



**HAL**  
open science

## Codification des connecteurs de zones pour les transports en commun

Ouassim Manout

► **To cite this version:**

Ouassim Manout. Codification des connecteurs de zones pour les transports en commun. Architecture, aménagement de l'espace. 2014. dumas-01305732

**HAL Id: dumas-01305732**

**<https://dumas.ccsd.cnrs.fr/dumas-01305732v1>**

Submitted on 21 Apr 2016

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

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



Université Lumière Lyon 2

École Nationale des Travaux Publics de l'État

---

# CODIFICATION DES CONNECTEURS DE ZONES POUR LES TRANSPORTS EN COMMUN

---

Mémoire

Pour l'obtention du diplôme de Master Recherche

Transports Espace Réseaux

Présenté et soutenu publiquement par

Ouassim MANOUT

15 septembre 2014

Membres du Jury :

**Patrick BONNEL** : Ingénieur en Chef des Travaux Publics de l'État, Enseignant-chercheur au Laboratoire d'Économie des Transports, Habilité à diriger des recherches, Chef du Département Transports à l'ENTPE et Tuteur du mémoire au LET.

---

**Louafi BOUZOUINA** : Chargé de Recherche du Ministère de l'Écologie, du Développement durable et de l'Énergie, Cotuteur du mémoire et Expert en accessibilité.

---

**Didier PLAT** : Ingénieur Divisionnaire des Travaux Publics de l'État, Enseignant-chercheur au Laboratoire d'Économie des transports, Habilité à diriger des recherches et Directeur du jury.



## NOTICE ANALYTIQUE

	NOM	PRÉNOM	
<b>AUTEUR</b>	MANOUT	OUASSIM	
<b>TITRE DU MÉMOIRE</b>	Codification des connecteurs de zones pour les transports en commun		
<b>MASTER</b>	Transports, Espace, Réseaux		
<b>ÉTABLISSEMENTS HABILITÉS</b>	École Nationale des Travaux Publics de l'État Université Lumière Lyon 2		
<b>ANNÉE SCOLAIRE</b>	2013/2014	<b>DATE DE SOUTENANCE</b>	15/09/2014
	<b>ORGANISME D'AFFILIATION ET LOCALISATION</b>		<b>NOM PRÉNOM</b>
<b>DIRECTEUR DU MÉMOIRE</b>	Laboratoire d'Économie des Transports. Lyon		Patrick BONNEL
	<b>NOMBRE DE PAGES : 122</b>	<b>NOMBRE D'ANNEXES : 7</b>	<b>NOMBRE DE RÉFÉRENCES : 42</b>
<b>COLLATION</b>			
<b>MOTS CLÉS</b>	Aire de chalandise, BD Topo <sup>®</sup> , carroyage Insee, centroïde de zone, connecteur de centroïde, distance de diffusion, distance de marche à pied, distance de rabattement, transport en commun, population localisé, SIRENE <sup>®</sup> .		
<b>TERMES GÉOGRAPHIQUES</b>	Agrégation, aire de chalandise, aire de service, biais d'agrégation, centroïdes de zones, connecteurs de zones, distance à vol d'oiseau, distance rectilinéaire, IRIS, MAUP, unités zonales modifiables, SIG, zonage, zones.		
<b>RÉSUMÉ</b>	La plupart des méthodologies de codification des connecteurs reposent sur la position géographique des centroïdes de zones. Cette pratique est, certes, simple à implémenter, mais elle semble être inadaptée aux connecteurs des transports en commun, d'autant plus que son impact sur les résultats de modélisation n'est pas négligeable. Le présent mémoire a pour objectif l'élaboration d'une nouvelle méthodologie de codification plus adaptée aux transports collectifs. Cette nouvelle codification doit s'affranchir du biais d'agrégation imposé par le zonage et rendre compte, à un niveau désagrégé, des pratiques de marche à pied vers et depuis les arrêts.		
<b>TITLE ABSTRACT</b>	Defining centroid connectors for the transit system Defining centroid connectors has long been based on the centroid position. It's a simple and quick way that might be inappropriate for the transit system connectors. Some results drawn from a transit assignment model are shown to be sensitive to the connectors' representation. Therefore, here we suggest a new centroid connectors' definition based on a disaggregated approach that seems to be more suitable for the transit system.		



# REMERCIEMENTS

---

Je tiens à remercier Patrick BONNEL pour la confiance qu'il placée en moi en me confiant ce sujet de mémoire. Son encadrement et ses conseils infailibles ont éclairé le cheminement de ce travail.

À Louafi BOUZOUINA j'exprime ma reconnaissance. Sa disponibilité, ses conseils, ses propositions et surtout sa bienveillance ont permis à ce travail d'avancer.

Je tiens à remercier Jorge CABRERA DELGADO qui n'a ménagé aucun effort pour me venir en aide. Je le remercie aussi en tant que collègue de bureau pour sa bonne humeur.

Je remercie Didier PLAT ainsi que toute l'équipe enseignante du master 'Transports Espace Réseaux'.

L'aboutissement de ce travail tient à la générosité d'esprit des diverses communautés actives dans le développement des plateformes libres. En particulier, j'exprime ma gratitude aux développeurs de Qgis et d'OpenStreetMap.

Merci à toute l'équipe du Laboratoire d'Économie des Transports pour son accueil et sa bienveillance.

Enfin, je remercie Madame Yousra TIMOUNAY pour son dévouement et sa patience.

# SOMMAIRE

---

Remerciements .....	1
Sommaire .....	2
Introduction .....	5
<b>Chapitre 1</b> .....	<b>6</b>
1 Développement de la problématique.....	6
1.1 Contexte.....	6
1.2 Constat .....	8
1.3 Objectifs.....	9
<b>Chapitre 2</b> .....	<b>11</b>
2 La méthodologie de codification des connecteurs selon SIMBAD .....	11
2.1 Présentation générale de SIMBAD.....	11
2.2 Méthodologie de codification des connecteurs dans SIMBAD.....	12
2.2.1 Description du zonage .....	12
2.2.2 Description du réseau.....	14
2.2.3 Description des connecteurs des transports collectifs.....	14
2.2.3.1 Présentation.....	14
2.2.3.2 Méthodologie.....	15
2.2.4 Limites de la méthodologie actuelle.....	16
2.2.4.1 hypothèses.....	16
2.2.4.2 Pertinence .....	18
2.2.4.3 Effets sur les charges.....	20
2.3 Conclusion.....	24
<b>Chapitre 3</b> .....	<b>25</b>
3 Nouvelle méthodologie de codification des connecteurs .....	25
3.1 Choix des arrêts à connecter .....	25
3.1.1 Présentation.....	25
3.1.2 Accessibilité des points de desserte.....	26
3.1.3 Agrégation des connecteurs.....	28
3.2 Formalisation de la génération des connecteurs.....	29
3.3 Formalisation de la méthodologie de calcul de la longueur des connecteurs ...	31
3.3.1 Formalisation générale .....	31
3.3.2 Adaptation de la méthodologie de codification des connecteurs aux données disponibles.....	33
3.3.2.1 Choix du modèle de régression .....	33

3.3.2.2	Discussion autour de la relation de régression .....	36
3.3.3	Prise en compte de l'effet distance .....	36
3.3.3.1	La fonction 'Tout ou rien' .....	37
3.3.3.2	La fonction 'Dégressive' .....	38
3.3.3.3	La fonction 'Combinée' .....	40
3.3.4	Distance d'accessibilité .....	40
3.3.5	Algorithme simplifié d'implémentation.....	43
3.3.6	Conclusion.....	43
<b>Chapitre 4</b>	.....	<b>44</b>
4	Données indispensables à l'implémentation de la nouvelle méthodologie.....	44
4.1	Présentation des données d'entrée de la population.....	44
4.1.1	Présentation de la Base Carroyée.....	45
4.1.2	Présentation de la BD TOPO® .....	47
4.1.3	Présentation des données de recensement de la population de 2010 .....	52
4.2	Présentation des données d'emploi .....	52
4.3	Exploitation des données d'entrée .....	55
4.3.1	Exploitation des données carroyées .....	55
4.3.1.1	Répartition ponctuelle aux centroïdes des carreaux.....	55
4.3.1.2	Répartition proportionnelle à la surface des carreaux.....	60
4.3.2	Exploitation des données carroyées et de la BD TOPO® .....	65
4.3.2.1	Répartition proportionnelle au volume bâti.....	65
4.3.2.1.1	Calcul du volume bâti.....	66
4.3.2.2	Variante 1 : découpage des bâtiments en portions.....	67
4.3.2.3	Variante 2 : répartition des bâtiments par la méthode des centroïdes.....	71
4.4	Comparaison des méthodes d'estimation de la population.....	74
4.5	Calage de la population estimée sur la population recensée.....	78
4.6	Conclusion.....	78
<b>Chapitre 5</b>	.....	<b>79</b>
5	Mise en application de la méthodologie de codification des connecteurs.....	79
5.1	Premier cas élémentaire : Un iris, une ligne de transport et un seul point de desserte .....	79
5.1.1	Calcul par la fonction 'Tout ou rien'.....	80
5.1.2	Calcul par la fonction 'Dégressive'.....	82
5.1.3	Calcul par la fonction 'Combinée' .....	82
5.1.4	Conclusion.....	83
5.2	Second cas élémentaire : Un iris, une ligne de transport et deux points de desserte .....	83

5.2.1	Calcul par la fonction ‘tout ou rien’ .....	86
5.2.2	Calcul par la fonction ‘degressive’ .....	86
5.2.3	Calcul par la fonction ‘combinée’ .....	86
5.2.4	Conclusion.....	87
5.3	Cas général.....	87
5.3.1	Implémentation de l’algorithme.....	87
5.3.2	Conclusion.....	91
	Conclusion.....	92
	<b>Annexes.....</b>	<b>94</b>
	<b>Bibliographie.....</b>	<b>111</b>
	<b>Liste des figures.....</b>	<b>114</b>
	<b>Liste des graphiques .....</b>	<b>116</b>
	<b>Liste des tables.....</b>	<b>118</b>

# INTRODUCTION

---

Un connecteur est un lien fictif qui traduit les conditions locales d'accès et de sortie du réseau de transport modélisé. Il permet de relier la demande en déplacement des zones à l'offre de transport. Ainsi, sa modélisation doit faire l'objet d'une attention particulière. Notamment pour les transports en commun dont la qualité de l'offre est fortement tributaire de leur accessibilité. Toutefois, peu d'études se sont intéressées à ce sujet et les pratiques en matière de codification des connecteurs demeurent, pour la plupart, *simplistes*.

En effet, la plupart des méthodologies de définition des connecteurs reposent sur la notion du centroïde de zone et de sa position géographique. Le centroïde de zone est un point représentatif de toute la zone. Il est supposé être la localisation de l'*individu moyen* qui résume l'ensemble des informations désagrégées de la zone. Les pratiques reposant, dans leur codification des connecteurs, sur ce point modélisent un connecteur comme étant le lien qui relie le centroïde au réseau de transport correspondant. Si cette approximation est simple à mettre en œuvre, elle n'est pas sans impact sur les sorties de la chaîne de modélisation, en particulier sur les résultats du système des transports collectifs.

Conscients de cette problématique, nous proposons, dans le présent mémoire, d'étudier la codification des connecteurs du réseau des transports collectifs, d'évaluer leurs impacts et de proposer, en conséquence, une nouvelle approche plus adaptée au cas de ce système de transport.

Pour ce faire, nous allons commencer par un développement de la problématique dans le premier chapitre. Nous enchaînerons par la suite sur la présentation et l'évaluation de l'incidence de la méthodologie actuellement utilisée par SIMBAD pour modéliser ses connecteurs. Le troisième chapitre sera réservé à la définition de la nouvelle approche de codification et le quatrième à la construction d'une nouvelle base de données indispensable à l'implémentation de cette dernière. Le cinquième chapitre détaille quelques exemples d'application et dresse un comparatif entre les deux méthodologies.

# CHAPITRE 1

---

## 1 DÉVELOPPEMENT DE LA PROBLÉMATIQUE

### 1.1 CONTEXTE

L'implémentation d'un modèle de prévision de trafic nécessite, d'un côté, l'intégration de données relatives à la demande de transport et de l'autre côté, de données détaillant l'offre de transport. En général, les données de demande sont déduites des caractéristiques démographiques et socioéconomiques des ménages - population, emplois, âge, catégories socio-professionnelles, etc. -. L'offre de transport représente, à un niveau de détails précis, les réseaux et les services de transport ainsi que leurs caractéristiques intrinsèques.

L'opération de collecte, de traitement et de prévision de ces données est souvent consommatrice de temps et d'argent, et peut être assujettie à une législation restrictive visant la protection de la vie privée (Bonnell, 2004; Martinez et al., 2010). Ceci dit et en fonction des objectifs de la modélisation, il n'est pas toujours nécessaire de décrire en détail l'offre et la demande de transport, d'autant plus que ces efforts sont proportionnels au niveau de finesse recherché. De ce fait, une description agrégée peut être suffisante. Ainsi, la plupart des modèles de prévision de trafic reposent sur une discrétisation de l'espace continu en zones dites Traffic Analysis Zones ou TAZ (Binetti and Ciani, 2002; Bonnell, 2004; Ortúzar S, Juan D and Willumsen, 2011). Seuls les modèles totalement désagrégés s'affranchissent de cette agrégation (Chapleau, 1992).

L'agrégation spatiale consiste en la discrétisation de l'espace continu en zones. L'établissement de ce zonage doit répondre à plusieurs critères qu'on peut résumer comme suit :

- Construire des zones homogènes selon les caractéristiques principales de la population ou des emplois – motorisation, CSP, revenu, usage du sol, etc. - (Baass, 1981; Bonnell, 2004; Ding, 1998, 1994; Martínez et al., 2007, 2009; Openshaw, 1978; Ortúzar S, Juan D and Willumsen, 2011).
- Concevoir un découpage compatible avec les systèmes de zonage existant et les unités de diffusion des données - IRIS, arrondissement, commune, etc. - (Baass, 1981; Bonnell, 2004; Ortúzar S, Juan D and Willumsen, 2011).
- Concevoir un découpage respectant les contraintes géographiques et les barrières artificielles – fleuve, voie rapide, voie ferrée, etc. - (Baass, 1981; Bonnell, 2004; Ortúzar S, Juan D and Willumsen, 2011).
- Maximiser les interactions entre zones ou minimiser les déplacements intra-zonaux (Baass, 1981; Martínez et al., 2007, 2009).

- Réaliser un découpage régulier des zones, c'est-à-dire éviter les formes irrégulières (Martínez et al., 2007, 2009; Ortúzar S, Juan D and Willumsen, 2011).
- Construire un zonage exclusif, autrement dit, éviter les chevauchements de zones (Baass, 1981; Ding, 1998, 1994; Martínez et al., 2007, 2009).
- Construire des zones ayant un pouvoir d'attraction et d'émission comparable (Ding, 1998, 1994; Martínez et al., 2007; Ortúzar S, Juan D and Willumsen, 2011).

Il semble donc difficile de respecter tous ces critères qui sont parfois contradictoires. D'où la difficulté accompagnant la mise en place d'un zonage. Pour cette raison quelques méthodes automatisées d'agrégation ont vu le jour (Baass, 1981; Martínez et al., 2010; Martínez et al., 2009). Ces algorithmes ont comme point commun la minimisation ou la maximisation d'une '*fonction objectif*', comme par exemple le taux de déplacements intra-zonaux ou l'homogénéité des caractéristiques des individus peuplant chaque zone. Toutefois les pratiques demeurent disparates et surtout guidées par l'expérience des modélisateurs et les objectifs propres à chaque étude.

Les zones (ou TAZ) résument les informations contenues à l'intérieur de leurs limites. Ces informations sont attribuées aux centres, dits aussi centroïdes de zones. Ces derniers n'ont pas d'équivalents sur le terrain et sont perçus comme des points flottant à l'intérieur des zones (Ortúzar S, Juan D and Willumsen, 2011). Dans ce cas, l'unité d'analyse n'est plus l'individu localisé mais l'individu moyen ou l'individu moyen catégorisé<sup>1</sup>. Cette simplification facilite, certes, l'implémentation et la validation des modèles, mais engendre une perte de l'information et induit un biais d'erreur dit biais d'agrégation (Ortúzar S, Juan D and Willumsen, 2011). Ce biais est d'autant plus prépondérant que le zonage est grossier (Bovy and Jansen, 1983; Chang et al., 2002; Horowitz, 2001). Cette erreur fait partie d'une problématique plus générale et commune aux modèles spatiaux agrégés dite la problématique des unités zonales modifiables (ou MAUP pour Modifiable Areal Unit Problem). La MAUP fait référence à la sensibilité des résultats de sortie d'un modèle au système de zonage adopté.

Dans ce sens, une des manifestations problématiques de l'agrégation zonale est la définition des connecteurs. Un connecteur est un lien fictif reliant la zone au réseau de transport (Bonnell, 2004; Friedrich and Galster, 2009; Ortúzar S, Juan D and Willumsen, 2011). Normalement, un connecteur doit traduire les conditions locales d'accès et de sortie du réseau modélisé. Suite à l'agrégation, ces conditions locales ne sont plus connues et un connecteur traduit désormais les conditions moyennes – Ou de l'individu moyen – d'entrée ou de sortie du réseau.

---

<sup>1</sup> Il est possible d'agréger les données de zones selon des catégories, par exemple, ménages motorisés, revenu, etc.

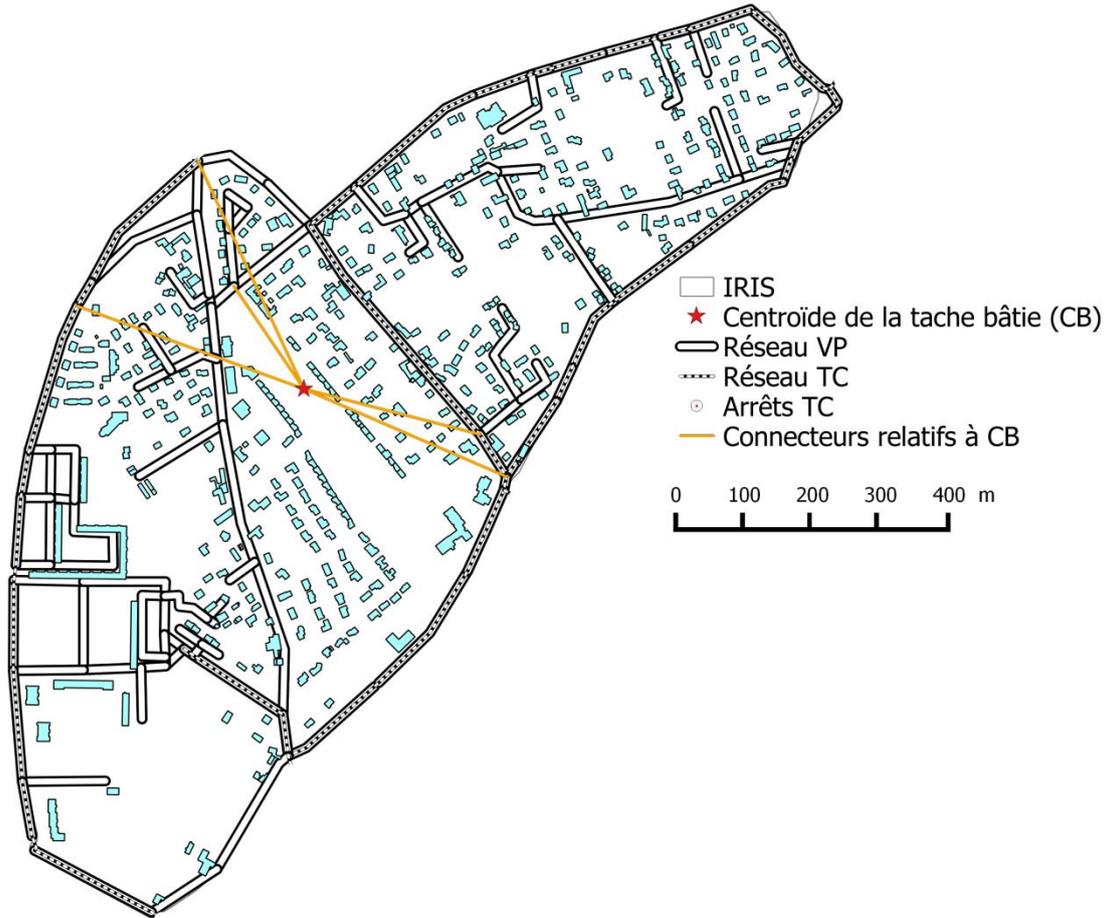


FIGURE 1: EXEMPLE DE DÉFINITION DE CONNECTEURS VP ET TC SUR LA BASE DE LA POSITION GÉOGRAPHIQUE DU CENTROÏDE DE ZONE (SOURCE SIMBAD)

## 1.2 CONSTAT

Parmi les attributs des connecteurs, le temps de parcours et les nœuds de connexion sont les plus importants. Le coût d'un connecteur est calculé à partir de sa longueur et de sa vitesse de parcours. Dans la plupart des études de modélisation, la vitesse de parcours est constante et est fixée en fonction de la nature du connecteur. Autrement dit, elle est indépendante de la charge du connecteur. D'ailleurs, ce dernier est souvent supposé avoir une capacité infinie ou assez élevée, l'épargnant ainsi des effets de la congestion. La vitesse est de l'ordre de 30km/h pour les connecteurs du système de transport individuel et de 4km/h pour le système des transports en commun. Ces vitesses traduisent les conditions locales d'accès et de sortie des réseaux correspondants, et sont donc, plus ou moins, connues des modélisateurs. Le choix de la méthodologie de calcul de la longueur et des nœuds de connexion est par contre moins évident et ses conséquences sont méconnues et mal maîtrisées.

