, 1979 ; MOZOLIN et alii, 2000). Cette supériorité dépend de la bonne estimation des marges de la matrice O-D dans la phase de génération. Dans la mesure où évaluer la transférabilité temporelle du modèle gravitaire revient à, dans un premier temps, tester le pouvoir prédictif du modèle *a posteriori*, c'est-à-dire en utilisant des variables en entrée effectivement observées à l'horizon de prévision, notre étude fonctionne comme si nous avions fait une prévision parfaite des émissions et attractions. L'utilisation du modèle doublement contraint est donc parfaitement adaptée et elle peut servir de référence pour d'autres types de modèles.

Les modèles gravitaires sont également caractérisés par la forme de la fonction de résistance retenue. Classiquement, on utilise une fonction exponentielle avec un seul paramètre. Celle-ci présente l'avantage de permettre une interprétation relativement aisée des évolutions paramétriques (voir, par exemple, les indications données par ELMI et alii, 1999 ; MIKKONEN, LUOMA, 1999). L'utilisation d'un seul paramètre peut ne pas être suffisante pour reproduire de manière acceptable la distribution des déplacements par tranche de distance ou temps. Le recours à une fonction avec deux paramètres (dite « exponentielle-puissance » ou TANNER, 1961) peut être avantageux dans un tel cas⁸.

Nous estimons donc trois types de modèles gravitaires doublement contraints. Il convient de présenter les spécifications en commençant par le modèle à la TANNER donné dans les équations (1) à (3) :

$$T_{ij}^p = A_i^p O_i^p B_j^p D_j^p c_{ij}^{\alpha_p} \exp(\beta^p c_{ij}) \quad (1)$$

$$A_i^p = \left(\sum_j B_j^p D_j^p c_{ij}^{\alpha_p} \exp(\beta^p c_{ij}) \right)^{-1} \quad (2)$$

$$B_j^p = \left(\sum_i A_i^p O_i^p c_{ij}^{\alpha_p} \exp(\beta^p c_{ij}) \right)^{-1} \quad (3)$$

où

⁷ C'est-à-dire à la fois sur les émissions et les attractions qui constituent les marges en ligne et en colonne de la matrice origine-destination.

⁸ En réalité, à partir du moment où l'on a des matrices O-D tous modes (lorsque la distribution précède la répartition modale, par exemple), les demi-sorties courtes sont très nombreuses et les « pics » sont mal reproduits par la fonction de résistance exponentielle.

T_{ij}^p est le nombre de demi-sorties ayant comme origine la zone i et ayant le motif principal p situé dans la zone j ;

O_i^p est le nombre de demi-sorties ayant comme origine la zone i et ayant le motif principal p (il s'agit donc des émissions de la zone i) ;

A_i^p et B_j^p sont des paramètres appelés facteurs de balancement estimés lors de la procédure itérative visant au respect de la contrainte aux marges dans un modèle doublement contraint ;

D_j^p est le nombre de demi-sorties ayant comme destination la zone j et ayant le motif principal p (il s'agit donc des attractions de la zone j) ;

c_{ij} est le temps généralisé pour aller de la zone i à la zone j .

$$c_{ij}^{\alpha^p} \exp(\beta^p c_{ij}) \quad (4)$$

est la fonction de résistance à deux paramètres, α^p et β^p . Le modèle gravitaire exponentiel simple correspond au cas où le paramètre α^p est nul. La fonction de résistance (4) peut alors s'écrire comme indiqué en (5).

$$\exp(\beta^p c_{ij}) \quad (5)$$

Il convient d'avoir un modèle de référence par rapport auquel on puisse comparer la performance de ces deux modèles. On peut prendre un modèle simple à facteurs de croissance (DEMING, STEPHAN, 1940 ; FURNESS, 1965), qui correspond au cas où α^p et β^p sont tous les deux nuls. Il s'agit d'un modèle purement théorique de référence dans la mesure où l'application des facteurs de croissance se fait d'habitude à partir d'une matrice O-D observée dans le passé.

3.4. CALIBRATION ET INDICATEURS

3.4.1. Estimation des paramètres par la méthode du maximum de vraisemblance

ERLANDER et STEWART (1990) ont mis en évidence un principe d'équivalence mathématique entre les modèles dérivés suivant chacune des différentes approches théoriques évoquées dans le début de la section 2. La traduction pratique de cette équivalence est que l'on peut supposer que l'une des approches est la plus adaptée à la modélisation, d'un point de vue théorique, et utiliser une autre approche pour effectuer les calculs. Ce principe d'équivalence sera utile, ici, au moment de proposer une méthode pour calibrer le modèle. Ainsi, nous pouvons supposer que notre modèle a été défini, de manière usuelle, à partir de la maximisation de l'entropie sous une contrainte de coût (ou de temps donnée par les équations (8) et (9)), et pour déterminer les estimateurs par maximum de vraisemblance, supposer que le nombre de

demi-sorties observés entre i et j , N_{ij}^p , suit une loi de Poisson de paramètre T_{ij}^p de sorte que $E(N_{ij}^p) = T_{ij}^p$.

BATTY et MACKIE (1972) et KIRBY (1974), ainsi que, plus tard, ERLANDER et STEWART (1990) et SEN et SMITH (1995) ont montré qu'il suffit de résoudre les systèmes définis par les équations (6) à (8) et (6) à (9) pour trouver les estimateurs par maximum de vraisemblance des paramètres du modèle gravitaire doublement contraint avec une fonction de résistance de type exponentielle et TANNER⁹, respectivement.

$$\sum_j T_{ij}^p = O_i^p \quad (6)$$

$$\sum_i T_{ij}^p = O_j^p \quad (7)$$

$$\sum_{ij} c_{ij} T_{ij}^p = \sum_{ij} c_{ij} N_{ij}^p \quad (8)$$

$$\sum_{ij} \log(c_{ij}) T_{ij}^p = \sum_{ij} \log(c_{ij}) N_{ij}^p \quad (9)$$

Les paramètres α^p et β^p sont ainsi définis de manière unique.

Afin de résoudre les systèmes non linéaires décrits par les équations (6) à (9), on utilise la procédure MSC (*modified scoring procedure*) décrite dans SEN et SMITH (1995, chapitre 5), que l'on a implémentée en utilisant le logiciel statistique R (R DEVELOPMENT CORE TEAM, 2011).

Il est nécessaire de préciser que les équations (8) et (9), connues comme contraintes sur les coûts (temps) moyens, jouent un rôle majeur dans la formulation théorique du modèle gravitaire. Leur introduction donne la forme de la fonction de résistance. Malgré cela, les méthodes de calibration usuellement retenues pour le modèle ne sont pas toujours compatibles avec ces contraintes. Dans le cas présent, la méthode retenue les prend en compte. Il s'agit d'une méthode de celles connues comme « ciblant le coût moyen » (SMITH, HUTCHINSON, 1981).

3.4.2. Évaluation de la qualité des calages

La mesure de la qualité de la reproduction des matrices observées par le modèle calibré nécessite la définition de certains indicateurs pertinents. Pour ce faire, on s'est appuyé sur la littérature existante. Les indicateurs retenus sont présentés dans ce qui suit.

Le premier indicateur découle naturellement de la méthode de calage retenue : il s'agit de la log-vraisemblance.

⁹ Ces équations sont, en effet, les conditions de premier ordre de la maximisation de la fonction de log-vraisemblance.