Au moment où des sujets comme l'affectation ou le choix modal ont fait l'objet d'une abondante littérature, peu d'auteurs se sont intéressés à la codification des connecteurs (Friedrich and Galster, 2009; Sean Qian and Zhang, 2012). Or l'incidence de ces derniers sur la chaîne de modélisation n'est pas négligeable (Chang et al., 2002; Ortúzar S, Juan D and Willumsen, 2011; Sean Qian and Zhang, 2012), en particulier sur les résultats de

l'affectation. En conséquence, quelques guides de 'bonnes pratiques', issus des expériences des autorités de transports, sont venus combler ce vide méthodologique et partager des pratiques qui, pour la plupart, ne sont pas soutenues par des travaux de recherche (Cambridge Systematics, Inc and AECOM Consult, 2007; WATS, 2008). On peut résumer, comme suit, ces pratiques :

- Le coût associé aux connecteurs – temps de parcours - doit être proche du coût réel supporté par les usagers.
- Les connecteurs doivent prendre en considération les effets de coupures – barrières naturelles et artificielles -.
- Les connecteurs doivent être reliés aux tronçons ayant le plus bas niveau dans la hiérarchie du réseau (réseau local, par exemple).
- Le nombre de connecteurs doit être suffisant pour éviter le phénomène de congestion fictive<sup>2</sup>.
- Les nœuds de connexion doivent éviter les intersections de routes, des entrées de voies rapides, etc.
- Le nombre de connecteurs doit être limité pour réduire les temps de calcul.

Ces règles sont pour la plupart destinées aux connecteurs du système de transport individuel. Les connecteurs des transports en commun doivent répondre en partie à ces critères, auxquels s'ajoutent :

- Toutes les lignes, considérées accessibles depuis une zone, doivent être connectées.
- Dans le cas d'une ligne avec plusieurs arrêts accessibles depuis une même zone, un seul connecteur '*moyen*' peut traduire les conditions moyennes d'accès aux différents arrêts. La répartition de la demande entre ces arrêts n'est ni indispensable ni maîtrisée.
- Le problème de la congestion fictive n'apparaît pas dans le cas des transports en commun.

L'application de ces règles n'est pas évidente et est souvent portée par l'expérience du modélisateur. De ce fait, peu de méthodologies '*scientifiques*' ont essayé d'éclairer ces pratiques, notamment pour le système des transports en commun.

### 1.3 OBJECTIFS

À notre connaissance aucune étude n'a abordé l'impact des connecteurs sur les résultats de l'affectation pour le système des transports collectifs. Une étude menée par Qian et Zhang (Sean Qian and Zhang, 2012) sur le système de transport individuel (voiture particulière) a démontré l'impact significatif de la codification des connecteurs sur les résultats de l'affectation. Les charges du réseau changent en fonction des connecteurs utilisés, engendrant ainsi une sous-estimation ou une surestimation des trafics. En effet, une expérience menée sur une voie rapide en Californie - State Road 41 - a montré que ses charges modélisées varient de presque 4000véh/h à 7000véh/h en choisissant aléatoirement les connecteurs. La charge de référence – considérée comme observée – est

---

<sup>2</sup> Un connecteur a, par construction, une capacité élevée voire infinie. La congestion fictive apparaît quand un connecteur véhicule trop de charges de déplacement vers une infrastructure dont la capacité réelle est limitée.

avoisine les 5000véh/h. Cet intervalle d'erreur se réduit en augmentant le nombre de connecteurs mais reste, toutefois, significatif. Il est de même pour les temps de parcours. Une des conclusions de cette étude est que les résultats de l'affectation routière varient *significativement* en fonction de la configuration des connecteurs (Sean Qian and Zhang, 2012).

Dans une autre étude, Chang et Khatib (Chang et al., 2002) ont souligné l'impact de la codification des centroïdes de zones et par conséquent des connecteurs sur la charge du réseau routier. Dans ce sens, ils ont comparé quatre configurations de codification des centroïdes : centre géographique de la zone (CG), centre de la commune principale de la zone (CC), centroïde pondéré par la population (CP) et centroïde pondéré par la densité des ménages (CM). Le ratio entre la charge modélisée et la charge observée a été analysé, entre autres, à la lumière de ces différentes positions des centroïdes. L'analyse montre une variation de ce ratio en fonction des configurations. Pour un même zonage, les codifications CM, CP et CC aboutissent souvent à un ratio plus satisfaisant<sup>3</sup> que la configuration CG.

Il semble que l'incidence de la définition des connecteurs sur les sorties de la chaîne de modélisation ne sont pas négligeables. En particulier, la modélisation des transports en commun semble être plus sensible à cette définition que la modélisation des transports individuels (Bonnell, 2004). En effet, si l'accès au réseau de transport individuel est motorisé et par conséquent plus rapide, l'accès aux arrêts de transport en commun est souvent accompli en marche à pied et est, par conséquent, non négligeable dans la chaîne de déplacement, d'autant plus qu'il est mal perçu par les usagers (Bonnell, 2004). Dans ce sens, une étude menée sur le réseau SIMBAD montre que la charge des lignes de transports collectifs est sensible à la définition des connecteurs (voir § 2.2.4.3).

Enfin, nous soulignons ici que l'incidence des connecteurs sur la chaîne de modélisation n'est pas totalement indépendante de la problématique des unités zonales modifiables (MAUP). En effet, il est reconnu qu'un zonage fin engendre des connecteurs de faible longueur avec un moindre impact sur les résultats finaux (Chang et al., 2002; Horowitz, 2001). Dans ce mémoire, nous nous limitons à l'étude de l'incidence brute des connecteurs sans isoler leurs effets de l'impact de la MAUP. Pour ce faire, nous utilisons un même système de zonage dans toutes les comparaisons procédées.

---

<sup>3</sup> Selon les auteurs, le ratio entre charges estimée et observée est satisfaisant s'il est entre 0.8 et 1.2.

# CHAPITRE 2

## 2 LA MÉTHODOLOGIE DE CODIFICATION DES CONNECTEURS SELON SIMBAD

### 2.1 PRÉSENTATION GÉNÉRALE DE SIMBAD

Simuler les MoBilités pour une Agglomération Durable - dit par la suite SIMBAD - est un projet de recherche mené par le Laboratoire d'Économie des Transports en collaboration avec l'Agence d'Urbanisme de Lyon dans le cadre du PREDIT 3. Il a comme objectif l'évaluation, à la lumière du développement durable, des politiques de transport et d'urbanisme au niveau de l'agglomération de Lyon (Figure 2). En pratique, il s'agit d'un modèle prospectif à l'horizon de 2025, intégrant les interactions entre transport et urbanisme. Dans ce sens il s'inscrit dans la catégorie des modèles LUTI - Land Use Transport Interaction -. Par construction, SIMBAD est plus un outil d'aide à la réflexion qu'un outil d'évaluation (Nicolas et al., 2009). Dans ce sens il s'inscrit dans la catégorie des modèles stratégiques.

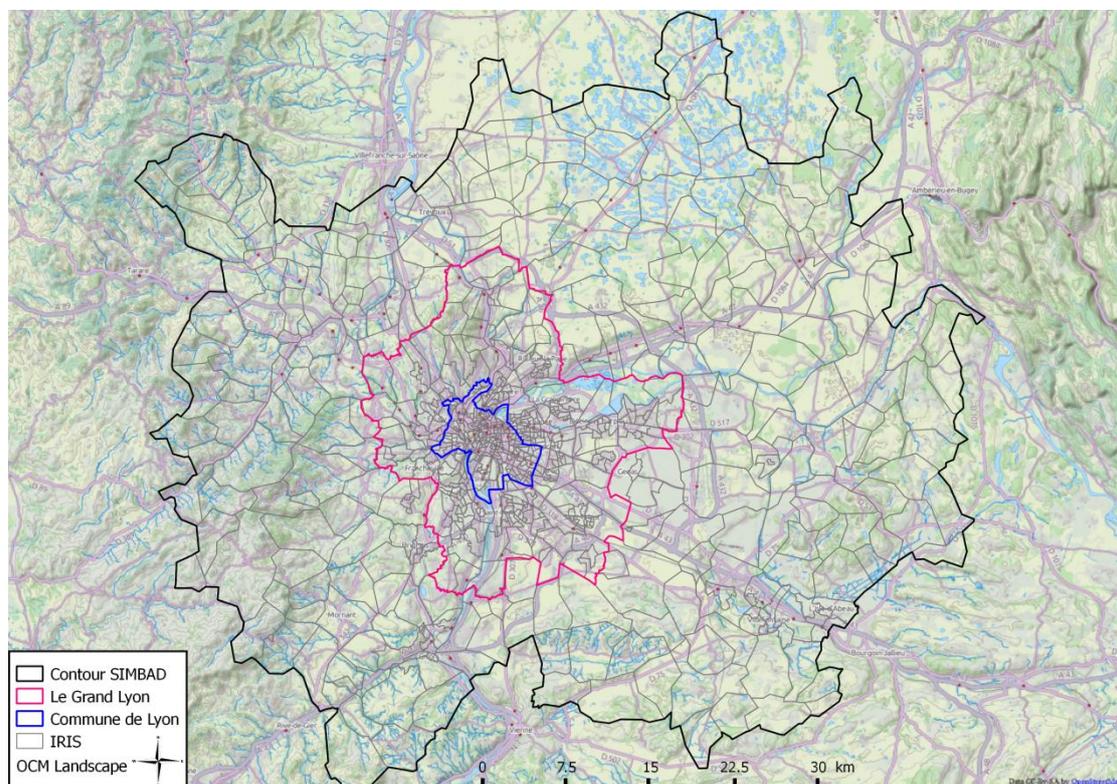


FIGURE 2 : PÉRIMÈTRE SIMBAD (FOND DE CARTE OPENSTREETMAP)

SIMBAD intègre plusieurs modules en interaction dont deux relatifs au transport. Pour prendre en considération ces différents modules et leurs interactions, SIMBAD adopte une structure systémique en boucle rétroactive (Figure 3).

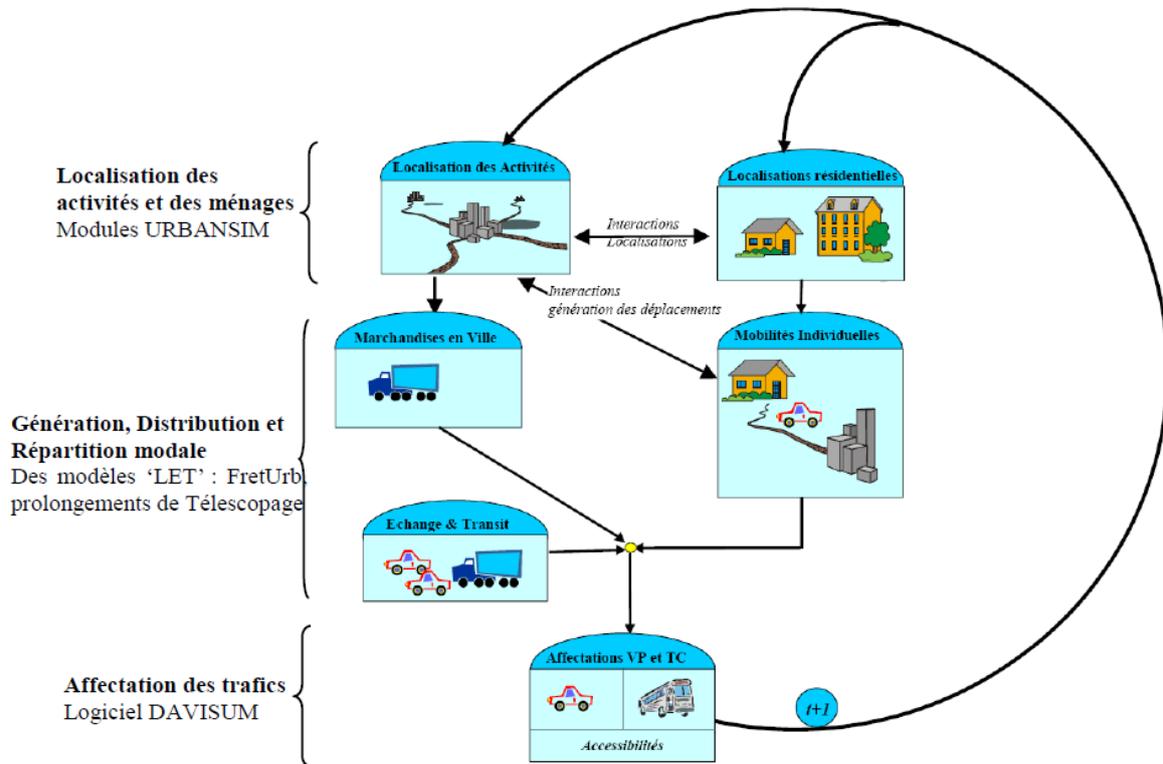


FIGURE 3 : AGENCEMENT DES MODULES INTÉGRÉS À SIMBAD (NICOLAS ET AL., 2009)

Nous nous intéressons ici aux modules transport et en particulier à la méthodologie suivie pour la codification des connecteurs TC.

## 2.2 MÉTHODOLOGIE DE CODIFICATION DES CONNECTEURS DANS SIMBAD

### 2.2.1 DESCRIPTION DU ZONAGE

Le zonage adopté par SIMBAD est le découpage en IRIS - Ilots Regroupés pour l'Information Statistique -. Il s'agit d'un découpage proposé par l'Insee pour diffuser les données démographiques infra-communales. Le périmètre SIMBAD regroupe 777 IRIS – ajouter à cela 10 zones externes modélisant les trafics d'échange -. Ce zonage a la particularité d'être fin au niveau de la commune de Lyon et ses alentours puis grossier ailleurs (Figure 4). Il est établi sur la base de critères géographiques et statistiques avec un effort d'homogénéité de l'habitat (INSEE, 2014). Ce découpage représente la maille la plus fine de diffusion des données démographiques de la population municipale – il existe un carroyage plus fin décrivant la population fiscale -.

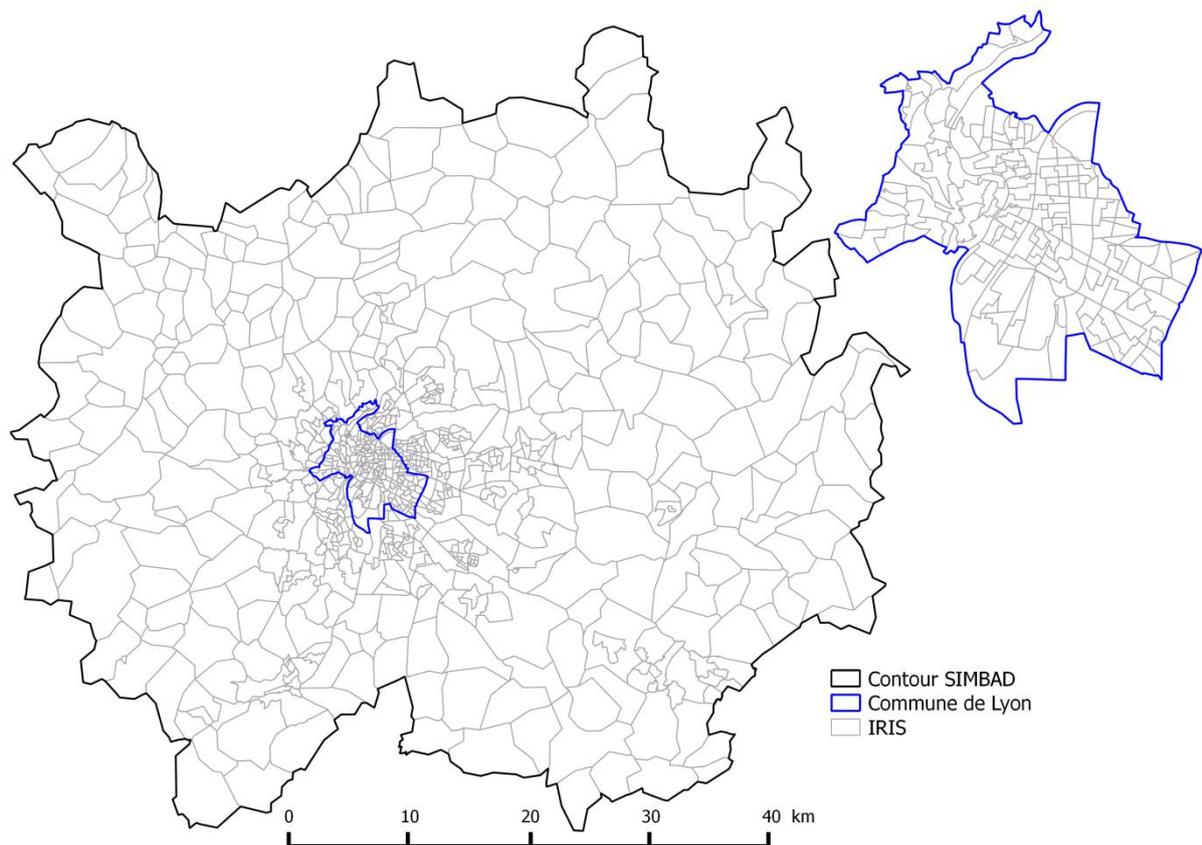


FIGURE 4 : DÉCOUPAGE EN IRIS DE L'AIRES SIMBAD, FOCUS SUR LA COMMUNE DE LYON

Ceci dit, on souligne que le zonage IRIS de l'aire SIMBAD enfreint quelques règles de 'bonnes pratiques' de découpage (voir § 1.1). En effet, il y a des zones qui se chevauchent complètement pour former des îles (voir cas 2 et 4 Figure 5). D'autres zones épousent des formes irrégulières (voir cas 5 Figure 5). Enfin, on a relevé le cas de deux IRIS discontinus et formés de deux zones disjointes ou quasi disjointes (voir cas 1 et 3 Figure 5). Ces irrégularités sont très préjudiciables à la génération des centroïdes de zones et dans certains cas des connecteurs si la définition de ces derniers dépend de la position du centroïde. Toutefois, nous nous contentons, ici, de signaler ce problème sans le traiter.

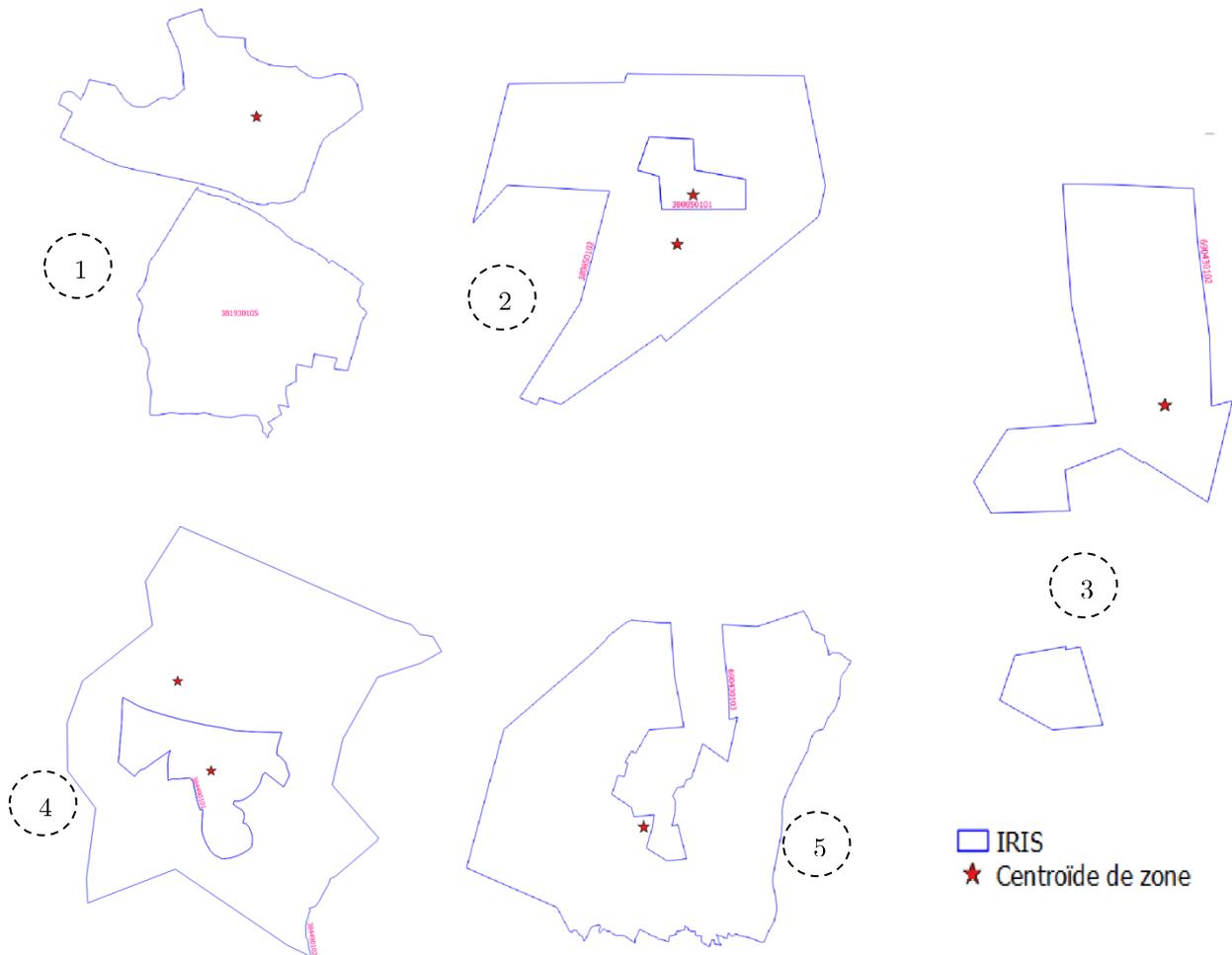


FIGURE 5 : CAS DE ZONES ISSUES DE SIMBAD NON CONFORMES AUX RÉGLES DE DÉCOUPAGE

## 2.2.2 DESCRIPTION DU RÉSEAU

L'offre de transport SIMBAD est décrite d'une façon détaillée. Ce choix a été dicté par la volonté de traiter simplement et automatiquement le maximum de données disponibles. Ainsi le réseau viaire et les services de transport en commun sont représentés finement. Toutes les lignes et tous les arrêts sont géolocalisés et leurs caractéristiques renseignées (profils des temps de parcours et services offerts, etc.). Toutefois, cette description fine de l'offre de transport ne s'apprête pas à une analyse aussi fine des résultats de sortie. En effet, l'objectif de SIMBAD n'est pas de simuler l'incidence locale d'une politique de transport ou d'urbanisme sur les trafics (Nicolas et al., 2009).

## 2.2.3 DESCRIPTION DES CONNECTEURS DES TRANSPORTS COLLECTIFS

### 2.2.3.1 PRÉSENTATION

Peu de recherches se sont penchées sur la problématique de définition des connecteurs pour les transports collectifs et leurs impacts sur les résultats de modélisation. La plupart

des travaux de modélisation, y compris SIMBAD, définissent les connecteurs sur la base des centroïdes de zones. Pour ce faire, une position géographique est attribuée aux centroïdes, à partir de laquelle une distance est calculée (Figure 1). Cette approche simplificatrice peut constituer une entrée en matière dans la codification des connecteurs, mais elle biaise les résultats de sortie (Bonnell, 2004; Ortúzar S, Juan D and Willumsen, 2011).

SIMBAD définit les connecteurs en fonction de la position géographique des centroïdes de zones. Cette position a été fixée, dans un premier lieu, au centre géographique de chaque zone puis au barycentre pondéré de la tache urbaine<sup>4</sup> (Godinot, 2008). À la suite du choix de l'arrêt à connecter, une ligne directe relie le centroïde à cet arrêt. Cette ligne porte le nom de connecteur de centroïde.

### 2.2.3.2 MÉTHODOLOGIE

Pour détailler la procédure de génération des connecteurs sous SIMBAD, nous empruntons largement à Nicolas et al (Nicolas et al., 2009).

Cette méthodologie, dite par la suite méthodologie SIMBAD, procède par étapes successives de la manière suivante :

1. Lors de la première étape, nous connectons les centroïdes aux nœuds de réseau comportant un arrêt de transports collectifs distant de moins de 500 mètres du centroïde. Nous avons choisi une distance rectilinéaire<sup>5</sup> qui permet de traduire correctement les distances réelles de marche à pied d'un tissu urbain. La limite de 500m a été fixée en fonction du rayon de l'aire de chalandise moyenne des arrêts de transports collectifs - généralement compris entre 300 et 700 mètres selon la nature du système de transports collectifs et la densité du réseau -. Lors de cette étape, nous avons limité le nombre de connecteurs possibles par centroïde à 8 afin de ne pas surcharger la description du réseau.
2. La seconde étape a consisté à renouveler la même opération pour tous les nœuds de réseau non connectés, mais en choisissant une distance de 1 km, sûrement plus adaptée à la situation péri-urbaine. À l'issue de cette étape, 220 nœuds de réseau et 244 IRIS ne sont pas connectés. L'analyse des nœuds de réseau montre qu'il s'agit d'arrêts peu importants et que les arrêts proches sont déjà connectés au sein des IRIS concernés. Si l'absence de connexion pour les nœuds de réseau n'est pas problématique, son absence pour les IRIS l'est, car elle empêche l'accès au réseau de transports collectifs.
3. Nous renouvelons donc l'opération pour les IRIS non connectés en portant successivement la distance à 1,5, 2, 3, 4, 5, 6 et enfin 7 km. Il ne reste plus que deux IRIS et 198 nœuds non connectés.

---

<sup>4</sup> Selon l'Insee, la tache urbaine correspond aux ensembles de plus de cinq bâtiments, ayant chacun une emprise au sol supérieure à 30 m<sup>2</sup> entourés d'une zone tampon de 20m.

<sup>5</sup> Il s'agit d'une distance rectilinéaire (DR) calculée à partir de la distance à vol d'oiseau (DVO) par la formule :  $DR = DVO * (1.1 + 0.3 * \text{Exp}(-DVO/20))$ . Cette distance rend compte des détours imposés par le tissu urbain (Gallez cité dans Godinot, 2008)

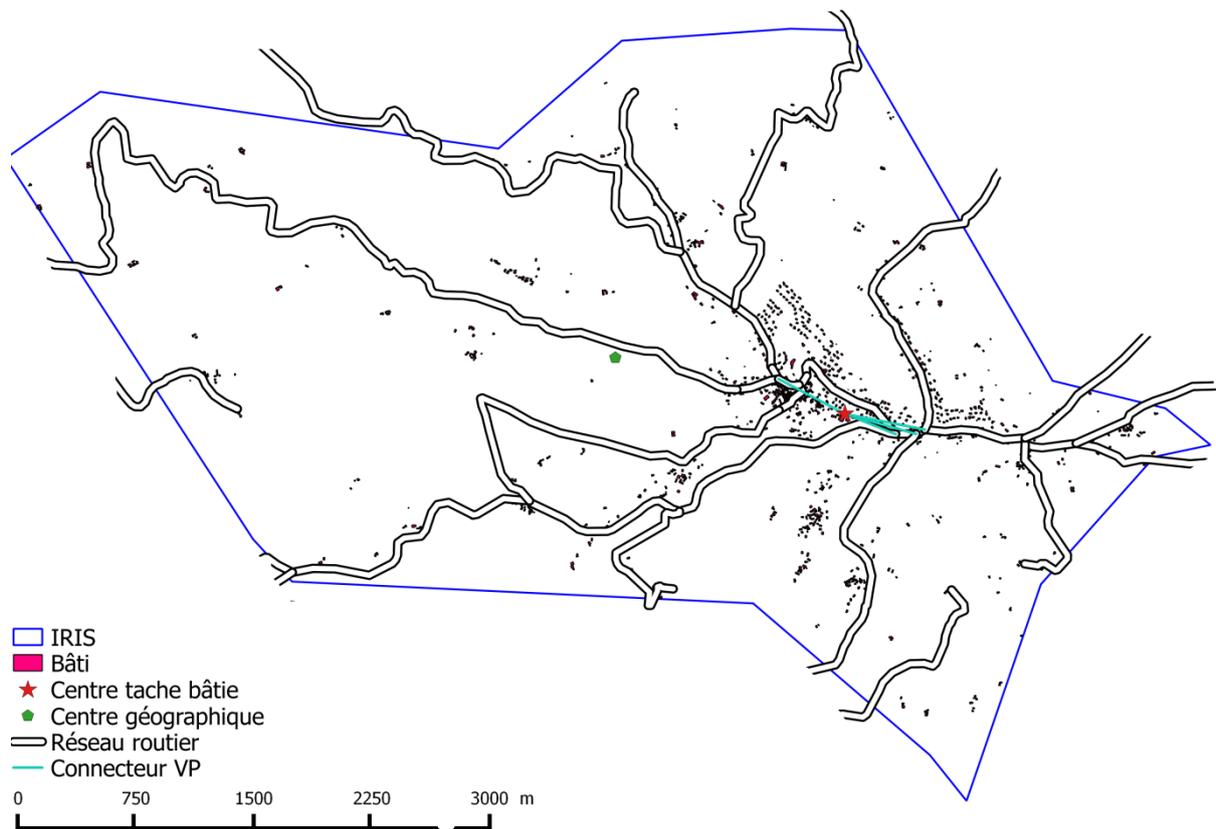


FIGURE 6 : EXEMPLE DE DEUX CONFIGURATIONS DE CODIFICATION DES CENTROÏDES

La simplicité d'implémentation de cette approche tient à l'hypothèse d'existence d'une position géographique des centroïdes de zones. De ce fait, la définition des connecteurs devient intimement liée à la définition des centroïdes. L'exemple ci-dessus (Figure 6) illustre l'impact du choix de la méthodologie de codification des centroïdes sur les connecteurs. Les deux approches – centre géographique de la zone et centroïde de la tache urbaine - aboutissent à deux centres différents et écartés de presque 1500m. Cet écart se répercutera sur la longueur des connecteurs générés selon l'une ou l'autre méthode.

## 2.2.4 LIMITES DE LA MÉTHODOLOGIE ACTUELLE

### 2.2.4.1 HYPOTHÈSES

Il est évident que la méthodologie actuelle de codification des connecteurs est tributaire de la méthode de positionnement des centroïdes. Ces derniers sont supposés représenter l'individu moyen. C'est d'ailleurs pour cette raison que le choix final de la position des centres des zones SIMBAD s'est fixé sur le barycentre de la tache urbaine. Cette tache est censée regrouper les différentes positions géographiques des individus résidant ou travaillant à l'intérieur de la zone.

A priori, cette approximation semble être cohérente avec la définition des connecteurs qui doivent représenter la *distance moyenne* d'accès au réseau. Toutefois, sa pertinence tient à quatre grandes hypothèses détaillées dessous.

Positionner le centroïde au centre de la tache urbaine suppose l'uniformité de la densité de la population et des emplois. Autrement dit, ce positionnement ne prend pas en considération les volumes bâtis, ce qui peut être valide pour certaines zones dont la

typologie du bâti est homogène comme les zones pavillonnaires, mais demeure contestable ailleurs (Figure 7). Ainsi, à surface égale, un logement collectif sur plusieurs étages est équivalent à un logement individuel.

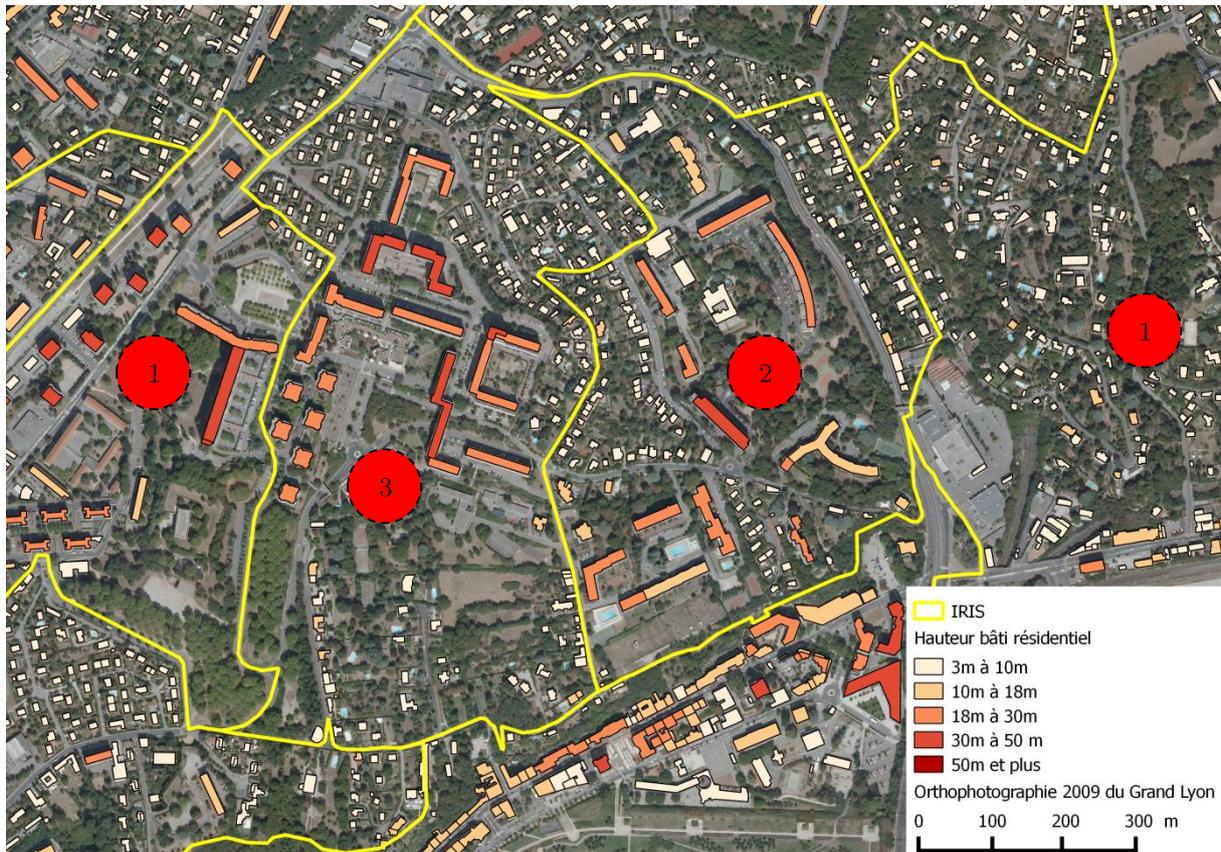


FIGURE 7 : EXEMPLE DE ZONES À CALUIRE-ET-CUIRE AVEC UNE DENSITÉ NON UNIFORME (ORTHOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2009)

Dans cet exemple, l'hypothèse d'uniformité est valide pour les IRIS 1 abritant des immeubles de même hauteur, mais elle est en défaut pour les IRIS 2 et 3 à bâti mixte.

En parallèle à cette hypothèse, la précédente méthodologie ne différencie pas le bâti résidentiel du reste - commercial, industriel ou autres -. De ce fait, toutes les typologies d'usage du sol sont supposées équivalentes. C'est-à-dire qu'à surface égale, un bâtiment résidentiel est équivalent à un centre commercial ou à une université, tandis que les pratiques de mobilité des individus correspondant à chaque typologie de bâti sont différentes.

La troisième hypothèse prise en compte dans l'implémentation de cette méthodologie est l'égalité entre **distance à l'arrêt de l'individu moyen** et **distance moyenne des usagers de l'arrêt**. En effet, le positionnement au centre de l'individu moyen prend en compte tous les individus peuplant la tache urbaine. Or, la longueur d'un connecteur doit être 'assez proche de la distance moyenne effectivement parcourue par les utilisateurs des transports collectifs (et non pas une distance moyenne de la population aux arrêts)' (Bonnell, 2004). Dans ce cas, cette méthode aboutirait, probablement, à une surestimation de la longueur des connecteurs des transports collectifs et par conséquent à une dégradation relative de leur attractivité (Godinot, 2008).

## 2.2.4.2 PERTINENCE

Outre ces hypothèses sous-jacentes à l'implémentation de cette méthode, sa pertinence dépend de la continuité du tissu urbain de la zone (Godinot, 2008). Il est à noter que pour des zones englobant deux ou plusieurs taches urbaines déconnectées, le centroïde de zone est le barycentre des centres des taches. Par conséquent, il se peut que le centroïde se positionne loin des deux taches (Figure 8).