La très grande majorité des études de validation empirique du modèle gravitaire utilisent au moins une statistique générale de comparaison « cellule à cellule » des matrices modélisées et observées, le plus souvent en rapport avec l'erreur quadratique moyenne. Nous avons retenu pour cette étude deux indicateurs de ce type. Le premier est le $SRMSE$ ¹⁰ (*standardized root mean square error*), donné dans l'équation (10) suivante :

$$SRMSE = \frac{\sqrt{\frac{\sum_{ij} (N_{ij} - T_{ij})^2}{I \cdot J}}}{\frac{\sum_{ij} N_{ij}}{I \cdot J}} \quad (10)$$

où l'on a omis les motifs p pour faciliter la lecture. I et J sont respectivement le nombre d'origines et de destinations émettant ou attirant au moins une demi-sortie. D'après FOTHERINGHAM et KNUDSEN (1987), l'utilisation du $RMSE$ est préférable à celle des autres types de $SRMSE$ car ces derniers sont très sensibles aux ordres de grandeur des données. Dans la mesure où nous étudions des modèles sur trois périmètres différents, cette statistique est ainsi pertinente. La sensibilité du $SRMSE$ aux erreurs est linéaire, ce qui lui confère, d'après FOTHERINGHAM et KNUDSEN, une supériorité par rapport aux autres mesures générales de distance « cellule à cellule » pour comparer un même modèle dans deux contextes différents.

La deuxième mesure de distance entre matrices est une forme particulière de $RMSE$, utilisée par ELMi et alii (1999). Nous l'avons introduite dans un souci de comparabilité des résultats. Elle prend la forme suivante :

$$RMSE = \left[\frac{\sum_{ij} \frac{(N_{ij} - T_{ij})^2}{T_{ij}}}{\sum_{ij} T_{ij}} \right]^{\frac{1}{2}} \quad (11)$$

3.4.3. Évaluation de la transférabilité des modèles

L'analyse de la transférabilité prend appui en général sur les indicateurs de qualité de calage. Il s'agit de comparer les résultats des simulations faites avec les paramètres de calage correspondant à la date des données observées et ceux utilisant des paramètres d'une date antérieure.

Comme le notent McARTHUR et alii (2011), les revues récentes de la littérature relative à la transférabilité temporelle des modèles de prévision de la demande de transport (ORTÚZAR, WILLUMSEN, 2001, sont évoqués explicitement mais on pourrait y ajouter KARASMAA, 2008 et FOX et HESS, 2010) retiennent deux indicateurs principaux : la statistique de test de

¹⁰ Racine carrée normalisée de l'erreur quadratique moyenne.

transférabilité (TTS, du sigle en anglais) et l'indice de transfert (TI, idem)¹¹. Cependant, McARTHUR et alii remarquent que l'indicateur TTS semble mal s'adapter à l'étude du modèle gravitaire en raison des non-linéarités inhérentes à celui-ci. On peut penser que cette raison a également conduit ELMi et alii (1999) à ne pas prendre en compte la statistique TTS et à privilégier l'utilisation d'autres indicateurs. Nous avons donc décidé de procéder de la même manière et de ne pas l'utiliser et de ne retenir que l'indice TI. Celui-ci (KOPPELMAN, WILMOT, 1982) est moins une mesure d'égalité paramétrique qu'une mesure de performance relative :

$$TI_a(\theta_r) = \frac{LL_a(\theta_r) - LL_a(\theta_{REF})}{LL_a(\theta_a) - LL_a(\theta_{REF})} \quad (12)$$

où $LL_a(\theta_a)$ est la valeur de la log-vraisemblance du modèle calibré à partir de données observées à la date a . θ_a correspond aux paramètres issus de ce calage¹². $LL_a(\theta_r)$ est la valeur de la log-vraisemblance du modèle simulé à la date d'application a avec des données d'émissions, d'attractions et de temps généralisé de la date a , mais à partir des paramètres de calage θ_r estimés à une date antérieure r . θ_{REF} correspond aux paramètres d'un modèle de référence ou nul (facteurs de croissance). L'indicateur décrit le degré avec lequel la log-vraisemblance du modèle transféré améliore un modèle nul ou de référence, par rapport à l'amélioration apportée par un modèle développé dans le contexte d'application.

Les indicateurs donnés par (10) et (11) sont des mesures absolues de l'erreur produite. On peut les utiliser pour produire un indicateur permettant de comparer la performance des modèles en description et en prévision. ELMi et alii (1999) utilisent le *RATE* (*relative aggregate transfer error*) qui compare la valeur des erreurs issues d'une simulation faite avec les paramètres transférés à celle des erreurs issues du calage du modèle correspondant à la date d'application. Il est défini dans les équations (13) et (14) ci-dessous :

$$RATE_S = \frac{SRMSE_a(\theta_r)}{SRMSE_a(\theta_a)} \quad (13)$$

$$RATE_R = \frac{RMSE_a(\theta_r)}{RMSE_a(\theta_a)} \quad (14)$$

correspondant aux deux indicateurs de distance entre matrices définis par (10) et (11), respectivement.

¹¹ Les études sur la transférabilité des modèles portent très majoritairement sur les modèles désagrégés. Les exemples d'utilisation de ces indicateurs dans le cas de modèles gravitaires sont peu nombreux, à la connaissance des auteurs. On peut citer ELMi et alii (1999) et McARTHUR et alii (2011), notamment.

¹² Soit β , pour le modèle avec une fonction de résistance exponentielle, et α et β , pour celui avec une fonction de résistance de type TANNER.

Pour l'étude de l'allongement des distances de déplacement, il convient de regarder la distribution de celles-ci. Notre approche diffère néanmoins de l'utilisation classique de cette distribution. En effet, la pratique habituelle est de regarder la distribution de la variable utilisée pour mesurer la « séparation » entre les zones de l'aire d'étude. Dans le cas présent, cela reviendrait à analyser la distribution des demi-sorties par tranche de temps (c_{ij}). L'objectif est ici de voir si l'application du modèle permet de reproduire l'augmentation des distances moyennes de déplacement observées. Il est plus intéressant de regarder la distribution des distances à vol d'oiseau entre les zones de l'aire d'étude. Il convient de rappeler que la méthode d'estimation retenue permet de reproduire le temps moyen observé, mais il n'y a aucune contrainte sur les distances à vol d'oiseau. L'indicateur qui s'impose naturellement pour étudier la proximité des distributions des distances observées et simulées est le χ^2 de Pearson donné par :

$$\chi_{dist}^2 = \sum_{d=1}^D \frac{(N_d - T_d)^2}{T_d} \quad (15)$$

où d est une classe de distance appartenant à $\{1 \dots D\}$. N_d et T_d sont le nombre de demi-sorties observées et modélisées, respectivement, appartenant à la classe de distance d .

4. SOURCES DES DONNÉES, PÉRIMÈTRES ET DÉFINITION DU ZONAGE

La calibration d'un modèle gravitaire fait appel à deux types de données : le premier correspond aux données permettant d'établir les matrices origine-destination observées (N_{ij}). Il s'agit des données de la demande (4.1.). Le deuxième type correspond aux données de « séparation » entre origine et destination. Notre étude utilise le temps généralisé en voiture particulière (c_{ij}). Celui-ci est obtenu à partir de la modélisation du réseau routier. On parle de données d'offre ou de niveau de service du réseau (4.2.). Nous terminons cette section avec quelques données sur l'évolution de l'offre de transport au sein de l'agglomération lyonnaise (4.3.).