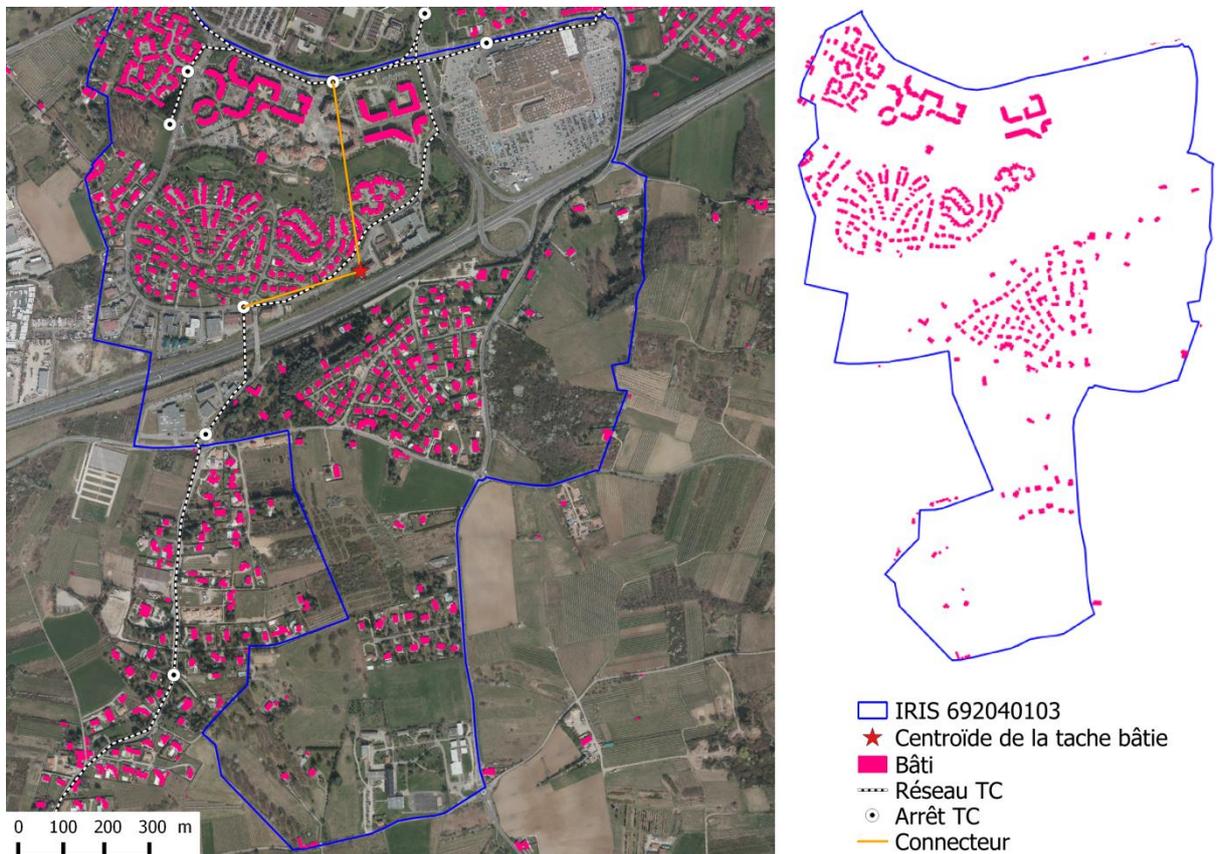


FIGURE 8 : EXEMPLE D'UNE ZONE AVEC UN TISSU URBAIN DISCONTINU (ORTHOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012)

Dans cet exemple, deux taches urbaines distinctes, séparées par une voie rapide, sont regroupées au sein d'un même IRIS. Chacune des deux taches est bien desservie par les transports en commun. Toutefois, l'application de la méthodologie précédente oblige au positionnement du centroïde de l'IRIS entre les deux taches en s'éloignant des arrêts. Ainsi, la longueur des connecteurs est surestimée pour les deux sous-zones.

La discontinuité du tissu urbain caractérise particulièrement les zones périurbaines. Grâce à une analyse des données CORINE Land Cover 2006 - CLC06 - appliquées à l'aire SIMBAD, il paraît que la majorité des IRIS sont formés de tissus discontinus<sup>6</sup>. Seuls l'hyper-centre lyonnais et le quartier de Montchat ont un tissu continu au sens de CLC06 (Figure 9). Dans ce cas particulier, l'application de cette méthodologie peut être contestable voire préjudiciable.

<sup>6</sup> Un tissu urbain est dit discontinu si les surfaces artificialisées (bâti, voies, etc.) ont plus de 30% et moins de 80% de la surface totale. Au-delà de 80%, le tissu est dit continu (Commissariat général au développement durable, 2009).

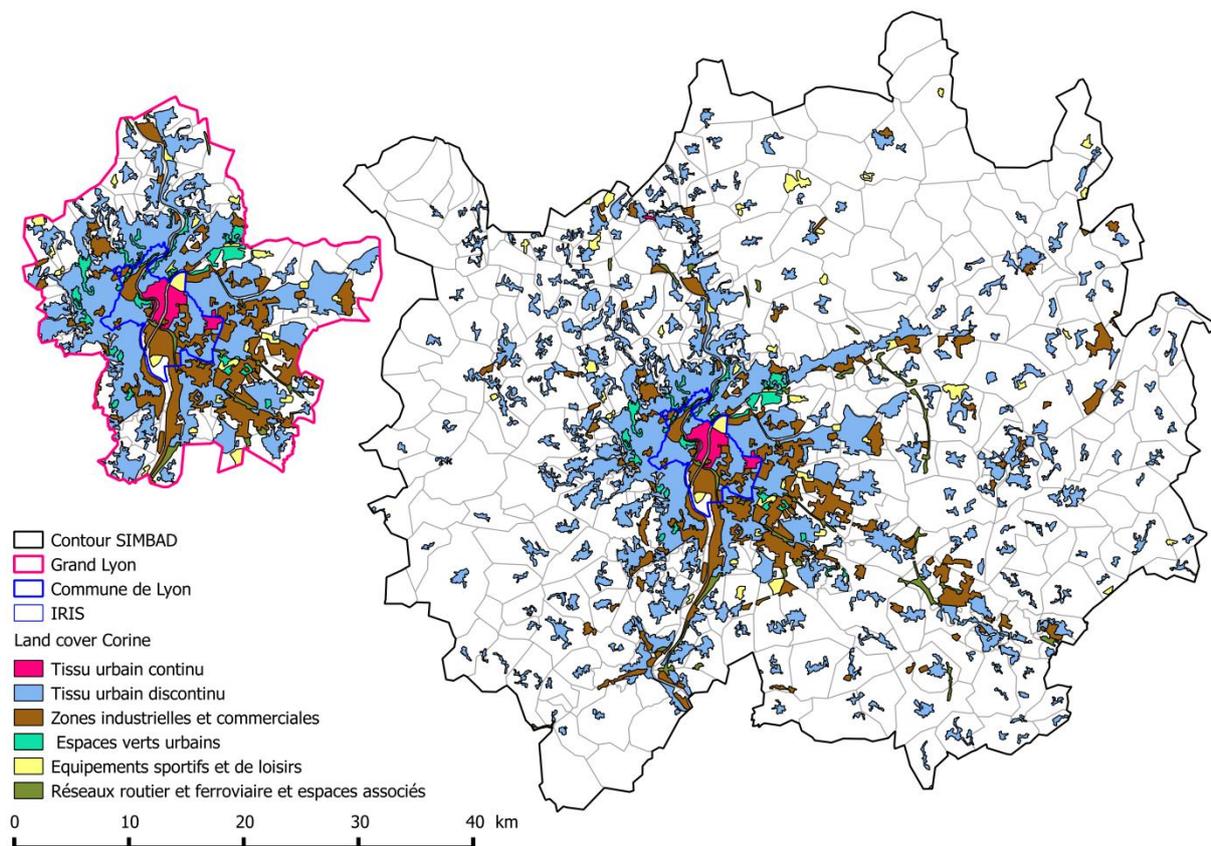


FIGURE 9 : NATURE DU TISSU URBAIN SUR L'AIRE SIMBAD (DONNÉES CORINE LAND COVER 2006)

Pour finir, nous mettons l'accent sur une problématique commune aux méthodologies de génération de connecteurs utilisant la distance à vol d'oiseau. À l'encontre de la distance réelle dite distance réseau ou viaire, la distance à vol d'oiseau ne tient pas compte de l'effet de coupure comme les obstacles naturels et artificiels infranchissables (fleuves, voie rapide, voie ferrée, etc.). Par conséquent, plusieurs connecteurs enjambent ces obstacles et 'court-circuitent' à tort le trajet (Figure 10). Dans l'exemple ci-dessous, les connecteurs en rouge traversent le Rhône sans emprunter de pont. Ainsi, ils ont une longueur réduite de presque 311m. Le trajet viaire – représenté ici en jaune - passant par le pont Winston Churchill est, quant à lui, d'une longueur totale de 1600m. Cette imprécision de calcul peut être nuisible dans le cas des connecteurs TC.

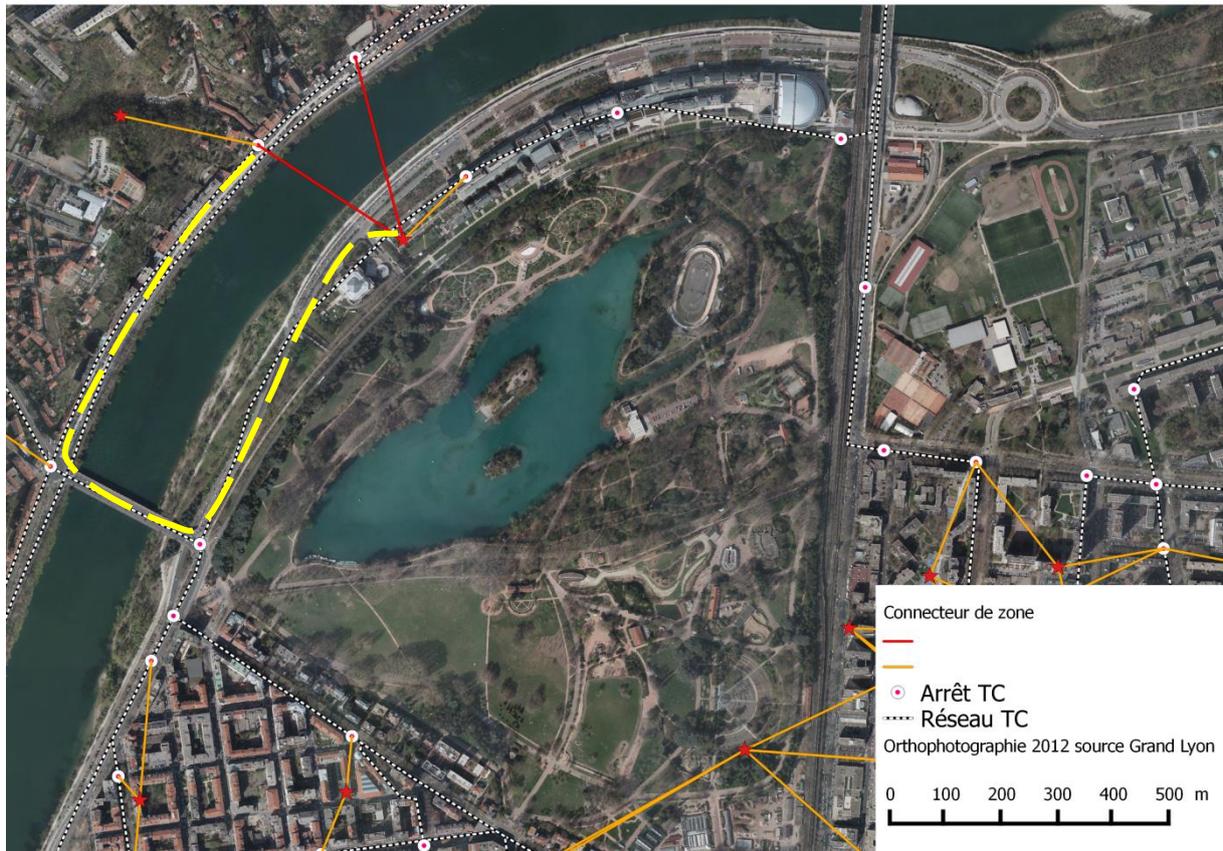


FIGURE 10 : EXEMPLE DE CONNECTEURS DE L'IRIS 693860101 ENJAMBANT LE RHÔNE (ORTHOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012)

Il est possible d'enrayer ce problème en interdisant automatiquement le passage des connecteurs à travers des obstacles non franchissables, toutefois, il faut connecter 'à la main' les zones correspondant à ces situations. Une autre solution serait d'utiliser la distance viaire pour la génération des connecteurs. Dans ce cas, la longueur des connecteurs serait égale à la distance parcourue à pied sur le réseau routier. Dans ce mémoire, nous utilisons la distance à vol d'oiseau et sa variante rectifiée dite la distance rectilinéaire. Le calcul de la distance viaire exige un effort de traitement considérable et non indispensable, à notre avis, à l'élaboration de la nouvelle méthodologie.

#### 2.2.4.3 EFFETS SUR LES CHARGES

Dans cette partie, nous essayons de mettre en évidence l'incidence de la définition des connecteurs sur les résultats SIMBAD. Il paraît logique que la portée de cet impact atteigne toutes les étapes intégrant dans leur calcul le temps généralisé des transports en commun. Ainsi, le choix modal et l'affectation seraient les plus concernées. Nous avons choisi cette dernière étape pour l'évaluation de l'incidence, car son implémentation nous est plus rapide et plus facile<sup>7</sup>. Pour ce faire, nous utilisons le réseau SIMBAD et sa codification actuelle des connecteurs. Quatre tests sont conduits sous le logiciel commercial PTV VISUM<sup>®</sup> édition 12.5, édité par la société PTV AG. À chaque expérience correspond un système de connecteurs. Les résultats de ces expériences sont comparés aux résultats de l'affectation de SIMBAD, dits par la suite de référence.

<sup>7</sup> Nous nous sommes limités à affecter les charges sans caler pour chaque configuration les fonctions correspondantes.

Pour les quatre expériences, nous nous sommes basés sur la définition actuelle des connecteurs que nous avons modifiée pour avoir quatre nouvelles configurations. La première configuration consiste en une modification aléatoire des longueurs de connecteurs. Nous avons multiplié ou divisé les longueurs de référence par des coefficients proches de 1 pour avoir de nouvelles longueurs. Les trois autres expériences consistent en une modification des connecteurs les uns par rapport aux autres, de sorte que chaque zone, ayant deux connecteurs ou plus, ait au moins un connecteur de référence et autre modifié. Ainsi, pour chaque zone, nous avons gardé les connecteurs de référence de moindre coût et modifié le reste en multipliant leur longueur par 1.5, 1.7 et 2.

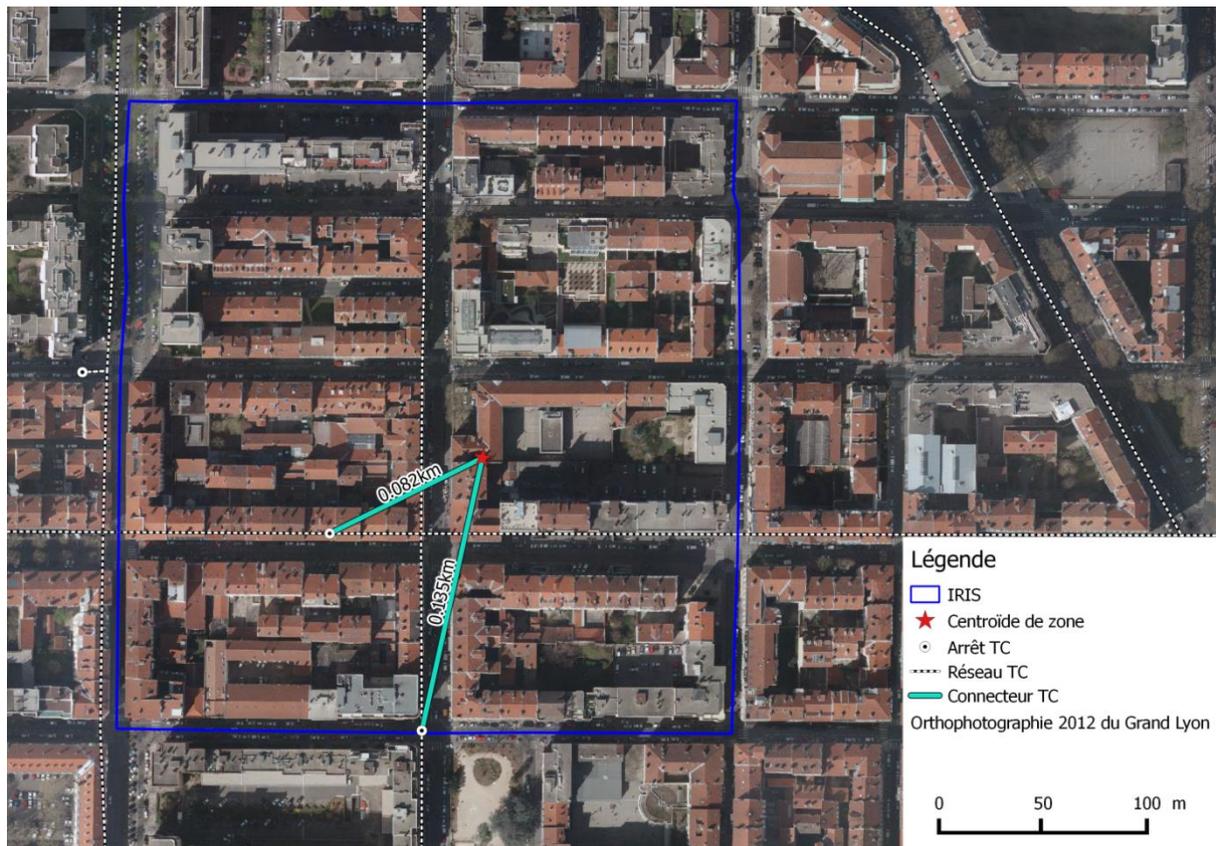


FIGURE 11 : EXEMPLE D'UNE SITUATION DE RÉFÉRENCE D'UNE CONFIGURATION DE CONNECTEURS

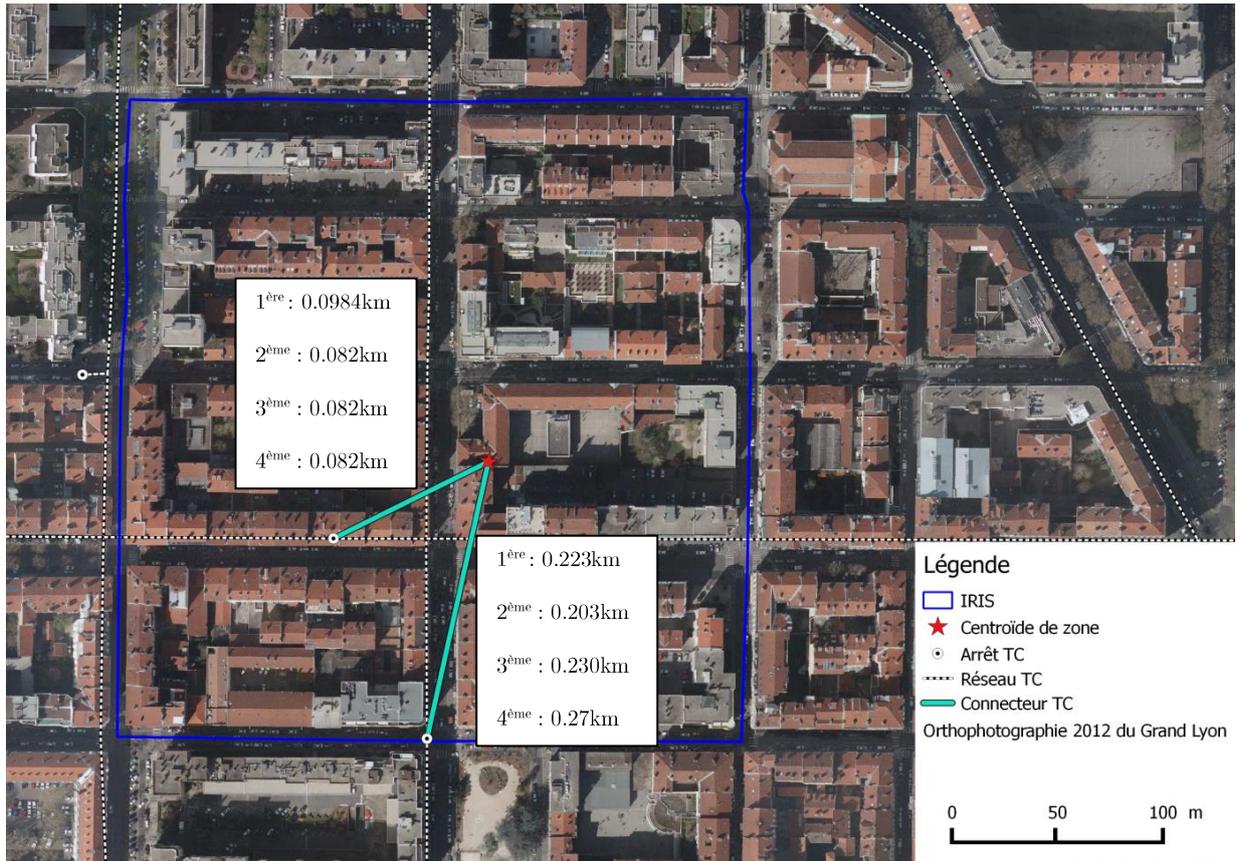


FIGURE 12 : LES QUATRE CONFIGURATIONS TESTÉES

L'incidence d'une configuration de codification est étudiée à travers l'écart entre les charges estimées et les charges de référence des lignes de transport en commun. Cet écart est approché par la racine carrée de l'erreur quadratique moyenne, plus connue sous le sigle RMSE - Root-Mean-Square Error - et son pourcentage PRMSE - Pourcentage of RMSE -. Le RMSE est semblable dans son calcul à l'écart-type et le PRMSE au coefficient de variation.