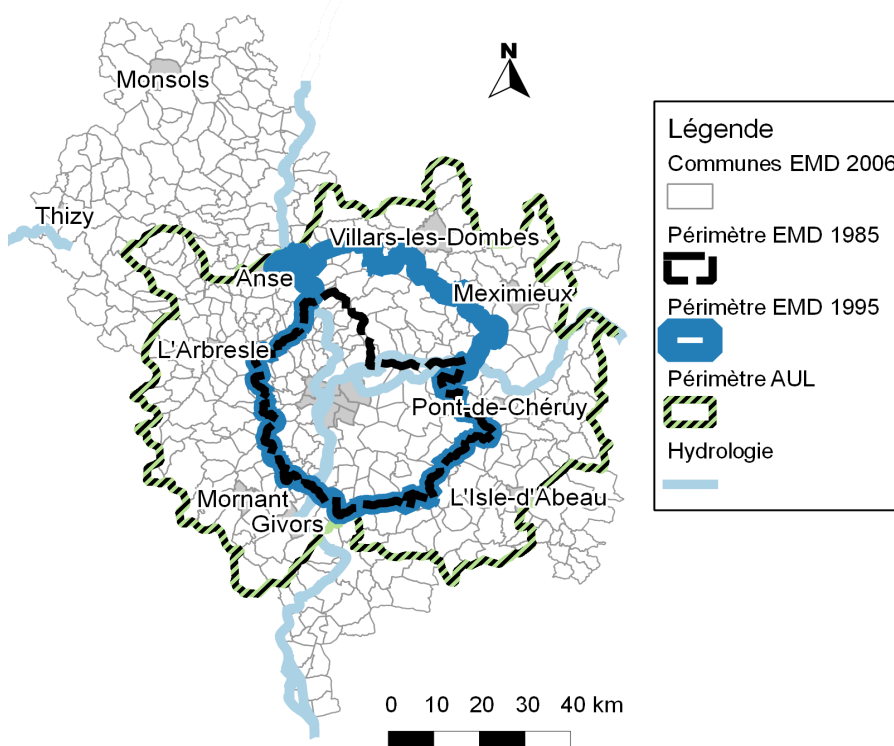
4.1. LES DONNÉES RELATIVES À LA DEMANDE

Les données relatives à la demande, c'est-à-dire au nombre de sorties par paire d'O-D, proviennent des trois dernières enquêtes ménages déplacements (EMD) réalisées sur l'agglomération lyonnaise (1985, 1995 et 2006). Les EMD sont l'un des instruments essentiels pour l'étude de la mobilité et de son évolution en France (CERTU, 2008). Les EMD visent à reconstituer la mobilité des habitants pour un jour ouvrable « moyen » de semaine, hors vacances scolaires. Tous les déplacements de toutes les personnes de cinq ans et plus habitant dans le logement sont interrogées en personne. L'échantillonnage vise à une représentativité statistique des déplacements des

résidents de plus de 5 ans du périmètre enquêté¹³.

Le respect de la méthode d'enquête « standard Certu » assure, en principe, la comparabilité des données dans le temps. Dans la pratique, il est nécessaire de prendre certaines précautions pour s'assurer de la comparabilité des différents champs nécessaires au calage du modèle gravitaire. Le premier élément à prendre en compte est l'évolution du périmètre : le périmètre des enquêtes a été étendu au cours du temps. En 1985, 1995 et 2006, l'enquête portait respectivement sur 71, 99 et 453 communes. La Figure 2 donne un aperçu des communes appartenant aux différents périmètres des EMD.

Figure 2 : Communes de l'agglomération lyonnaise incluses dans le périmètre de l'EMD de 2006



Le deuxième aspect à prendre en compte est la question du zonage¹⁴. En réalité, pour comparer les modèles calibrés à plusieurs dates, il est nécessaire de disposer d'un zonage commun pour la définition des différentes matrices

¹³ Les données EMD disposent d'un coefficient de redressement permettant de retrouver la population totale à partir de la population enquêtée. Dans ce travail, on a utilisé des coefficients normalisés par rapport au taux de sondage afin de conserver le poids relatif des observations sans gonfler artificiellement l'effectif observé.

¹⁴ Les adresses exactes des ménages enquêtés dans les EMD ne sont pas géocodées et les zonages changent pour chaque nouvelle enquête.

O-D. Autrement, il est possible que les différences observées entre deux calages soient en partie liées aux différences de zonage. Un cas extrême est étudié par OPENSHAW (1977) qui montre qu'il est possible de calibrer un modèle gravitaire en choisissant des valeurs paramétriques *a priori* et en modifiant le zonage pour faire la correspondance. Par ailleurs, comme le notent MOZOLIN et alii (2000), il n'est pas rare d'observer une moindre performance des modèles lorsque le niveau de zonage est très fin. Ces aspects font partie de ce que la littérature géographique appelle problème de l'unité surfacique modifiable (*modifiable areal unit problem*, ou encore MAUP, voir OPENSHAW, 1984 ; BRIANT et alii, 2010). Le traitement de la question dans le contexte des modèles gravitaires a été fait par BATTY et SIKDAR (1982a ; 1982b ; 1982c ; 1982d) et va au-delà de la portée de la présente étude.

On a retenu deux zonages pour la construction des matrices O-D observées. Le premier, très fin, correspond aux IRIS2000, unité spatiale la plus fine de diffusion des données définie par l'INSEE, et le second, plus agrégé, correspond aux communes (arrondissements pour Lyon). Le zonage à l'IRIS permet d'avoir un nombre de demi-sorties intrazones faible. Il correspond à l'échelle de codification du réseau.

Les zonages à l'IRIS2000 et communal comportent 523 et 79 zones respectivement dans le périmètre de l'EMD 1985, 555 et 107 zones dans le périmètre de l'EMD 1995 et 777 et 302 zones dans le périmètre de l'aire urbaine lyonnaise (AUL) qui est inclus à l'intérieur du périmètre de l'EMD 2006 (Cf. Carte 2). Le Tableau 2 permet d'avoir un aperçu des données de demande. On y remarque un accroissement du nombre de déplacements par personne dans le périmètre de l'EMD 1985, entre 1985 et 1995 puis une stagnation. On remarque parallèlement que l'augmentation du nombre de sorties par personne est moins importante, ce qui témoigne d'une complexification de celles-ci. Le passage du zonage en IRIS2000 aux communes induit une très forte augmentation des intrazones car une grande partie des demi-sorties courtes visibles sur la Figure 1 sont réalisées dans le centre de l'agglomération.

Tableau 2 : Récapitulatif des données présentes dans les enquêtes ménages déplacements

EMD/Périmètre	enquête	redressé	Déplacements par personne	Sorties par personne	Sorties intrazones (%)	
					IRIS2000	Commune
1985/P1985	11 449	1 017 893	3,48	1,46	9,37	53,01
1995/P1985	13 579	1 138 161	3,90	1,55	8,51	50,05
1995/P1995	13 586	1 146 520	3,62	1,47	8,41	49,59
2006/P1985	13 997	1 195 190	3,90	1,56	8,80	51,09
2006/P1995	14 523	1 209 987	3,64	1,48	9,05	50,25
2006/AUL	20 302	1 537 593	3,70	1,49	9,87	49,60

Source : EMD 1985, 1995 et 2006.

4.2. LES DONNÉES DE NIVEAU DE SERVICE TRANSPORT

Un des principaux obstacles à l'évaluation de la performance prospective du modèle gravitaire est le manque de données comparables sur plusieurs dates, notamment en ce qui concerne les temps de parcours. Ainsi, par exemple, OPENSHAW (1979) décrit Durham comme un cas d'étude spécial parce que des données de temps de parcours fiables et comparables sont disponibles à deux dates. En réalité, en l'absence de données de temps comparables, les variations des valeurs des paramètres ne traduisent pas seulement des changements dans les comportements de mobilité. Elles peuvent, tout simplement, être le reflet des inconsistances entre les mesures de temps utilisées. La propriété de consistance fait ici référence à l'utilisation d'un même outil pour mesurer l'évolution d'une variable entre deux dates.

Les temps déclarés provenant des différentes enquêtes sont parfois utilisés, mais ils peuvent ne pas être comparables entre eux, notamment lorsque les besoins de désagrégation des données sont importants¹⁵. Dans la mesure où la modélisation s'appuie sur une codification des réseaux pour produire des temps de parcours dans les exercices de prévision, il est apparu naturel d'avoir recours à la codification de l'historique des réseaux. En suivant certaines règles prédéfinies, on peut produire des données de temps comparables sur plusieurs dates, nécessaires à ce travail. Une telle donnée était indisponible en France et une partie de ce travail a donc été consacrée à la production de données. La codification de réseaux étant un travail lourd et extrêmement chronophage, il semblait raisonnable de prendre appui sur des efforts de codification existants, plutôt que de se lancer dans la codification de réseaux *ex nihilo*. Ainsi, nous avons décidé de prendre comme point de départ les réseaux codifiés pour le projet SIMBAD (NICOLAS et alii, 2009). En suivant la logique de codification définie pour le projet SIMBAD, on a construit les réseaux pour les dates des EMD (1985, 1995 et 2006) de manière à s'assurer de leur comparabilité dans le temps en recherchant le plus possible l'automatisation de la codification (GODINOT, BONNEL, 2008). Le réseau a été codifié en utilisant le logiciel DAVISUM, sur le périmètre de l'aire urbaine lyonnaise de 1999. On dispose donc de matrices O-D de temps généralisés consistants pour les dates des EMD. A la date de rédaction de l'article, seul le réseau routier a été complètement finalisé. C'est la raison pour laquelle nous avons utilisé uniquement le temps généralisé voiture. Cette simplifica-

¹⁵ Si les temps déclarés sont consistants au niveau agrégé, ils le sont beaucoup moins lorsqu'on les analyse à un niveau plus désagrégé. Le faible nombre de personnes enquêtées par O-D rend leur utilisation délicate en modélisation. En effet, les temps déclarés sont sujets à des effets de perception et d'arrondis. Un exemple permet d'illustrer le problème. Si un déplacement domicile-travail d'une personne en voiture prend 6 minutes à la date t_1 et 8 minutes à la date t_2 , les effets d'arrondis peuvent faire que, lors de l'enquête, elle déclarera mettre 5 minutes en t_1 et 10 minutes en t_2 (arrondi au multiple de 5 le plus proche) pour réaliser le trajet. À la place d'avoir un écart de 2 minutes, on en a un de 5 minutes.

tion n'est toutefois pas susceptible de remettre significativement en cause les conclusions auxquelles nous parvenons car la pratique modélisatrice habituelle est de retenir pour le modèle de distribution, le coût/temps généralisé du mode le plus rapide, qui est généralement la voiture à l'exception de certains déplacements dans les zones centrales, les temps/coûts généralisés des transports collectifs étant fortement pénalisés par les temps de marche et d'attente perçus négativement par les individus.

4.3. Principales évolutions du système de transport de l'agglomération lyonnaise

L'agglomération lyonnaise fait aujourd'hui partie des agglomérations françaises disposant des meilleurs réseaux de transports collectifs que l'on considère des indicateurs d'offre ou de demande. Ce réseau s'est construit progressivement depuis l'ouverture des premières lignes de métro (A, B et C) à la fin des années 70, puis la ligne D entièrement automatique en 1991. Mis à part quelques prolongements de lignes de métro, le développement des axes structurants s'est ensuite organisé autour des lignes de tramway T1 et T2 en 2001, puis T3 en 2006 et des lignes fortes dotées de véhicules articulés, de fortes fréquences et de mesures de protection vis-à-vis de la circulation offrant un assez bon niveau de régularité. Ces développements des transports collectifs ont dans un premier temps (jusqu'à la fin des années 90) été réalisés sans mesure de maîtrise significative de la place de la voiture, voire en parallèle d'investissements routiers (construction de nombreux parcs de stationnement, périphérique Nord, rocade de contournement par l'Est débutée au début des années 80 et terminée en 1993). La part de marché des transports collectifs par rapport à la voiture est ainsi restée constante entre 1976 et 1995 d'après les données des enquêtes ménages déplacements. Dans le courant des années 90, puis plus nettement au 21^e siècle, la politique de transport s'est orientée vers un développement plus marqué en faveur des transports collectifs dans un premier temps, puis de l'ensemble des modes « doux » (mise en place des vélos en libre-service VELO'V en 2005, aménagement des places du centre ville...) avec une maîtrise de plus en plus marquée de la place de la voiture. Cette évolution est forte dans les zones les plus centrales, mais de moins en moins visible à mesure que l'on s'éloigne vers la périphérie de l'agglomération. On observe ainsi dans les données des EMD une baisse de la part de la voiture en nombre de déplacements marquée dans les déplacements liés au centre, mais qui s'amenuise à mesure que l'on s'en éloigne¹⁶. Toutefois, ces évolutions ont peu impacté les tendances concernant les distances automobiles qui continuent à croître malgré la

¹⁶ Nous ne souhaitons aborder ici la question de la causalité car il est évident que d'autres facteurs ont également pu jouer pour expliquer ces évolutions : vieillissement de la population, crise économique, prix des carburants... Une analyse plus détaillée des résultats des EMD est disponible dans SYTRAL (1990 ; 1996 ; 2007) et LET, SEMALY (1997).

diminution du nombre de déplacements en voiture du fait de la croissance des distances moyennes (Cf. Tableau 1).

5. RÉSULTATS

Nous présentons tout d'abord les résultats de la calibration des modèles et ensuite leur performance en prévision. Le modèle à deux paramètres donne systématiquement de meilleurs résultats que le modèle exponentiel classique. Ce résultat est conforme à ce que l'on attendait au vu des données retenues pour la constitution de la matrice O-D observée, celles-ci incluant les demi-sorties tous modes confondus. En conséquence, la présentation se focalise sur le modèle à la TANNER.

5.1. RÉSULTATS DE CALAGE

Le Tableau 3 donne les résultats des calages du modèle gravitaire à la Tanner au niveau des IRIS2000. On y trouve les valeurs des coefficients α et β ainsi que les indicateurs présentés dans la section 3.4.2.

Tableau 3 : Résultats des calages du modèle TANNER par année et périmètre pour les motifs Travail et Loisirs (zonage IRIS2000)

EMD/Périmètre	α	β	$LL_a(\theta_a)$	RMSE	SRMSE
<i>p=Travail</i>					
1985/P1985	-0,551	-0,085	-1,742E+04	1,07	2,52
1995/P1985	-0,468	-0,080	-1,820E+04	1,05	2,34
1995/P1995	-0,480	-0,078	-1,920E+04	1,07	2,4
2006/P1985	-0,551	-0,076	-2,020E+04	0,999	2,35
2006/P1995	-0,559	-0,075	-2,175E+04	1,03	2,37
2006/AUL	-0,622	-0,067	-2,881E+04	1,17	2,99
<i>p=Loisir</i>					
1985/P1985	-0,779	-0,134	-8,586E+03	1,9	6,85
1995/P1985	-0,735	-0,133	-1,039E+04	1,74	4,58
1995/P1995	-0,783	-0,124	-1,092E+04	1,86	4,79
2006/P1985	-0,681	-0,147	-1,253E+04	1,57	4,48
2006/P1995	-0,754	-0,136	-1,322E+04	1,85	4,65

Note de lecture : la première colonne du tableau spécifie les données utilisées : par exemple, à la deuxième ligne, « 1995/P1985 » indique qu'il s'agit des données de l'EMD 1995, dont on n'utilise que la partie couvrant le périmètre de l'EMD 1985, dans la construction de la matrice observée pour le motif Travail.