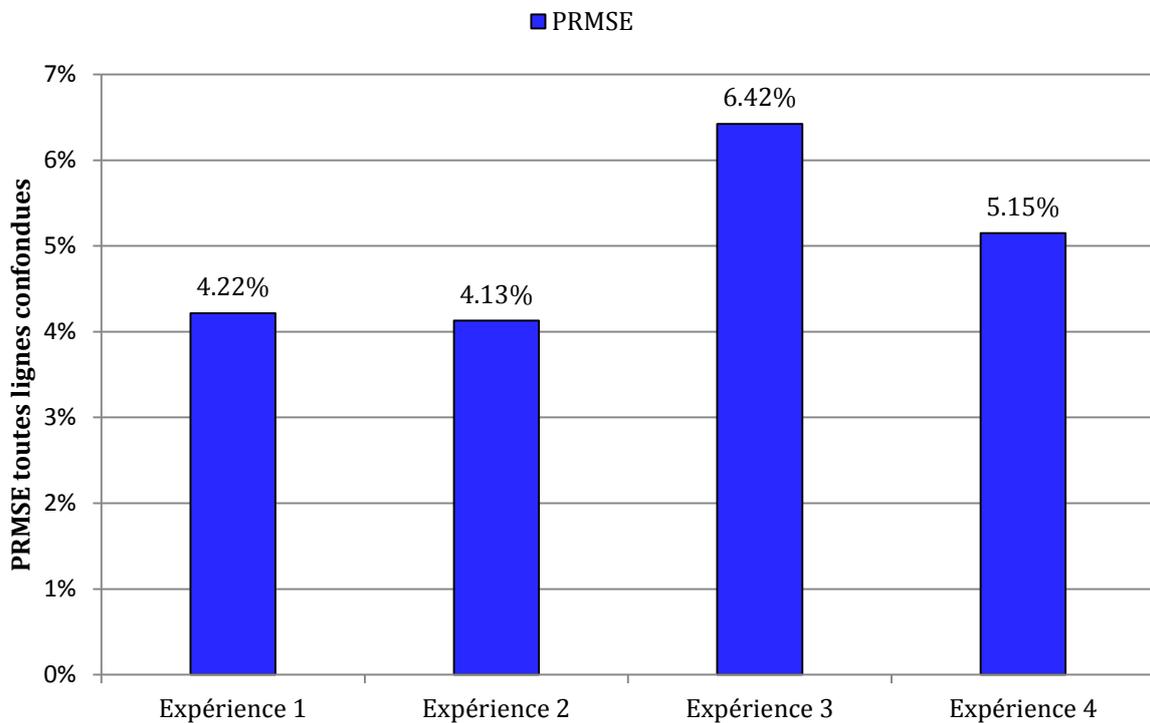
$$RMSE = \sqrt{\frac{\sum_1^n (\tilde{C}_i - C_i)^2}{n}}$$

$$PRMSE = \frac{RMSE}{\tilde{C}}$$

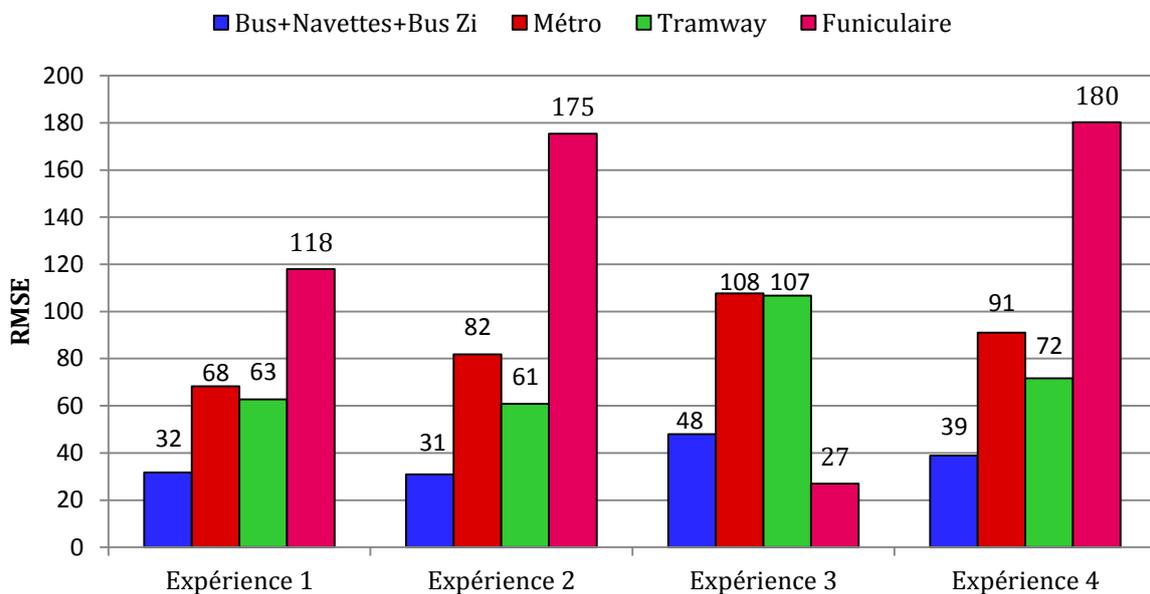
Où  $\tilde{C}_i$  et  $C_i$  sont respectivement la charge estimée et de référence du tronçon  $i$ .  $\tilde{C}$  est la charge moyenne de référence.

Le calcul du RMSE est censé être pertinent en comparant entre des lignes ayant des charges de même ordre. Ainsi, dans un premier temps, nous calculons cet indicateur sans différencier entre les lignes, puis nous catégorisons les lignes selon leur mode. On note que le calcul des écarts est procédé au niveau de la charge partielle des lignes, c'est-à-dire au niveau des arcs entre les stations.

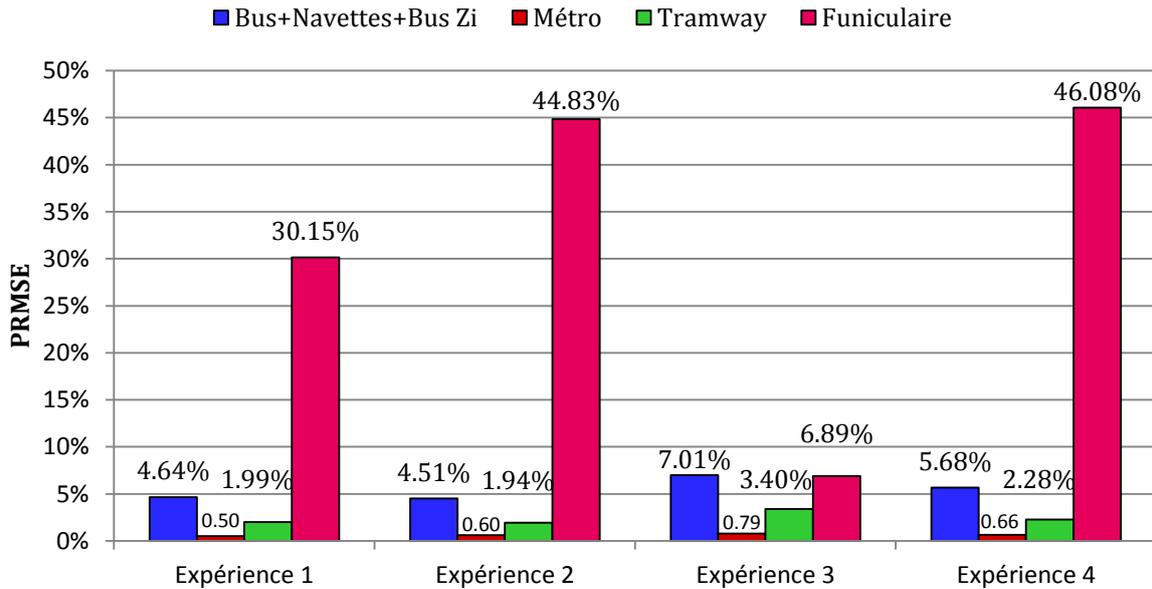
Nous affectons à l'horaire une matrice de demande relative aux transports en commun. Cette affectation prend en considération le temps d'attente ou de correspondance des usagers selon les services suivants (voir Annexe.1).



GRAPHIQUE 1 : PRMSE DE TOUTES LES LIGNES DE TRANSPORT EN COMMUN



GRAPHIQUE 2 : RMSE DES QUATRE EXPÉRIENCES SELON LA CATÉGORIE DE LA LIGNE



GRAPHIQUE 3 : PRMSE DES QUATRE EXPERIENCES SELON LA CATEGORIE DE LA LIGNE

Ce test montre que les résultats de l'affectation SIMBAD sont sensibles à la définition des connecteurs. De par la grandeur de ses charges, le métro affiche le moins de variation autour de la moyenne – valeurs du PRMSE -. Les charges du funiculaire, par contre, varient sensiblement de 30% à 46%. Ceci est dû, à ses faibles charges. Les charges du tramway et du bus varient légèrement autour de la moyenne.

## 2.3 CONCLUSION

Malgré la simplicité d'implémentation de la méthodologie actuellement utilisée par SIMBAD, sa pertinence pour le système des transports en commun peut être contestable. Ceci est principalement dû au biais d'agrégation induit par l'utilisation d'une position géographique des centroïdes de zones pour la définition des connecteurs. Conscients de cette limite, nous nous sommes posés comme objectif principal de définir les connecteurs indépendamment de la définition des centroïdes.

Dans ce qui suit, nous proposons une nouvelle approche de codification des connecteurs. L'élaboration de cette méthodologie a pour principe directeur la réduction des effets du biais d'agrégation déjà présenté. C'est-à-dire, utiliser, autant que faire se peut, des données désagrégées. Dans ce sens, nous nous sommes investis dans la construction d'une base de données décrivant la population localisée de l'aire SIMBAD. À cette dernière s'ajoute une base des emplois localisés.

# CHAPITRE 3

---

## 3 NOUVELLE MÉTHODOLOGIE DE CODIFICATION DES CONNECTEURS

Lors de la codification des connecteurs, deux questions principales se posent : Quels nœuds connecter ? Et quelle longueur attribuer au connecteur correspondant ? La nouvelle approche de codification des connecteurs essaie de répondre à ces questions dans le cadre du système des transports en commun. Le choix de ce système en particulier découle de deux principales raisons. D'une part, l'offre des transports collectifs semble être plus sensible à la distance de rabattement que le système de transport individuel. D'autre part, peu de travaux de recherche ont sondé cette sensibilité, et notamment celle des transports en commun.

Pour s'affranchir des limites propres à la méthodologie SIMBAD, nous avons fait le choix d'adopter une approche désagrégée pour la génération des connecteurs. Cette approche est censée rendre compte des pratiques de marche à pied vers et depuis les arrêts.

Nous allons commencer par éclairer le choix des nœuds de connexion, pour aborder ensuite le calcul de la longueur des connecteurs.

### 3.1 CHOIX DES ARRÊTS À CONNECTER

#### 3.1.1 PRÉSENTATION

Un autre attribut, aussi important que la longueur de connecteur, est le nœud auquel il se connecte (Ortúzar S, Juan D and Willumsen, 2011). Dans le cas des transports en commun ce nœud représente un ou plusieurs arrêts ou stations.

Une zone peut être parcourue par plusieurs lignes de transport et accueillir plusieurs arrêts (Figure 13). De ce fait, différentes configurations de codification des connecteurs sont possibles.

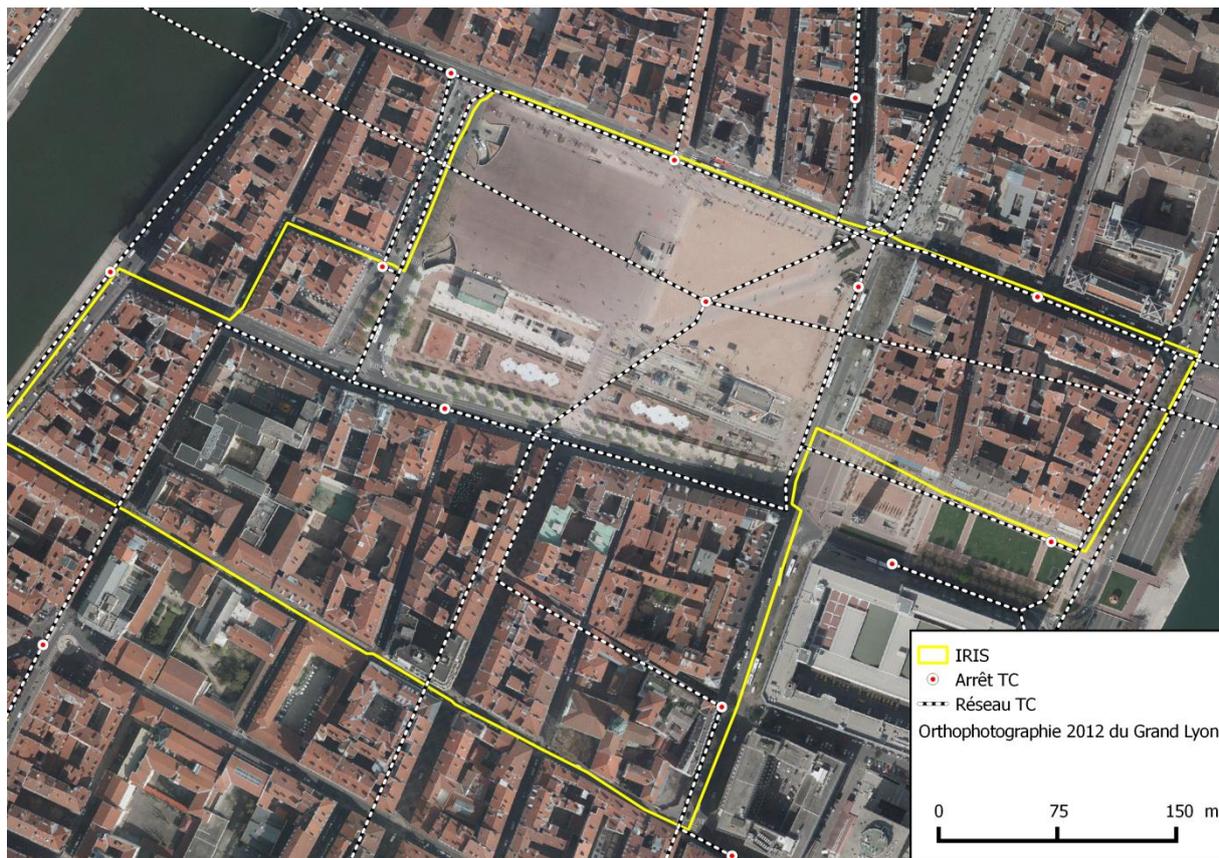


FIGURE 13 : EXEMPLE D'IRIS ET DE DISPOSITIONS GÉOGRAPHIQUES DES ARRÊTS DES LIGNES TC (ORTHOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012)

Dans cette partie, nous allons expliciter les critères de choix des arrêts à connecter aux zones. Ces critères se réfèrent essentiellement à la nature des lignes desservant les arrêts - catégorie de la ligne - et à leur accessibilité -distance de rabattement-.

### 3.1.2 ACCESSIBILITÉ DES POINTS DE DESSERTE

Nous appelons un '*point de desserte*' un arrêt relatif à un unique service commercial de transport en commun. Il est desservi par une seule ligne orientée. Son accessibilité est comprise au sens de la distance de rabattement ou de diffusion parcourue à pied et ayant comme origine ou destination cet arrêt. Indépendamment de la nature de la ligne desservant le point de desserte, ce dernier est considéré comme accessible depuis une zone s'il est situé à une distance maximale  $d$  d'une entité résidentielle ou d'activité appartenant à cette zone. Dans l'exemple ci-dessous, la première entité de l'IRIS 12760000 interceptant la ligne de bus est à 2km. Si  $d$  est inférieure à 2km alors cet IRIS n'a pas accès à la ligne de bus. À l'inverse l'IRIS 692800000 a accès à cette même ligne.

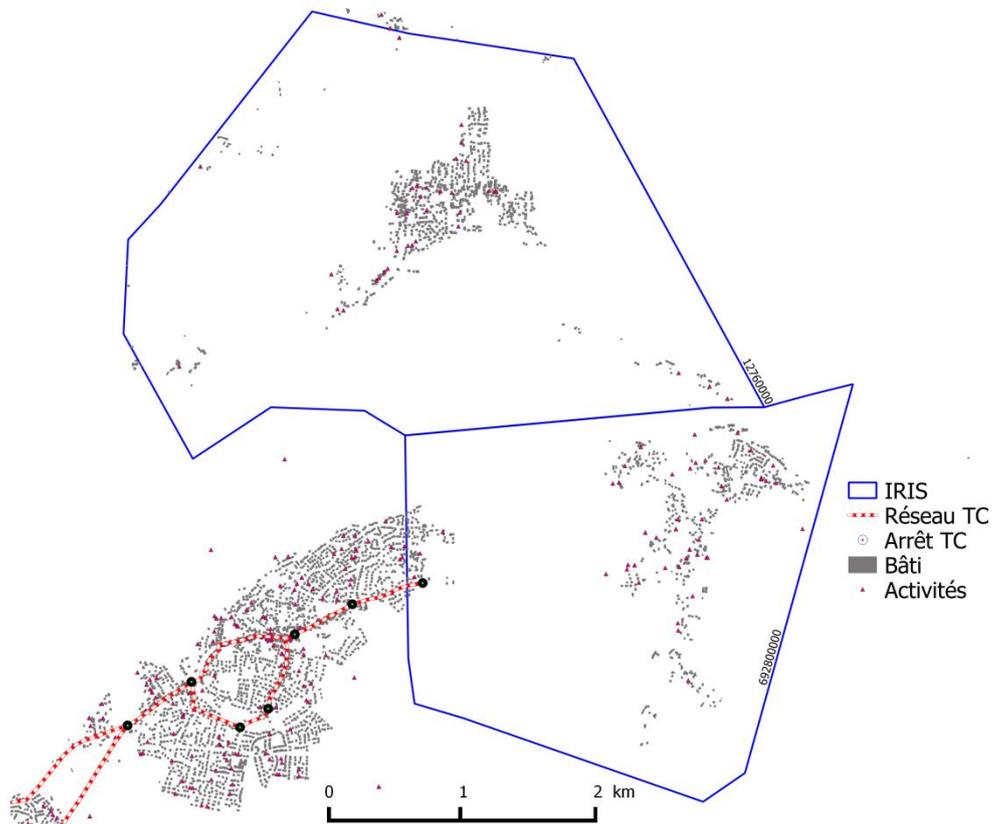
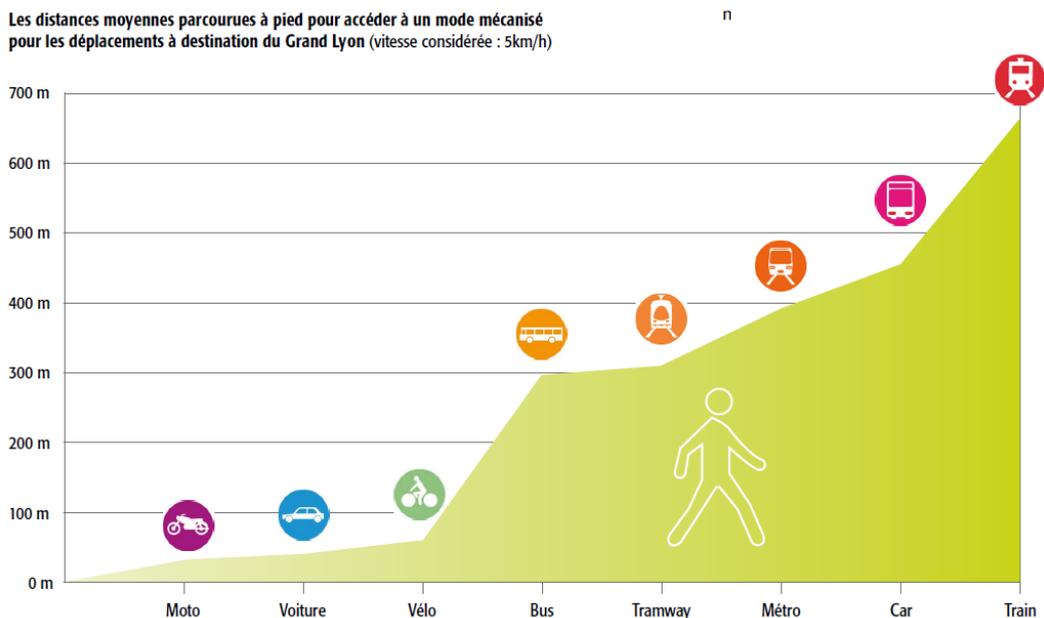


FIGURE 14 : EXEMPLE D'ACCESSIBILITÉ DES IRIS

Plusieurs études et enquêtes ont montré que la distance de rabattement ou de diffusion varie en fonction des caractéristiques des lignes (Agence d'Urbanisme de Lyon et al., 2013). En effet, un usager est susceptible de parcourir plus de distance pour accéder à une meilleure offre de transport. De ce fait, la distance  $d$  dépend de la nature de la ligne desservant le point de desserte.



GRAPHIQUE 4 : DISTANCE D'ACCÈS AUX ARRÊTS EN FONCTION DU MODE (SOURCE AGENCE D'URBANISME DE LYON ET AL., 2013)

Désormais, un point de desserte est dit '*accessible*' depuis une zone s'il intercepte une entité résidentielle ou d'activité appartenant à cette zone dans la limite d'une distance  $d_{cat\_ligne}$  où '*cat\_ligne*' est la catégorie de ligne desservant le point de desserte en question. Ces différentes valeurs de la distance sont détaillées dans le § 3.3.4.

Dans un premier temps, on catégorise les lignes selon leur mode. Par la suite, pour chaque mode, on peut différencier entre les lignes selon d'autres critères comme la fréquence moyenne, la nature de la ligne (TCSP et autres), la congestion à bord, etc.

On associe à chaque point de desserte accessible depuis une zone donnée un connecteur. Il y a donc autant de connecteurs par zone que de points de desserte accessibles depuis cette zone.

Dans ce qui suit, nous étendons la notion d'accessibilité aux lignes. Une ligne est dite '*accessible*' depuis une zone si elle dessert un point accessible depuis cette même zone.

### **3.1.3 AGRÉGATION DES CONNECTEURS**

Sur la base du critère d'accessibilité, on a identifié tous les points de desserte, et par conséquent tous les connecteurs possibles pour les différentes zones.

Pour une même zone, un seul connecteur par ligne semble être, à la fois, nécessaire et suffisant. En effet, pour conserver la diversité de l'offre de transport en commun, toutes les lignes TC considérées accessibles depuis cette zone doivent y être connectées. De même, il n'est pas indispensable de relier une même zone à plusieurs arrêts d'une même ligne de transport. Un seul connecteur permettant de traduire les conditions moyennes d'accès à cette ligne via ses points de desserte accessibles peut être suffisant. La longueur de ce dernier est la moyenne des longueurs des connecteurs correspondants pondérées par le poids de chaque point de desserte (voir formalisation 3.2).

À ce stade, il y a autant de connecteurs par zone que de lignes accessibles depuis cette zone.

Dans la pratique, un arrêt est souvent mutualisé entre plusieurs lignes. C'est le cas, par exemple, de la plupart des arrêts de bus (Figure 15). De même, et en fonction des objectifs de l'étude, une simplification de la codification du réseau peut amener à modéliser plusieurs points de desserte juxtaposés en un seul point, dit point d'arrêt dans le vocabulaire VISUM<sup>®</sup> (PTV VISUM, 2012).



FIGURE 15 : MUTUALISATION D'UN POINT DE DESSERTE ENTRE LIGNES (BUS 1 ET 34) ET DIRECTIONS (ALLER-RETOUR) DE BUS (FOND GOOGLE MAPS)

Dans ce cas et pour une même zone, l'application de la méthodologie précédente peut aboutir à la génération de plusieurs connecteurs pour une même position géographique d'un arrêt. Pour ne garder qu'un seul connecteur, une agrégation de ces derniers est indispensable.

On appelle '*point d'arrêt*' la codification simplifiée de plusieurs points de desserte en un seul point. Il est donc la superposition de plusieurs points de desserte et est, par conséquent, desservi par plusieurs lignes. Dans l'exemple, précédent, le point d'arrêt '*Alsace*' regroupe quatre points de desserte relatifs à deux lignes orientées de transport. Dans ce cas la longueur du connecteur reliant la zone au point d'arrêt '*Alsace*' est la moyenne pondérée des longueurs des connecteurs des quatre points de desserte.

### 3.2 FORMALISATION DE LA GÉNÉRATION DES CONNECTEURS

Soit  $I$  une zone quelconque et  $Pnt_{desserte} = \{Pnt_{desserte_i} | i \in \llbracket 1; N \rrbracket\}$  l'ensemble des  $N$  points de desserte de l'aire d'étude. Soit  $d_{(I, Pnt_{desserte_i})}$  une métrique entre la zone  $I$  et un point de desserte  $i$ . La définition exacte de cette métrique ainsi que sa méthodologie de calcul sont détaillées dans le § 3.3.

$Pnt_{desserte, I} = \cup_i \{Pnt_{desserte_i} | d_{(I, Pnt_{desserte_i})} \leq d_{cat.ligne}\}$  est l'ensemble de tous les points de desserte accessibles depuis la zone  $I$ .

Soit  $L(Pnt_{desserte_i})$  la ligne desservant le point de desserte  $Pnt_{desserte_i}$ .

Pour une codification en points de desserte,  $C_I = \{C_{L(Pnt_{desserte_i}),I} \mid Pnt_{desserte_i} \in Pnt_{desserte,I}\}$  est l'ensemble des connecteurs de la zone  $I$ .  $C_{L(Pnt_{desserte_i}),I}$  est le connecteur reliant la zone  $I$  au point de desserte  $i$ . Sa longueur est définie telle que:

$$\|C_{I,L(Pnt_{desserte_i})}\| = \sum_k \frac{poids(Pnt_{desserte_k}) * d_{(I,Pnt_{desserte_k})}}{\sum_k poids(Pnt_{desserte_k})}, L(Pnt_{desserte_k}) = L(Pnt_{desserte_i})$$

Où  $poids(Pnt_{desserte_k})$  est une pondération associée à chaque point de desserte (population ou emplois environnant, par exemple) (voir § 3.3).

$L(Pnt_{desserte_k}) = L(Pnt_{desserte_i})$  est la condition d'agrégation des points de desserte accessibles depuis une même zone et appartenant à une même ligne.

Cette formalisation peut être adaptée à une codification simplifiée des arrêts en points d'arrêt. Dans ce cas, un connecteur relie une zone à un point d'arrêt  $Pnt_{Arr_i}$  - comme dans SIMBAD d'ailleurs -. Par définition  $Pnt_{Arr_i}$  regroupe plusieurs points de desserte  $Pnt_{desserte_k}$ . On écrit alors  $Pnt_{desserte_k} \in Pnt_{Arr_i}$ .

Dans ce cas, la longueur d'un connecteur  $C_{I,Pnt_{Arr_i}}$  est :

$$\|C_{I,Pnt_{Arr_i}}\| = \sum_k \frac{poids(L(Pnt_{desserte_k})) * \|C_{I,L(Pnt_{desserte_k})}\|}{\sum_k poids(L(Pnt_{desserte_k}))}, Pnt_{desserte_k} \in (Pnt_{Arr_i} \cap Pnt_{desserte,I})$$

La condition  $Pnt_{desserte_k} \in (Pnt_{Arr_i} \cap Pnt_{desserte,I})$  regroupe tous les points de desserte  $Pnt_{desserte_k}$  accessibles depuis la zone  $I$  et compris dans le point d'arrêt  $Pnt_{Arr_i}$ . Le poids de la ligne est tel que :

$$poids(L(Pnt_{desserte_k})) = \begin{cases} \sum_j poids(Pnt_{desserte_j}); & L(Pnt_{desserte_j}) = L(Pnt_{desserte_k}) \\ poids(Pnt_{desserte_k}); & Sinon \end{cases}$$

Pour être en concordance avec la logique de codification du projet SIMBAD, nous allons utiliser la deuxième formalisation correspondant à une codification simplifiée des arrêts en points d'arrêt.

Cette formalisation calcule  $\|C_{I,L(Pnt_{desserte_k})}\|$  pour chaque ligne et indépendamment des autres lignes considérées aussi accessibles. Autrement dit, la méthodologie proposée ne prend pas en compte la concurrence entre les différentes lignes. Dans les modèles à quatre étapes, la concurrence entre les lignes n'est prise en compte qu'à la dernière étape de la chaîne de modélisation, c'est-à-dire l'affectation TC (voir Figure 3). De ce fait, toute prise en compte de la concurrence à l'étape de génération des connecteurs peut constituer une incohérence interne du modèle. Cette incohérence peut être levée, éventuellement, par une modélisation globale en boucle. Ainsi, les flux des arrêts en sortie du modèle peuvent alimenter la génération des connecteurs, qui, à leur tour, modifient les charges et ainsi de suite. Pour le moment, nous adoptons, l'hypothèse simplificatrice de non concurrence.

### 3.3 FORMALISATION DE LA MÉTHODOLOGIE DE CALCUL DE LA LONGUEUR DES CONNECTEURS

#### 3.3.1 FORMALISATION GÉNÉRALE

La longueur d'un connecteur reliant une zone à un arrêt de transport en commun doit traduire approximativement la distance moyenne parcourue par les usagers issus de cette zone pour accéder à cet arrêt. Il s'agit '[d']une distance assez proche de la distance moyenne effectivement parcourue par les utilisateurs des transports collectifs' (Bonnell, 2004). Soit  $i$  un individu résidant ou travaillant à l'intérieur d'un iris  $I$  et utilisant une ligne de transport desservant une station  $Pnt_{desserte_k}$ . La longueur du connecteur reliant  $I$  à  $Pnt_{desserte_k}$  est :

$$\|C_{I,Pnt_{desserte_k}}\| \cong d_{moy,(I,Pnt_{desserte_k})} = \frac{1}{n} \sum_{i=1}^n d_{(I,Pnt_{desserte_k}),i} \quad (1)$$

Où  $n$  est la somme des flux entrant et sortant du point de desserte  $Pnt_{desserte_k}$  et ayant comme origine ou destination la zone  $I$ .  $d_{(I,Pnt_{desserte_k}),i}$  est la distance parcourue par l'individu  $i$  entre son origine ou sa destination et le point de desserte.

Dans tout ce qui suit, nous cherchons à construire des connecteurs avec des longueurs proches des distances parcourues par les usagers potentiels des transports en commun. L'intégration, par la suite, de ce calcul dans la chaîne de modélisation à quatre étapes soulève quelques questions et contraint à adopter certaines hypothèses. En effet, si la longueur calculée selon la logique du modèle (1) est adaptée à l'affectation de la 'matrice O-D pour les transports en commun', où les déplacements utilisent effectivement les TC, elle l'est moins pour les autres étapes. De la distribution au choix modal, l'application de cette distance sur une 'matrice OD tous modes confondus' peut fausser l'estimation de la demande des transports collectifs. Conscients de cette limite, nous l'acceptons quand même pour éviter de définir des connecteurs propres à chaque étape de modélisation.

S'il est facile de déterminer le nombre d'usagers d'un point de desserte (exploitation des données billettiques par exemple), il est plutôt difficile sinon onéreux de localiser leur provenance exacte ou leur destination finale (nécessité d'une enquête, réticence des enquêtés, inexactitude des réponses, etc.) (O'Sullivan and Morrall, 1996). Ceci est d'autant plus vrai pour l'année de référence que pour les années de prévision. Ainsi, en l'absence d'une telle donnée – localisation géographique x-y des usagers TC –, la méthode (1) ne peut pas être appliquée telle quelle.

Pour lever cette contrainte, on diminue la résolution de notre analyse en introduisant une partition spatiale de l'IRIS  $I$  en  $m$  sous-zones. Les individus ne sont plus localisés exactement avec leurs coordonnées x-y mais avec la position de la sous-zone de départ ou d'arrivée. La distance par conséquent est calculée depuis le centre de cette sous-zone. Voici la variante correspondant à cette simplification :

$$\|C_{I,Pnt_{desserte_k}}\| \cong \|\tilde{C}_{I,Pnt_{desserte_k}}\| \cong \frac{\sum_{i=1}^m d_{(Z_i,Pnt_{desserte_k})} * U_{(Z_i,Pnt_{desserte_k})}}{\sum_{i=1}^m U_{(Z_i,Pnt_{desserte_k})}} \quad (2)$$

Où  $Z_I = \{Z_i | i \in \llbracket 1; m \rrbracket\}$  est une partition de l'iris  $I$  en  $m$  sous-zones.  $d_{(Z_i,Pnt_{desserte_k})}$  est la distance entre le centroïde de la sous-zone  $Z_i$  et le point de desserte  $Pnt_{desserte_k}$ .

$U_{(Z_i, Pnt_{desserte_k})}$  est le nombre d’usagers issus de  $Z_i$  et utilisant  $Pnt_{desserte_k}$ . On a selon l’équation (1),  $\sum_1^m U_{(Z_i, Pnt_{desserte_k})} = n$ .

La qualité de cette approximation dépend de la partition  $Z_I$  et de la méthodologie de calcul de  $d_{(Z_i, Pnt_{desserte_k})}$ . Toutes choses étant égales par ailleurs, la partition qui minimise l’erreur entre  $C_{I, Ligne}(Pnt_{desserte_k})$  et  $\tilde{C}_{I, Ligne}(Pnt_{desserte_k})$  est constituée de  $m$  sous-zones, dont les centroïdes coïncident avec les  $m$  lieux d’origine ou de destination (x-y) des  $n$  usagers.



FIGURE 16 : EXEMPLE DE PARTITION EN CARROYAGE RÉSUMANT L’INFORMATION DÉSAGRÉGÉE

Dans le cadre de notre étude, nous ne disposons ni des données relatives aux flux des usagers des arrêts ni de leur géolocalisation exacte. Par conséquent, nous avons exploré l’estimation de ces flux potentiels à partir d’une régression sur des variables explicatives faciles à construire et à prédire.

Ces modèles de régression, dits ‘modèles directs’ par opposition aux modèles séquentiels<sup>8</sup>, sont souvent adoptés par les petites autorités de planification des transports pour leur facilité, leur rapidité et leur faible coût d’implémentation ainsi que leur fine échelle d’analyse (Gutiérrez et al., 2011).

Une revue de littérature a montré l’existence de deux grandes familles de variables explicatives :

<sup>8</sup> Les modèles séquentiels décomposent, en plusieurs étapes, la modélisation des choix de déplacement. C’est le cas du modèle à quatre étapes (Bonnell, 2004).

- **Variables exogènes :**
  - *Variables économiques et sociodémographiques* : nombre de résidents, d'emplois, revenu moyen, âge... (Frank and Pivo, 1994; Gutiérrez et al., 2011; Parsons Brinckerhoff Quade & Douglas, Inc., 1993).
  - *Variables liées à l'environnement de la station* : accessibilité de la station, existence des passages piétons dans la zone, structure du réseau routier environnant, etc. (Ewing and Cervero, 2001).
- **Variables endogènes :**
  - *Variables liées aux services de transport offerts par la station* : modes de transport, tarification, type de station (terminale, de correspondance, etc.), nombre de lignes, nombre de correspondances pour aller au centre-ville...(Chu, 2004; Frank and Pivo, 1994).

Ainsi, un modèle direct sous sa forme la plus générale peut donc s'écrire :

$$U_{(Z_i, Pnt_{desserte_k})} = \vec{\alpha} * \vec{V}_{Z_i, Indiv} + \vec{\beta} * \vec{V}_{Pnt_{desserte_k}} + \vec{\theta} * \vec{V}_{Z_i, ConfigSpatiale} + \gamma \quad (i)$$

Où  $U_{(Z_i, Pnt_{desserte_k})}$  est le flux *potentiel* du point de desserte ayant pour origine ou pour destination la zone  $Z_i$ .  $\vec{V}_{Z_i, Indiv}$  est l'ensemble des variables explicatives relatives aux caractéristiques économiques et sociodémographiques de  $Z_i$ .  $\vec{V}_{Pnt_{desserte_k}}$  est l'ensemble des variables explicatives relevant des caractéristiques du point de desserte.  $\vec{V}_{Z_i, ConfigSpatiale}$  est l'ensemble des variables explicatives se référant à la configuration spatiale autour de  $Pnt_{desserte_k}$ .

En introduisant la relation (ii) au modèle (2) on obtient le nouveau modèle :

$$\| \vec{C}_{i, Pnt_{desserte_k}} \| = \sum_{i=1}^m d_{(Z_i, Pnt_{desserte_k})} \left[ \frac{\gamma + \vec{\alpha} * \vec{V}_{Z_i, Indiv} + \vec{\beta} * \vec{V}_{Pnt_{desserte_k}} + \vec{\theta} * \vec{V}_{Z_i, ConfigSpatiale}}{\sum_1^m (\gamma + \vec{\alpha} * \vec{V}_{Z_i, Indiv} + \vec{\beta} * \vec{V}_{Pnt_{desserte_k}} + \vec{\theta} * \vec{V}_{Z_i, ConfigSpatiale})} \right] \quad (3)$$

Dans ce qui suit, nous allons adopter cette formulation générale pour l'estimation de la longueur des connecteurs TC.