Pour chacun des périmètres, seules les demi-sorties internes au périmètre réalisées par les résidents du périmètre sont prises en compte. Cette remarque s'applique également aux données du Tableau 4.

Si l'on s'intéresse aux paramètres obtenus, le premier constat que l'on peut faire est qu'ils sont tous négatifs. Il s'agit encore de l'impact des sorties courtes. On observe également que les paramètres changent mais pas de

manière uniforme. Pour le paramètre α , l'écart le plus important est d'environ 30 % pour le motif Travail et 20 % pour le motif Loisirs. Il est intéressant de remarquer que l'écart maximum se trouve pour des données de la même année (2006) pour le motif Loisirs. Pour le paramètre β , l'écart le plus important est d'environ 25 % pour le motif Travail et 20 % pour le motif Loisirs. Entre les motifs, les paramètres du motif Loisirs sont plus grands en valeur absolue que ceux du motif Travail, ce qui, comme tous les paramètres sont négatifs, est en accord avec le fait que l'on se déplace plus longtemps pour aller travailler. Les variations à l'intérieur d'un motif sont moins importantes et souvent une augmentation en valeur absolue d'un paramètre correspond avec une variation dans le sens inverse de l'autre paramètre. Il est donc plus difficile d'en tirer une interprétation en relation avec une volonté de se déplacer plus longtemps.

On remarque que la qualité de l'ajustement cellule à cellule, telle que mesurée par les indicateurs *RMSE* et *SRMSE* est moyenne. On obtient, dans la majorité des cas, des valeurs supérieures à 1, ce qui veut dire que l'erreur moyenne est supérieure à la valeur moyenne de la cellule. Cependant les valeurs de ces indicateurs ne sont pas très élevées compte tenu du fait qu'ils ont été établis pour un zonage très fin. Il convient de remarquer ici que les valeurs obtenues sont du même ordre ou inférieures à celles qu'on a pu trouver dans la littérature (ELMI et alii, 1999, par exemple, obtiennent de valeurs pour le *RMSE* entre 7 et 15, alors qu'ici on obtient des valeurs allant de 0,999 à 2,68). Si l'on regarde motif par motif, on constate que l'ajustement cellule par cellule est bien meilleur pour le motif Travail. Les indicateurs *RMSE* et *SRMSE* sont environ deux fois plus grands pour le motif Loisirs que pour le motif Travail.

5.2. PERFORMANCE EN PRÉVISION

Le Tableau 4 donne des indicateurs de qualité d'ajustement permettant d'étudier la performance prospective des modèles de type TANNER calibrés en 1985 et appliqués en 1995 et 2006, sur les trois périmètres étudiés.

Tableau 4: Performance prospective du modèle de type TANNER avec les coefficients du calage de 1985/P1985 (zonage IRIS2000)

EMD/Périmètre	$\Delta \overline{dist}$ (%)	Π	RATE ₂	RATE ₅	χ^2 calage		χ^2 prévision avec coeff's de 1985		
					Valeur	$P(X < x)$	Valeur	$P(X < x)$	
<i>p</i> =Travail									
1995/P1985	1,942	0,989	1,035	1,032	14,126	0,721	12,316	0,83	
1995/P1995	3,186	0,988	1,042	1,029	13,784	0,909	14,908	0,866	
2006/P1985	1,078	0,994	1,020	1,025	12,144	0,84	6,423	0,994	
2006/P1995	2,763	0,992	1,033	1,026	11,485	0,967	9,829	0,988	
2006/AUL	8,185	0,984	1,102	1,023	19,123	0,981	48,441	0,0516	
<i>p</i> =Loisir									
1995/P1985	-2,781	0,999	1,028	0,999	22,776	0,064	15,912	0,319	
1995/P1995	1,967	0,998	1,090	0,996	28,958	0,0106	21,175	0,0972	
2006/P1985	-7,737	0,999	0,974	1,005	36,796	0,000792	40,989	0,000179	
2006/P1995	-4,432	1,000	0,996	1,001	45,657	3,19E-05	43,856	6,25E-05	
2006/AUL	3,092	0,999	1,305	0,997	60,419	1,94E-05	57,523	5,16E-05	

L'indice de transférabilité (TI) est très élevé. Sa valeur minimale est supérieure à 0,95 et, pour certaines années/périmètres elle s'approche de 1 (les effets d'arrondis font apparaître la valeur 1,000 dans ce cas). On peut en déduire que les modèles « transférés », c'est-à-dire ceux conservant les paramètres de calage d'une année antérieure à la date d'application, contiennent une grande partie de l'information obtenue lors du calage réalisé à la date d'application, et ce malgré l'élargissement du périmètre d'étude.

Si on s'intéresse aux indicateurs relatifs $RATE_S$ et $RATE_R$, on retrouve une particularité récurrente dans les études de transférabilité réalisées pour le modèle gravitaire : il existe certains cas pour lesquels les indicateurs sont inférieurs à 1, ce qui revient à dire que l'ajustement du modèle transféré est meilleur que celui du modèle calibré à la date d'application¹⁷. On peut expliquer ce type de configuration par le fait que les indices de type $SRMSE$ ou $RMSE$ accordent un poids plus important aux erreurs relatives que ne le fait la déviance normalisée (*scaled deviance*), qui est l'indicateur minimisé lors d'une estimation par maximum de vraisemblance. Autrement dit, minimiser la déviance normalisée ne correspond pas tout à fait à minimiser le $RMSE$. Ceci n'est pas sans rappeler la différence entre le calage du modèle gravitaire par la méthode des moindres carrés et par maximum de vraisemblance.

Les valeurs de $RATE_S$ et $RATE_R$ sont proches de 1, ce qui permet de retrouver un résultat donné par OPENSHAW (1979) : la qualité des prévisions avec le modèle gravitaire est très proche de sa capacité descriptive, qui, elle, est moyenne. Dans le cas présent, on a établi que la qualité des calages pour le motif Travail était meilleure que celle pour le motif Loisirs. Il s'en suit que la qualité des prévisions avec les coefficients de 1985 est meilleure pour le motif Travail que pour le motif Loisirs.

Si l'on s'intéresse aux distances, on peut, dans un premier temps, noter que l'application des coefficients de calage de 1985 aux marges et temps de déplacement de 1995 et 2006 donne une très bonne prévision. Pour le motif Travail, on sous-estime légèrement la distance moyenne. Elle n'est sous-estimée que d'environ 8 % pour le motif Travail quand on applique les coefficients de 1985 aux données de 2006 sur un périmètre qui est bien plus large que celui sur lequel les coefficients ont été estimés. Lorsque les périmètres de calage et d'application sont plus proches, la sous-estimation est bien plus faible. Pour le motif Loisirs, on observe à la fois de légères surestimations et sous-estimations.

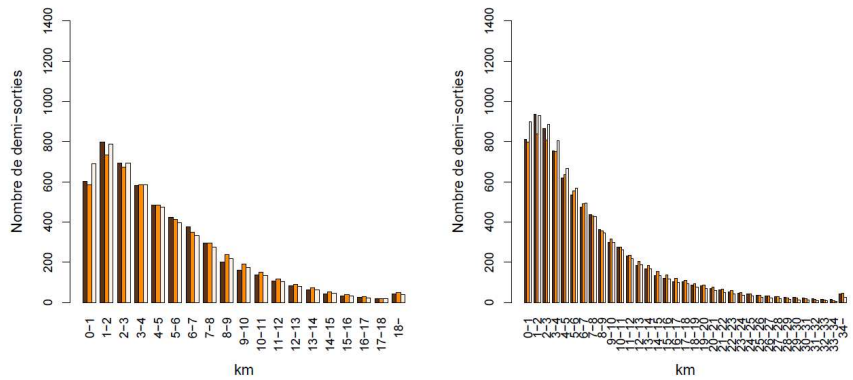
Pour étudier la distribution des distances, on a retenu des classes de 1 km, lorsque l'effectif le permettait. La dernière classe regroupe les demi-sorties pour les distances « très longues » (distances au-delà d'un certain seuil défini

¹⁷ Ce type de comportement « surprenant » peut être observé dans OPENSHAW (1979) et McARTHUR et alii (2011), entre autres.

en fonction de l'effectif). Le χ^2 appliqué aux classes de distance donne des résultats très satisfaisants pour le motif Travail, tant au niveau de la matrice simulée lors du calage que de la matrice simulée à partir des coefficients de 1985. Le calage reproduit bien la distribution des distances pour l'ensemble des années/périmètres. En prévision, on a un résultat similaire. Toutefois, pour la distribution des distances avec les données de 2006 sur le périmètre de l'aire urbaine, l'hypothèse selon laquelle les écarts entre la distribution observée et la distribution modélisée peuvent provenir des aléas liés à l'échantillon (H_0) est acceptée de très peu, avec un niveau de confiance de 95 %. Pour le motif Loisirs, les résultats sont moins satisfaisants : l'hypothèse H_0 est rejetée dans 4 cas sur 5 pour les distances simulées lors du calage et dans 3 cas sur 5 en prévision. Les cas où l'hypothèse est acceptée correspondent à ceux où le contexte de prévision est plus proche de 1985 (données de 1995 pour le périmètre de 1985 et, en prévision, pour le périmètre de 1995). La Figure 3 donne les distributions des distances entre l'origine et la destination des demi-sorties observées, simulées à la date d'application et prédites à partir des coefficients de 1985, pour les motifs Travail et Loisirs. Elle contient, pour les deux motifs de déplacement, le cas d'une prévision à dix ans sur le même périmètre et le cas d'une prévision à vingt ans sur un périmètre élargi. Il s'agit des distributions qui ont permis de calculer les χ^2 correspondants, présentés dans le Tableau 4. Les Figures 3a, 3b et 3c correspondent à des cas où la reproduction des distances est bonne. Les écarts entre le nombre de déplacements observés, modélisés lors du calage et prédits avec les coefficients de 1985 sont faibles. La Figure 3d correspond à un cas où le test du χ^2 a rejeté H_0 . On observe que les distributions sont relativement proches à l'exception de certaines classes de distance pour lesquelles les écarts sont importants. Par exemple, pour la classe de distance allant de 1 à 2 km, les demi-sorties modélisées lors du calage et prédites sont nettement inférieures aux demi-sorties observées. Inversement, pour les classes comprises entre 3 et 8 km, les demi-sorties modélisées lors du calage et prédites sont plus nombreuses que les demi-sorties observées.

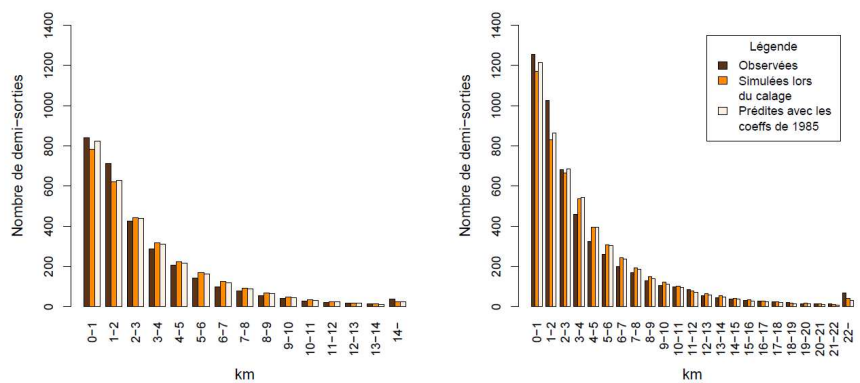
La qualité du calage semble donc aller de pair avec la capacité du modèle gravitaire à reproduire les distances de déplacement. Pour le motif Travail, la fonction de résistance à la TANNER permet d'obtenir une qualité de calage relativement meilleure que pour le motif Loisirs. En prévision, cette différence est retrouvée : les distances de déplacement sont bien reproduites pour le motif Travail mais moins bien pour le motif Loisirs.

Figure 3 : Capacité des matrices simulées, à partir du modèle à la TANNER, à reproduire la distribution des distances observées



(a) Motif Travail, données de 1995 et périmètre de 1985

(b) Motif Travail, données de 2006 et périmètre de l'AUL



(c) Motif Loisirs, données de 1995 et périmètre de 1985

(d) Motif Loisir, données de 2006 et périmètre de l'AUL

6. CONCLUSION

Le modèle gravitaire reste depuis plus de 40 ans le modèle le plus utilisé dans la modélisation transport tant pour des horizons de court/moyen terme (de l'ordre d'une dizaine d'années) que pour des horizons temporels plus lointains. Nous retrouvons les résultats déjà présentés par de nombreux auteurs à savoir que la qualité de l'estimation est souvent jugée moyenne. Ce résultat n'est pas étonnant dans la mesure où l'utilisation de zonages fins pour limiter le poids des intrazonnes conduit à utiliser des matrices observées de piètre qualité pour estimer les modèles. Du fait des effectifs enquêtés, les observations de chacune des cases de la matrice sont connues avec des intervalles de confiance qui sont extrêmement larges. La matrice comporte un très grand nombre de zéros pour lesquels on ne sait pas s'il n'y a effectivement aucun déplacement ou s'il y a seulement peu de déplacements.

En revanche, la matrice estimée par construction ne comportera aucun zéro dès lors que les marges de la matrice n'en comportent pas non plus. Le degré de reproduction des flux entre les différentes O-D peut donc ne pas être très bon et n'est pas forcément un indicateur de la performance prédictive. Dès lors, il ne semble pas raisonnable d'espérer que l'utilisation du modèle gravitaire en prévision puisse conduire à des résultats meilleurs que ceux obtenus lors de l'estimation des paramètres. De fait, l'analyse des résultats a montré que le degré d'erreur dans les prévisions est similaire à celui de l'estimation. Cette étude corrobore donc les résultats de travaux précédents.