### 3.3.2 ADAPTATION DE LA MÉTHODOLOGIE DE CODIFICATION DES CONNECTEURS AUX DONNÉES DISPONIBLES

#### 3.3.2.1 CHOIX DU MODÈLE DE RÉGRESSION

Plusieurs auteurs (Cervero and Kockelman, 1997; Frank and Pivo, 1994; Gutiérrez et al., 2011; Kuby et al., 2004; Parsons Brinckerhoff Quade & Douglas, Inc., 1993) ont montré l'existence d'une corrélation positive entre la densité de la population et des emplois autour d'une station et la fréquentation de cette dernière. Dans ce sens, Kuby (Kuby et al., 2004) a estimé dans une méta-analyse sur 268 stations de 9 villes aux États-Unis qu'il y a 9 nouveaux passagers pour chaque 100 nouveaux résidents autour d'une station LRT – Light Rail Transit - et 2 nouveaux passagers pour chaque 100 nouveaux emplois autour.

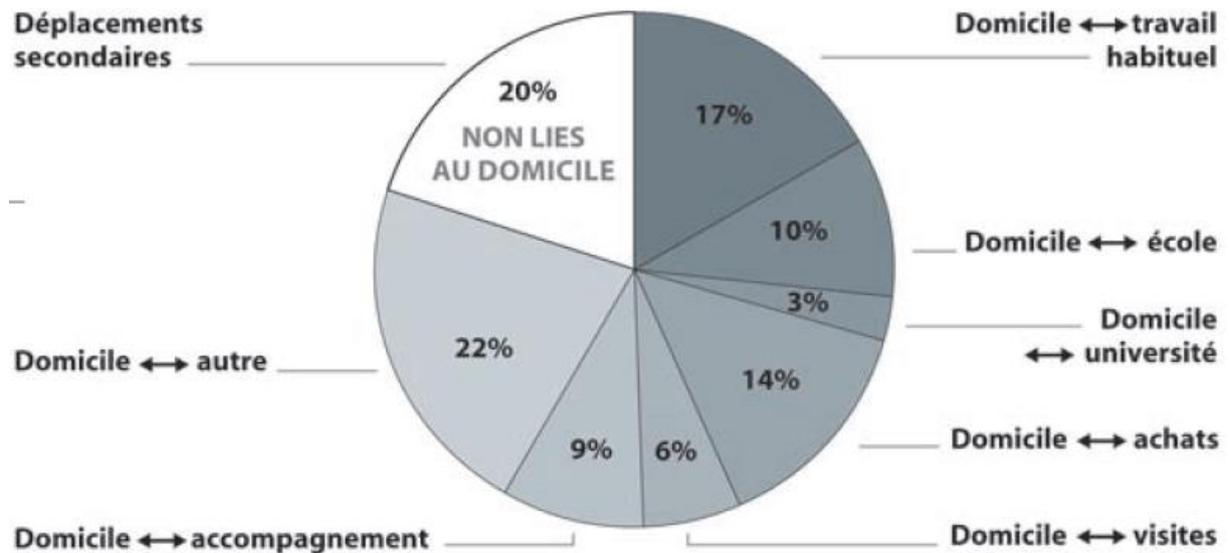
Ces deux variables – population et emplois - nous semblent faciles à collecter et à prédire au fil des ans et peuvent donc être utilisées comme variables explicatives dans un modèle de régression. Une telle relation, sous sa forme générale, peut, donc, s'écrire:

$$U_{(Z_i, Pnt_{desserte_k})} = \gamma + \alpha * Pop_{(Z_i, Pnt_{desserte_k})} + \beta * Emp_{(Z_i, Pnt_{desserte_k})} \quad (ii)$$

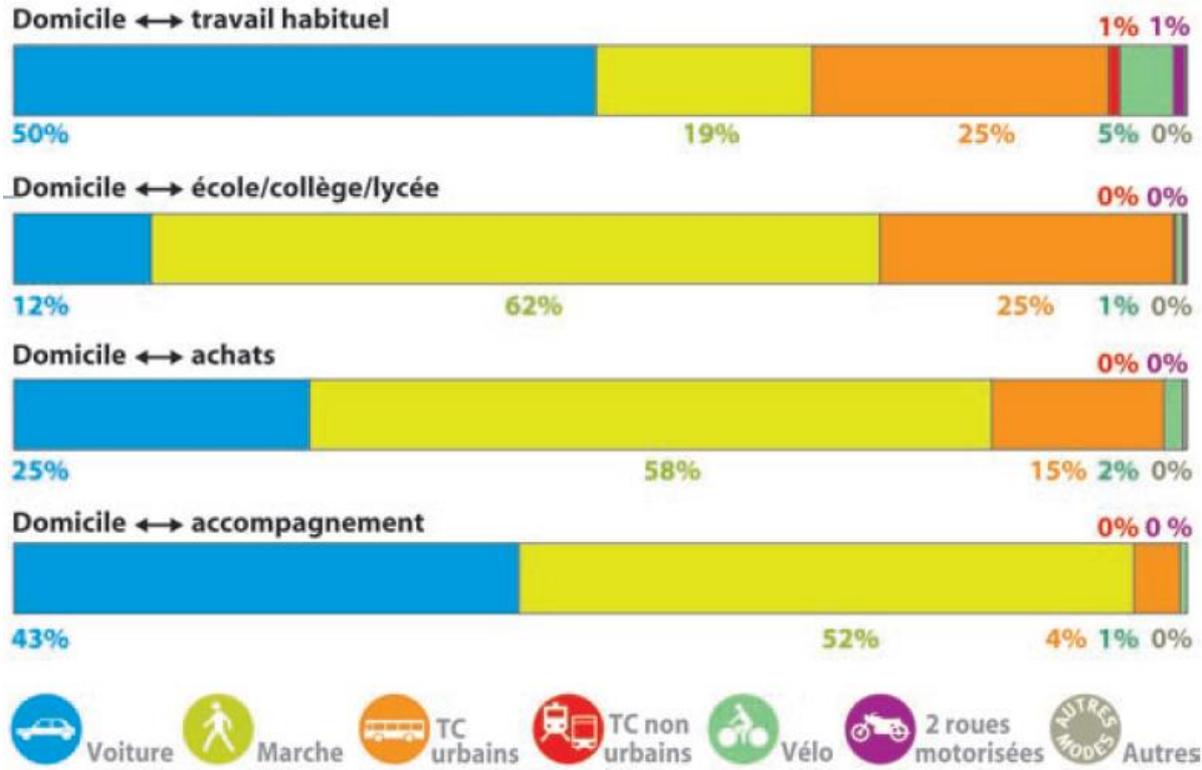
Où  $Pop_{(Z_i, Pnt_{desserte_k})}$  est la population de la zone  $Z_i$  à proximité du point de desserte  $Pnt_{desserte_k}$ .  $Emp_{(Z_i, Pnt_{desserte_k})}$  sont les emplois de la zone  $Z_i$  à proximité de  $Pnt_{desserte_k}$ .

Dans ce qui suit, nous allons nous intéresser au motif de déplacement ‘Domicile-Travail’. Pour ce motif particulier, la calibration de la relation (ii) doit aboutir à une valeur faible de  $\gamma$ . C’est-à-dire que les zones  $Z_i$  ne comportant ni population ni emplois ne doivent pas être capables de générer des usagers potentiels de transport en commun. Cette conclusion est moins vraie pour les autres motifs comme ‘Loisir’, ‘Accompagnement’ ou autres.

En effet, l’Enquête Ménages-Déplacements de 2006 (SYTRAL and Agence d’Urbanisme de Lyon, 2011) a montré que presque 80% des déplacements des résidents de Lyon-Villeurbanne sont liés au domicile. Dans ces 80%, le motif ‘Domicile-Travail’ représente 17% des déplacements dont le quart est réalisé en transports en commun (voir Graphique 5 et Graphique 6). Le motif ‘Domicile-Accompagnement’, par exemple, représente 9% des déplacements des résidents de Lyon-Villeurbanne et utilise minoritairement les transports collectifs (4% seulement des déplacements) (voir Graphique 5 et Graphique 6).



GRAPHIQUE 5 : RÉPARTITION DES DÉPLACEMENTS DES RÉSIDENTS DE LYON-VILLEURBANNE SELON LE MOTIF (SYTRAL AND AGENCE D’URBANISME DE LYON, 2011)



GRAPHIQUE 6 : PART MODALE SELON LE MOTIF DE DÉPLACEMENT SUR LYON-VILLEURBANNE (SYTRAL AND AGENCE D'URBANISME DE LYON, 2011)

Ainsi, pour éviter de générer des connecteurs propres à chaque motif de déplacement, nous allons appliquer les longueurs calculées sur la base de l'équation (ii) à tous les motifs.

La relation (ii) devient :

$$U_{(Z_i, Pnt_{desserte_k})} = \alpha * Pop_{(Z_i, Pnt_{desserte_k})} + \beta * Emp_{(Z_i, Pnt_{desserte_k})} \quad (iii)$$

En intégrant la relation (iii) au modèle (ii) on obtient :

$$\| \tilde{C}_{I, Pnt_{desserte_k}} \| = \frac{\sum_{i=1}^m d_{(Z_i, Pnt_{desserte_k})} * (\alpha * Pop_{(Z_i, Pnt_{desserte_k})} + \beta * Emp_{(Z_i, Pnt_{desserte_k})})}{\sum_1^m (\alpha * Pop_{(Z_i, Pnt_{desserte_k})} + \beta * Emp_{(Z_i, Pnt_{desserte_k})})} \quad (4)$$

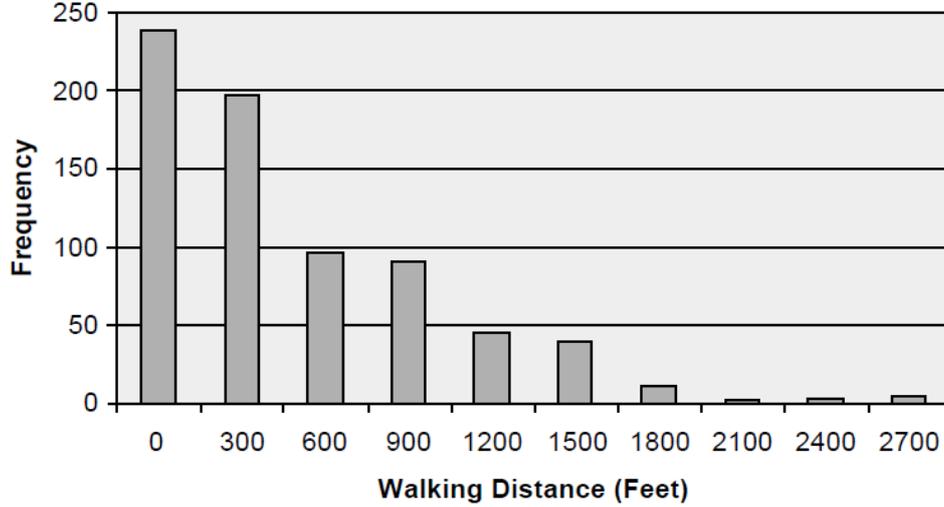
Soit  $\tau = \alpha/\beta$ , (4) devient :

$$\| \tilde{C}_{I, Pnt_{desserte_k}} \| = \sum_{i=1}^m d_{(Z_i, Pnt_{desserte_k})} * \left[ \frac{\tau * Pop_{(Z_i, Pnt_{desserte_k})} + Emp_{(Z_i, Pnt_{desserte_k})}}{\sum_1^m (\tau * Pop_{(Z_i, Pnt_{desserte_k})} + Emp_{(Z_i, Pnt_{desserte_k})})} \right] \quad (4)$$

$\tau$  traduit la différence des comportements de mobilité des actifs du reste de la population.

### 3.3.2.2 DISCUSSION AUTOUR DE LA RELATION DE RÉGRESSION

Une enquête menée en Floride du Sud a montré que la fréquentation des stations de transport en commun diminue en fonction de la distance de marche à pied parcourue vers cette station.



GRAPHIQUE 7 : EFFECTIFS DES USAGERS ACCEDANT AUX STATIONS EN FONCTION DE LA DISTANCE DE MARCHE-À-PIED (ZHAO ET AL., 2003)

Ainsi les effets de  $Pop_{(Z_i, Pnt_{desserte_k})}$  et  $Emp_{(Z_i, Pnt_{desserte_k})}$  sur le nombre d'usagers induits diffèrent en fonction de  $d_{(Z_i, Pnt_{desserte_k})}$  : distance entre la zone  $Z_i$  et le point de desserte  $Pnt_{desserte_k}$ .

La relation (iii) utilisent des coefficients  $\alpha$  et  $\beta$  constants pour toutes les zones  $Z_i$ . C'est-à-dire que toutes les zones sont supposées avoir le même poids dans l'estimation de la régression. Autrement dit, on suppose que l'effet de 100 nouveaux résidents sur le nombre d'usagers est le même à 100m, à 500m ou à 1000m d'une station. De même pour les emplois. Il s'agit de l'hypothèse de stabilité spatiale des coefficients.

### 3.3.3 PRISE EN COMPTE DE L'EFFET DISTANCE

Pour outrepasser l'hypothèse de stabilité spatiale, plusieurs modèles spatiaux travaillent avec des fonctions décroissantes de pondération, dites '*Distance Decay Functions*' (Daniels and Mulley, 2011; El-Geneidy et al., 2014; Gutiérrez et al., 2011; Hsiao et al., 1997; O'Sullivan and Morrall, 1996; Zhao et al., 2003). Ainsi, les zones les plus proches ont plus de poids dans le calcul de la régression que les zones éloignées de la station. Cette nouvelle relation peut être formalisée de la façon suivante :

$$U_{(Z_i, Pnt_{desserte_k})} = \alpha * f_{pop} \left( Pnt_{desserte_k}; d_{(Z_i, Pnt_{desserte_k})} \right) * Pop_{(Z_i, Pnt_{desserte_k})} + \beta * f_{emp} \left( Pnt_{desserte_k}; d_{(Z_i, Pnt_{desserte_k})} \right) * Emp_{(Z_i, Pnt_{desserte_k})} \quad (iv)$$

Où  $f_{pop/emp} \left( Pnt_{desserte_k}; d_{(Z_i, Pnt_{desserte_k})} \right)$  est une fonction de pondération décroissante en fonction de  $d_{(Z_i, Pnt_{desserte_k})}$ . Elle dépend de la nature de la ligne desservant le point de desserte (Bus, Tramway, Métro, Train) et s'applique aux caractéristiques de la zone  $Z_i$ . Pour simplifier la notation, on pose  $f_{pop/emp} \left( Pnt_{desserte_k}; d_{(Z_i, Pnt_{desserte_k})} \right) = f_{pop/emp}$ . On

remarque que cette formalisation suppose un effet différencié de la distance sur les composantes Population et Emplois.

Ainsi le modèle (4) devient :

$$\|\tilde{C}_{I,Pnt_{desserte_k}}\| = \sum_{i=1}^m d_{(Z_i,Pnt_{desserte_k})} \left[ \frac{\tau * f_{pop} * POP_{(Z_i,Pnt_{desserte_k})} + f_{emp} * Emp_{(Z_i,Pnt_{desserte_k})}}{\sum_1^m \tau * f_{pop} * POP_{(Z_i,Pnt_{desserte_k})} + f_{emp} * Emp_{(Z_i,Pnt_{desserte_k})}} \right] \quad (5)$$

Pour déterminer l'expression de ces fonctions, une revue de littérature a été menée et a permis d'identifier les fonctions suivantes:

	<i>Fonction 'tout ou rien'</i>	<i>Fonction 'dégressive'</i>	<i>Fonction 'combinée'</i>
<i>Distance à vol d'oiseau</i>	Méthode du tampon	Méthode de pondération	Méthode de pondération délimitée
<i>Distance réseau</i>	Méthode du tampon déformé	Méthode de pondération	Méthode de pondération délimitée

TABLE 1 : MÉTHODOLOGIE DE PONDÉRATION EN FONCTION DE LA DISTANCE

Pour calculer la distance  $d_{(Z_i,Pnt_{desserte_k})}$ , trois des six méthodes font appel à la notion simplificatrice de la distance euclidienne dite 'distance à vol d'oiseau', les trois autres méthodes utilisent la distance viaire dite 'distance réseau'.

### 3.3.3.1 LA FONCTION 'TOUT OU RIEN'

Elle reprend le principe d'une zone d'attractivité ou de chalandise limitée autour d'un arrêt de transport en commun. Tous les individus contenus dans cette zone ont la même probabilité d'accéder à la station et par la suite le même poids. En dehors de ce périmètre, aucun usager n'est induit. Mathématiquement elle est exprimée par la fonction binaire :

$$f_{pop}(Pnt_{desserte_k}; d_{(Z_i,Pnt_{desserte_k})}) = \begin{cases} 1, & \text{sur } Z_i \cap Ch_{Pnt_{desserte_k},pop} \\ 0, & \text{ailleurs} \end{cases}$$

$$f_{emp}(Pnt_{desserte_k}; d_{(Z_i,Pnt_{desserte_k})}) = \begin{cases} 1, & \text{sur } Z_i \cap Ch_{Pnt_{desserte_k},emp} \\ 0, & \text{ailleurs} \end{cases}$$

Où  $Ch_{Pnt_{desserte_k},pop/emp}$  est la zone de chalandise d'une station donnée. Elle est définie par une distance (ou un temps) maximale d'accès, par exemple, 300m autour de la station ou 5min de marche à pied. Pour simplifier la formalisation, on suppose que  $Ch_{Pnt_{desserte_k},pop} = Ch_{Pnt_{desserte_k},emp}$ .

En fonction de la distance utilisée – euclidienne ou viaire -  $Ch_{Pnt_{desserte_k}}$  adopte une forme circulaire dont le centre est la station ou un polygone polymorphe contenu dans un cercle dont le rayon est la distance maximale à vol d'oiseau (Figure 17).



FIGURE 17 : DIFFÉRENCE ENTRE AIRES DE CHALANDISE POUR UNE MÊME STATION SELON LA MÉTHODOLOGIE DE CALCUL DE LA DISTANCE (SZU-HAN CHEN, 2013)

Le choix de cette distance dépend de plusieurs facteurs :

- Le mode étudié : on suppose que  $Ch_{Bus} \subset Ch_{Tram} \subset Ch_{Métro} \subset Ch_{Train}$ .
- Le tissu urbain de l'iris contenant la station : continu, discontinu, dense, etc.
- Le nombre et le type de stations contenues ou à proximité de l'iris : chevauchement des aires de service, concurrence entre modes, etc.

<i>Mode</i>	<i>Ch<sub>s</sub></i>
Bus	300m
Tramway	320m
Métro	400m
Train	650m

TABLE 2 : EXEMPLE DE DISTANCES MOYENNES DE RABATTEMENT À PIED DANS LE GRAND LYON (AGENCE D'URBANISME DE LYON ET AL., 2013)

Dans la suite de cette étude, nous allons tester plusieurs distances (voir § 3.3.4).

La méthode ‘Tout ou rien’ a la particularité d’être facile et rapide à implémenter, toutefois, elle ignore tous les usagers extérieurs à son périmètre d’application et suppose toujours une stabilité spatiale des coefficients à l’intérieur de la zone de chalandise.

Pour cette raison, la méthode ‘tout ou rien’ surestime souvent la population servie pour une station donnée (Biba et al., 2010; Guerra et al., 2011; Zhao et al., 2003).