C'est la raison pour laquelle la qualité de l'estimation est également mesurée à travers d'autres indicateurs plus agrégés. En particulier, les courbes de distance de déplacements (en temps généralisé) sont fréquemment utilisées car le respect des distributions de distance est souvent considéré comme important au regard des objectifs des modélisations transport. Compte tenu de notre problématique, qui est d'analyser la capacité du modèle gravitaire à simuler l'allongement des distances (euclidiennes) de déplacement, nous avons retenu ce dernier indicateur au lieu du temps généralisé. Les résultats consignés dans le Tableau 4 et la Figure 3 montrent que le modèle gravitaire permet à la fois de bien reproduire une situation de référence et de proposer une projection correcte à un horizon de 20 ans. Les indicateurs que nous avons retenus montrent en effet que la qualité de reproduction de la courbe de distances de déplacements observées en 2006 est aussi bonne lorsque l'on simule la distribution avec les paramètres de calage de 1985 que lorsque l'on utilise les paramètres de 2006. Le résultat le plus intéressant et qui n'est absolument pas intuitif est que ce constat reste valable lorsqu'on considère le périmètre très élargi de l'enquête de 2006 et non plus seulement celui beaucoup plus restreint de l'enquête de 1985. La capacité prédictive reste donc valable même en cas d'élargissement du périmètre d'étude. Ces résultats sont valables pour les deux motifs contraint (Travail) et non contraint (Loisirs). Toutefois, la qualité de l'estimation étant nettement moindre pour le motif Loisirs, la qualité de la prévision des courbes de distribution des distances de déplacement l'est également moins. Il convient toutefois de nuancer ces résultats. Ils correspondent à un exercice de prospective effectué dans un passé connu et donc de prévision parfaite des variables en entrée du modèle. Les bons résultats obtenus en prévision signifient que les variations paramétriques du modèle jouent peu et que la bonne prévision des variables en entrée de celui-ci -émissions et attractions, issues de la phase de génération, et niveaux de service des réseaux- est déterminante.

On peut également en tirer un autre enseignement extrêmement intéressant. L'évolution des émissions et attractions estimées lors du modèle de génération résultent de l'évolution de la localisation des ménages et des activités. L'évolution des temps généralisés résulte de l'évolution de la performance des réseaux de transport. L'allongement des distances (Tableau 1) que l'on a

observé au cours des 20 dernières années résulte donc non pas d'un besoin (subi ou souhaité selon les interprétations) d'aller toujours plus loin ou d'élargir son espace d'activité mais bien de l'évolution des localisations et de la performance des réseaux de transport¹⁸. On aurait donc pu prévoir l'allongement des distances observé ces vingt dernières années, mais à condition de bien prévoir l'évolution des localisations et les flux émis et attirés qui en résultent, ainsi que l'évolution des niveaux de service des réseaux de transport. On retrouve donc tout l'enjeu des modèles d'interaction transport-urbanisme, auxquels le modèle gravitaire de distribution semble pouvoir s'intégrer.

REMERCIEMENTS

Les auteurs tiennent à remercier l'ADEME qui a financé dans le cadre du PREDIT (contrat 08 66 C0084) la recherche qui a servi de support à cet article, la Région Rhône-Alpes qui a financé l'allocation doctorale de Jorge CABRERA DELGADO dans le cadre du Cluster de recherche Transports, Territoires et Sociétés, et les deux référés anonymes pour leurs conseils et remarques qui ont permis d'améliorer cet article.

RÉFÉRENCES

ANAS A. (1983) Discrete choice theory, information theory and the multinomial logit and gravity models. **Transportation Research Part B: Methodological**, Vol. 17, n° 1, pp. 13-23.

AXHAUSEN K.W. (2000). Definition of movement and activity for transport modelling. In D.A. HENSHER, K.J. BUTTON (ed.) **Handbook of Transport Modelling**, Amsterdam, Pergamon, pp. 271-284.

BATTY M., MACKIE S.N.H. (1972) The calibration of gravity, entropy, and related models of spatial interaction. **Environment and Planning**, Vol. 4, n° 2, pp. 205-233.

BATTY M., SIKDAR P.K. (1982a) Spatial aggregation in gravity models. 1. An informationtheoretic framework. **Environment and Planning A**, Vol. 14, n° 3, pp. 377-405.

BATTY M., SIKDAR P.K. (1982b) Spatial aggregation in gravity models. 2. One-dimensional population density models. **Environment and Planning A**, Vol. 14, n° 4, pp. 525-553.

¹⁸ L'interprétation mécanique en termes de causalité qui résulte de la structure de modélisation à quatre étapes est évidemment sujette à discussion sur le sens des causalités tant il est souvent difficile de mettre en évidence des relations de causalité univoque. Elle traduit à notre sens au moins une très forte corrélation entre l'allongement des distances de déplacements et la localisation des ménages et des activités d'une part et la performance de l'offre de transport d'autre part.

- BATTY M., SIKDAR P.K. (1982c) Spatial aggregation in gravity models. 3. Two-dimensional trip distribution and location models. **Environment and Planning A**, Vol. 14, n° 5, pp. 629-658.
- BATTY M., SIKDAR P.K. (1982d) Spatial aggregation in gravity models. 4. Generalisations and large-scale applications. **Environment and Planning A**, Vol. 14, n° 6, pp. 795-822.
- BLACK W. R. (1995) Spatial interaction modeling using artificial neural networks. **Journal of Transport Geography**, Vol. 3, n° 3, pp. 159-166.
- BONNAFOUS A., PUEL H. (1983) **Physionomies de la Ville**. Paris, Éditions ouvrières.
- BONNEL P. (2004) **Prévoir la Demande de Transport**. Paris, Presses de l'École Nationale des Ponts et Chaussées.
- BRIANT A., COMBES P.-Ph., LAFOURCADE M. (2010) Dots to boxes: do the size and shape of geographical units jeopardize economic geography estimations? **Journal of Urban Economics**, Vol. 67, n° 3, pp. 287-302.
- CABANNE I. (2005) **Modélisation à long terme de l'évolution des trafics voyageurs à longue distance en France**. Lyon, Université Lumière Lyon 2, Thèse pour le doctorat de Sciences Économiques option Économie des Transports.
- CAMBRIDGE SYSTEMATICS, INC (2010) **Travel Model Validation and Reasonableness Checking Manual Second Edition**. Washington D.C., Federal Highway Administration, Travel Model Improvement Program.
- CERTU (2002) **Modélisation des déplacements urbains de voyageurs : Guide des Pratiques**. Lyon, Éditions du Certu.
- CERTU (2008). **L'enquête ménages déplacements « standard Certu » : Guide méthodologique**. Lyon, Éditions du Certu.
- COCHRANE R.A. (1975) A possible economic basis for the gravity model. **Journal of Transport Economics and Policy**, Vol. 9, n° 1, pp. 34-49.
- DEMING W.E., STEPHAN F.F. (1940) On least squares adjustment of a sampled frequency table when the expected marginal totals are known. **The Annals of Mathematical Statistics**, Vol. 11, n° 4, pp. 427-444.
- DUFFUS L.N., ALFA A.S., SOLIMAN A.H. (1987) The reliability of using the gravity model for forecasting trip distribution. **Transportation**, Vol. 14, n° 3, pp. 175-192.
- ELMI A.M., BADOE D.A., MILLER E.J. (1999) Transferability analysis of work-trip distribution models. **Transportation Research Record: Journal of the Transportation Research Board**, Vol. 1676, pp. 169-176.

ERLANDER S., STEWART N.F. (1990) **The Gravity Model in Transportation Analysis: Theory and Extensions**. Utrecht, VSP (Topics in Transportation).

EVANS A.W. (1971) The calibration of trip distribution models with exponential or similar cost functions. **Transportation Research**, Vol. 5, n° 1, pp. 15-38.

EVANS S.P., KIRBY H.R. (1974) A three-dimensional Furness procedure for calibrating gravity models. **Transportation Research**, Vol. 8, n° 2, pp. 105-122.

FLYVBJERG B., HOLM M.K.S., BUHL S. (2005) How (in)accurate are demand forecasts in public works projects. **Journal of the American Planning Association**, Vol. 71, n° 2, pp. 131-146.

FOTHERINGHAM A.S., KNUDSEN D.C. (1987) **Goodness of Fit Statistics**. Norwich, Geo Books (Numéro 46 de Catmog).

FOX J., HESS S. (2010) Review of evidence for temporal transferability of mode-destination models. **Transportation Research Record: Journal of the Transportation Research Board**, Vol. 2175, pp. 74-83.

FURNESS K.P. (1965) Time function iteration. **Traffic Engineering and Control**, Vol. 7, n° 7, pp. 458-460.

GODINOT C., BONNEL P. (2008) **Phase d'affectation : mise en forme du réseau routier sur l'aire urbaine de Lyon**. Lyon, LET, Rapport intermédiaire du projet SIMBAD pour le compte de la DRI (MEEDDAT) et de l'ADEME dans le cadre du groupe 11 du PRÉDIT 3, 173 p.

HYMAN G.M. (1969) The calibration of trip distribution models. **Environment and Planning**, Vol. 1, n° 1, pp. 105-112.

KARASMAA N. (2008) **The Transferability of Travel Demand Models: An Analysis of Transfer Methods, Data Quality and Model Estimation**. Saarbrücken, VDM Verlag Dr. Müller Aktiengesellschaft & Co. KG.

KIRBY H.R. (1974) Theoretical requirements for calibrating gravity models. **Transportation Research**, Vol. 8, n° 2, pp. 97-104.

KOPPELMAN F., WILMOT C. (1982) Transferability analysis of disaggregate choice models. **Transportation Research Record: Journal of the Transportation Research Board**, Vol. 895, pp. 18-24.

LET, SEMALY (1997) **Développement d'un modèle stratégique de simulation des déplacements-20 ans de rétrospective à travers les enquêtes déplacements de l'agglomération lyonnaise (1976-1986-1995)**. Lyon.

MCARTHUR D.P., KLEPPE G., THORSEN I., UBØE J. (2011) The spatial transferability of parameters in a gravity model of commuting flows. **Journal of Transport Geography**, Vol. 19, n° 4, pp. 596-605.

MIKKONEN K., LUOMA M. (1999) The parameters of the gravity model are changing-how and why? **Journal of Transport Geography**, Vol. 7, n° 4, pp. 277-283.

MOZOLIN M., THILL J.C., LYNN USERY E. (2000) Trip distribution forecasting with multilayer perceptron neural networks: a critical evaluation. **Transportation Research Part B: Methodological**, Vol. 34, n° 1, pp. 53-73.

NGUYEN-LUONG D. (2011) SIMAURIF : Un modèle dynamique de simulation de l'interaction transport-urbanisation en région Île-de-France. In J.-Ph. ANTONI (éd.) **Modéliser la Ville : Formes Urbaines et Politiques de Transport**. Paris, Economica (avec le soutien du PRÉDIT), pp. 78-118.

NICOLAS J.-P. (2011) **SIMBAD : Un outil pour intégrer le développement durable dans les politiques publiques**. In J.-Ph. ANTONI (éd.) **Modéliser la Ville : Formes Urbaines et Politiques de Transport**. Paris, Economica (avec le soutien du PRÉDIT), pp. 119-164.

NICOLAS J.-P., BONNEL P., CABRERA J., GODINOT C., HOMOCIANU M., ROUTHIER J.-L., TOILIER F., ZUCCARELLO P. (2009) **Simuler les mobilités pour une agglomération durable. Rapport final du projet SIMBAD**. Lyon, LÉT, Rapport de recherche pour le compte de la DRI (MEEDDAT) et de l'ADEME dans le cadre du groupe 11 du PRÉDIT 3. Avec le soutien de l'Agence d'Urbanisme pour le développement de l'agglomération lyonnaise.

OPENSHAW S. (1976) An empirical study of some spatial interaction models. **Environment and Planning A**, Vol. 8, n° 1, pp. 23-41.

OPENSHAW S. (1977) Optimal zoning systems for spatial interaction models. **Environment and Planning A**, Vol. 9, n° 2, pp. 169-184.

OPENSHAW S. (1979) Alternative methods of estimating spatial interaction models and their performance in short-term forecasting. In C.P.A. BARTELS, R.H. KETELLAPPER (ed.) **Exploratory and Explanatory Statistical Analysis of Spatial Data**. Boston, The Hague, London, Martinus Nijhoff Publishing, pp. 201-225.

OPENSHAW S. (1984) **The Modifiable Areal Unit Problem: Concepts and Techniques in Modern Geography**. Norwich, Geo Books.

OPENSHAW S. (1998) Neural network, genetic, and fuzzy logic models of spatial interaction. **Environment and Planning A**, Vol. 30, n° 10, pp. 1857-1872.

ORTÚZAR J.DE D., WILLUMSEN L.G. (2001) *Modelling Transport*. Chichester, John Wiley & Sons.

R DEVELOPMENT CORE TEAM (2011) **R: A Language and Environment for Statistical Computing**. Vienna, R Foundation for Statistical Computing.

SEN A., SMITH T.E. (1995) *Gravity Models of Spatial Interaction Behavior*. Berlin, Springer.

SMITH D.P., HUTCHINSON B.G. (1981) Goodness of fit statistics for trip distribution models. **Transportation Research Part A: General**, Vol. 15, n° 4, pp. 295-303.

SNICKARS F., WEIBULL J.W. (1977) A minimum information principle: theory and practice. **Regional Science and Urban Economics**, Vol. 7, n° 1-2, pp. 137-168.

SOUTHWORTH F. (1979) Spatial structure and parameter disaggregation in trip distribution models. **Regional Studies**, Vol. 13, n° 4, pp. 381-394.

SOUTHWORTH F. (1983) Temporal versus other impacts upon trip distribution model parameter values. **Regional Studies**, Vol. 17, n° 1, pp. 41-47.

SYTRAL (1990) **Comment se déplacent les Lyonnais ? Résultats de l'enquête ménages sur l'agglomération lyonnaise en 1986**. Lyon, 29 p.

SYTRAL, CETE DE LYON, CERTU, INSEE (1996) **Principaux résultats de l'enquête déplacements auprès des ménages de l'agglomération lyonnaise, exploitation standard CERTU, enquête réalisée en 1994-1995**. Lyon, SYTRAL.

SYTRAL, AGENCE D'URBANISME DE LYON (2007) **Enquête ménages déplacements 2006 de l'aire métropolitaine lyonnaise, principaux résultats**. Lyon, SYTRAL, 40 p.

TANNER J. C. (1961) **Factors affecting the amount of travel**. London, Road Research Technical Paper n° 5.

TRB (2007) **Metropolitan Travel Forecasting: Current Practice and Future Direction**. Washington D.C., National Academy of Sciences, Transportation Research Board, Special Report 288.

VOLET P., HUTCHINSON B.G. (1986) Explanatory and forecasting capabilities of trip distribution models. **Canadian Journal of Civil Engineering**, Vol. 13, n° 6, pp. 666-673.

WADDELL P., ULFARSSON G.F., FRANKLIN J.P., LOBB J. (2007) Incorporating land use in metropolitan transportation planning. **Transportation Research Part A: Policy and Practice**, Vol. 41, n° 5, pp. 382-410.

WILLIAMS I. (1976) A comparison of some calibration techniques for doubly constrained models with an exponential cost function. **Transportation Research**, Vol. 10, n° 2, pp. 91-104.

WILSON A. G. (1967) A statistical theory of spatial distribution models. **Transportation Research**, Vol. 1, n° 3, pp. 253-269.