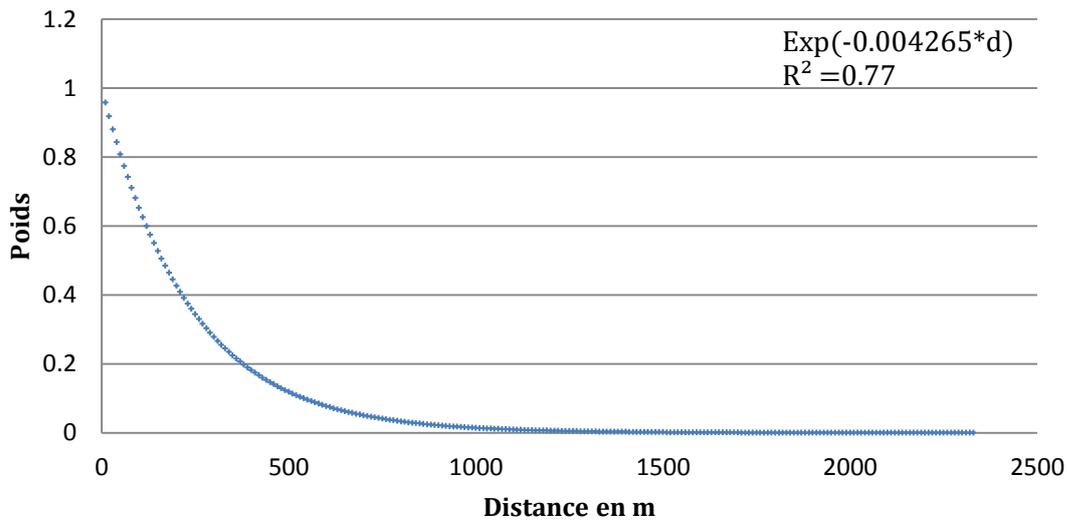
### 3.3.3.2 LA FONCTION ‘DÉGRESSIVE’

Cette fonction vient corriger l’hypothèse de stabilité spatiale de la fonction ‘Tout ou rien’ et élargir son domaine d’application au-delà d’une aire de chalandise. Elle stipule que l’attractivité d’une station diminue avec la distance. Ainsi, les individus habitant ou travaillant à proximité d’une station de transport ont plus de probabilité d’utiliser cette station que le reste. Voici quelques fonctions puisées dans la littérature :

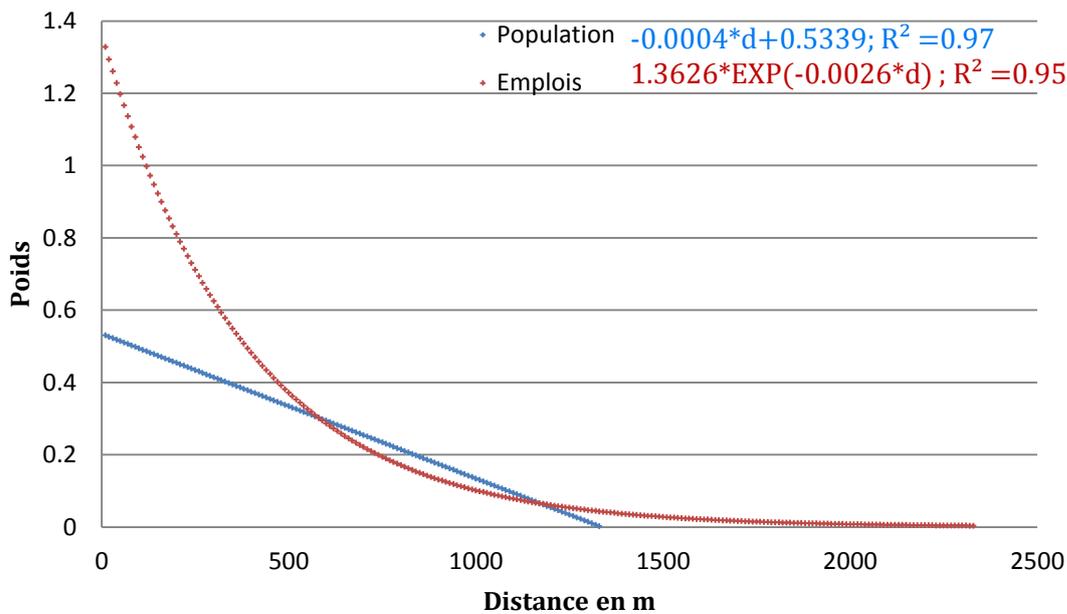
**Fonction dégressive**

<b>Agrégee</b>	$f_{pop/emp}(s; d_{(z_i,s)}) = 1.0126 * \exp(-0.004265 * d_{(z_i,s)}) \quad \text{sur } Z_i, \forall i$
<b>Désagrégée</b>	$\begin{cases} f_{pop}(s; d_{(z_i,s)}) = -0.0004 * d_{(z_i,s)} + 0.5339 \\ f_{emp}(s; d_{(z_i,s)}) = 1.3626 * \exp(-0.0026 * d_{(z_i,s)}) \end{cases} \quad \text{sur } Z_i, \forall i$

TABLE 3 : QUELQUES FONCTIONS DEGRESSIVES (ZHAO ET AL., 2003)



GRAPHIQUE 8 : FONCTION DEGRESSIVE AGRÉGÉE SELON ZHAO (ZHAO ET AL., 2003)



GRAPHIQUE 9 : FONCTION DÉGRESSIVE DÉSAGRÉGÉE SELON GUTIÉRREZ (GUTIÉRREZ ET AL., 2011)

La fonction agrégée a été calée sur des données issues d’une enquête à bord des lignes de transport en commun (Zhao et al., 2003). Cette enquête a été menée dans la Floride du sud en 2000. Elle a recueilli, entre autres, la localisation exacte, en adresse, des usagers des lignes en question (voir Annexe.2). La fonction désagrégée a été calée sur la base des données d’une enquête de mobilité du métro de Madrid en 2004. 17000 déplacements ont été géolocalisés pour calculer les distances à la station (Gutiérrez et al., 2011).

À l’inverse de la fonction ‘Tout ou rien’, la fonction ‘dégressive’ prend en considération tous les individus – dans notre cas, population et emplois - contenus dans l’iris  $I$  concerné. Elle associe à chaque sous-zone un poids en fonction de son éloignement de la station. Ainsi, les zones éloignées de la station participent peu à la génération d’usagers (Graphique 8 et Graphique 9). Cette fonction rend plus fidèlement compte du comportement des individus que la fonction ‘Tout ou rien’, toutefois son implémentation nécessite un temps et une capacité de calcul plus importants.

### 3.3.3.3 LA FONCTION ‘COMBINÉE’

Cette fonction combine la logique de la fonction dégressive et la rapidité d’implémentation de la fonction ‘Tout ou rien’. Elle est le produit de ces deux fonctions :

<i>Fonction combinée</i>	
<i>Agrégée</i>	$\begin{cases} f_{pop/emp}(s; d_{(Z_i,s)}) = 1.0126 * \exp(-0.004265 * d_{(Z_i,s)}), & \text{sur } Z_i \cap Ch_s \\ f_{pop/emp}(s; d_{(Z_i,s)}) = 0, & \text{ailleurs} \end{cases}$
<i>Désagrégée</i>	$\begin{cases} f_{pop}(s; d_{(Z_i,s)}) = -0.0004 * d_{(Z_i,s)} + 0.5339, & \text{sur } Z_i \cap Ch_s \\ f_{pop}(s; d_{(Z_i,s)}) = 0, & \text{ailleurs} \end{cases}$ $\begin{cases} f_{emp}(s; d_{(Z_i,s)}) = 1.3626 * \exp(-0.0026 * d_{(Z_i,s)}), & \text{sur } Z_i \cap Ch_s \\ f_{emp}(s; d_{(Z_i,s)}) = 0, & \text{ailleurs} \end{cases}$

TABLE 4 : QUELQUES FORMES DE LA FONCTION COMBINÉE (GUTIÉRREZ ET AL., 2011; ZHAO ET AL., 2003)

Zhao (Zhao et al., 2003), dans une comparaison entre les fonctions ‘Tout ou rien’ et ‘Combinée’ avec un périmètre de 400m, a estimé que la fonction ‘Combinée agrégée’ évalue à moins de 65% la population ayant accès à une station donnée que la fonction ‘Tout ou rien’.

### 3.3.4 DISTANCE D’ACCESSIBILITÉ

L’implémentation des précédentes fonctions nécessite la définition d’une distance limite (voir la définition de l’accessibilité d’un point de desserte au § 3.1.2). Cette distance permet de filtrer les IRIS ayant accès à un point de desserte donné pour définir par la suite les connecteurs correspondants. Pour les fonctions ‘Tout ou rien’ et ‘Combinée’, cette distance représente le rayon de l’aire de service. Par contre, pour la fonction dégressive, sa définition ne se cantonne pas à une aire de chalandise et s’applique à tous les éléments contenus dans l’IRIS, du moment que cet IRIS a accès à l’arrêt (voir application § 5.1.2).

O’Sullivan et Morrall (O’Sullivan and Morrall, 1996) ont conduit une enquête auprès des autorités planificatrices des transports au Canada et aux États-Unis sur les distances-tampon utilisées, et ont eu comme réponses :

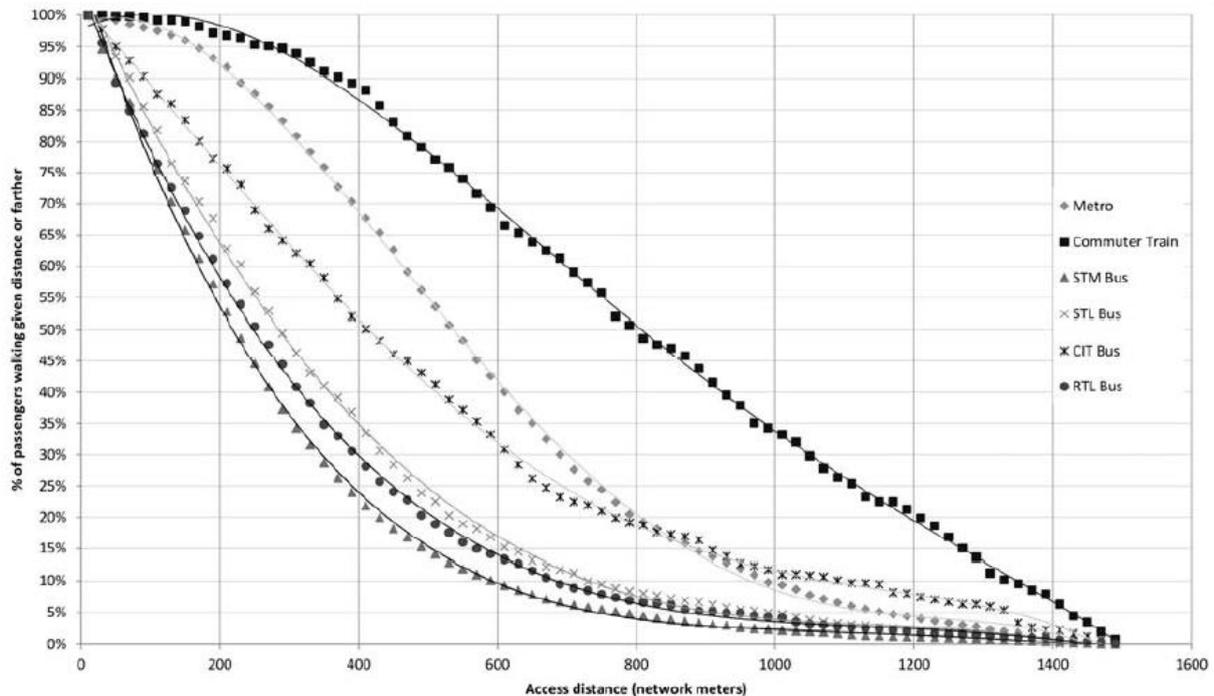
<i>AOT</i>	<i>Distance préconisée</i>	<i>Mode concerné</i>
Calgary Transit (Calgary Canada)	400m	Guide général
B.C Transit (Vancouver Canada)	900m	Light Rail Transit (LRT : Tramway et assimilés)
STCUM (Montréal Canada)	400m-500m	Guide général
Ottawa-Carleton Regional Transit Commission (Ottawa Canada)	400m-600m	Guide général
Toronto Transit Commission (Toronto Canada)	300m	Bus et tramway (Surface Transit)
Edmonton Transportation (Edmonton Canada)	400m	Guide général
Niagara Frontier Transportation Authority (Buffalo USA)	457m	Guide général
Regional Transportation District (Denver USA)	536m	Guide général
New Jersey Transit (Newark USA)	804m	Guide général
Southern Pennsylvania Transportation Authority (Philadelphia USA)	400m 804m	Une zone est ‘bien servie’ à 402m et ‘servie’ à 804m
Sacramento Regional Transit District (Sacramento USA)	609m	Guide général
Metropolitan Transit Development Board (San Diego USA)	538m	Guide général

TABLE 5 : ETENDUE DE L’AIRE DE SERVICE DE CERTAINS MODES DE TRANSPORT EN COMMUN AU CANADA ET AUX ÉTATS-UNIS (O’SULLIVAN AND MORRALL, 1996)

Dans une autre étude, l’analyse des données d’une enquête Origine-Destination menée sur la grande région métropolitaine de Montréal en 2003 (voir Annexe.3) a permis le calcul des distances viaires de rabattement et de diffusion des arrêts des transports en commun (El-Geneidy et al., 2014). Cette distance est calculée sur la base 16014 déplacements liés au domicile. Une géolocalisation des ménages grâce à l’adresse déclarée était nécessaire pour le calcul de la distance viaire. Les résultats de sortie sont tels que :

	All modes	Metro	Train	STM bus	RTL bus	STL bus	CIT bus
<b>Walking distance</b>							
Mean	374.80	564.80	817.98	275.96	314.50	347.01	488.73
Median	294.21	527.14	785.03	213.80	243.16	277.36	401.80
Maximum	1497.60	1496.83	1491.28	1490.49	1486.32	1440.36	1497.60
75th percentile	524.58	730.73	1102.84	371.10	427.20	471.51	653.97
85th percentile	678.34	873.35	1259.41	484.09	556.36	601.05	897.04
SD	308.64	297.37	360.10	245.50	282.00	285.27	370.37
75th percentile + 1.5 SD	987.53	1176.79	1642.99	739.35	850.20	899.41	1209.54

TABLE 6 : STATISTIQUES DE LA DISTANCE DE MARCHE À PIED VERS ET DEPUIS LES ARRÊTS (EL-GENEIDY ET AL., 2014)<sup>9</sup>



GRAPHIQUE 10 : FRÉQUENCE CUMULÉE DU NOMBRE DE PASSAGERS EN FONCTION DE LA DISTANCE DE MARCHE À PIED (EL-GENEIDY ET AL., 2014)

Il est clair que les pratiques de marche à pied diffèrent d'une ville à l'autre et d'un mode à l'autre, toutefois cette distance est souvent comprise entre 0m et 1500m, tous modes confondus.

Par la suite, nous allons tester plusieurs distances d'accessibilité variant entre la médiane et le 9<sup>ème</sup> décile des distances de marche à pied (Graphique 10).

<sup>9</sup> Voir Annexe.3 pour l'organisation des transports en commun de la région de Montréal.

<i>Mode</i>	<i>Distance d'accessibilité</i>
Bus	300m à 700m
Tramway	300m à 700m
Métro	600m à 1000m
Train	800m à 1400m

TABLE 7 : DISTANCE D'ACCESSIBILITÉ EN FONCTION DU MODE DE TRANSPORT

### 3.3.5 ALGORITHME SIMPLIFIÉ D'IMPLÉMENTATION

Enfin, cette nouvelle méthodologie de calcul peut être schématisée sous la forme de l'algorithme suivant :

1. Déterminer tous les points de desserte de l'aire d'étude.
2. Affecter à chaque point de desserte la distance d'accessibilité correspondant à la catégorie de la ligne le desservant.
3. Déterminer les points de desserte accessibles depuis chaque IRIS  $I$ .
4. Dédire les lignes accessibles depuis  $I$ .
5. Grouper les points de desserte en fonction des lignes accessibles desservant ces points.
6. Pour chaque ligne accessible et chaque IRIS, partager la population et les emplois entre les différents points de desserte appartenant à cette même ligne.
7. Pour chaque ligne accessible et chaque IRIS, calculer la distance entre les points de desserte de cette ligne et la population et les emplois de l'IRIS. Il s'agit des connecteurs des points de desserte.
8. Pour chaque ligne accessible et chaque IRIS, calculer la longueur moyenne pondérée des différentes longueurs de connecteurs des points de desserte. Il s'agit des connecteurs moyens des points de desserte.
9. Pour un groupe de points de desserte, attribuer la longueur moyenne récemment calculée au connecteur d'un point de desserte de ce groupe. De préférence, ce point de desserte doit appartenir au point d'arrêt qui contient le moins de points de desserte.
10. Pour chaque point d'arrêt, calculer la longueur moyenne pondérée des différentes longueurs des connecteurs moyens des points de desserte y appartenant.

### 3.3.6 CONCLUSION

Il est évident que l'accès aux transports en commun diffère en fonction de l'éloignement de la station. Une prise en compte de cet éloignement est donc nécessaire dans le modèle de régression. Pour ce faire, nous avons proposé trois fonctions de pondération. La fonction 'Tout ou rien' est la plus simple à implémenter, elle affecte une pondération binaire – 0 ou 1 – et par conséquent, ne s'applique qu'à une partie de l'IRIS. La fonction 'Dégressive' pondère, par contre, tous les individus de l'IRIS, ce qui nous semble plus cohérent avec la définition des connecteurs et leur utilisation (voir § 3.3.1). Cette fonction risque, toutefois, de solliciter plus de temps de calcul que la première. Pour cette dernière raison, la fonction 'Combinée' est définie sur la base des précédentes fonctions de sorte qu'elle combine la simplicité et la pertinence d'implémentation des deux. Le chapitre 5 illustre l'usage de ces fonctions et dresse un comparatif entre elles.

# CHAPITRE 4

---

## 4 DONNÉES INDISPENSABLES À L'IMPLEMENTATION DE LA NOUVELLE MÉTHODOLOGIE

Comme évoqué plus haut, la qualité de la méthodologie précédente tient, entre autres, à la partition  $Z_I = \{Z_i | i \in \llbracket 1; m \rrbracket\}$  de l'IRIS  $I$ . Cette qualité est d'autant plus meilleure que la partition est fine. Dans ce cadre, nous avons testé plusieurs partitions pour fixer notre choix sur l'échelle la plus fine, c'est-à-dire une partition constituée du bâti de résidence et d'activité.

Cette partition requiert la localisation fine de la population et des emplois. En France, l'échelle la plus fine de diffusion des données démographiques est l'IRIS<sup>10</sup>. Nous sommes, donc, dans l'obligation de construire cette partition à partir des bases de données existantes. Dans ce sens, nous nous sommes basés sur la nouvelle base carroyée diffusée par l'Insee ainsi que sur la base BD TOPO<sup>®</sup> de l'IGN et le Recensement Général de la Population de 2010 (RGP 2010). Pour les emplois, un travail préalable au mémoire a été réalisé par le LET pour localiser les activités et estimer le nombre de leurs salariés.

### 4.1 PRÉSENTATION DES DONNÉES D'ENTRÉE DE LA POPULATION

Pour estimer et localiser la population, un recouplement entre plusieurs bases de données est nécessaire. Les données du recensement sont diffusées au niveau des IRIS et permettent de caractériser la population à cette échelle. L'utilisation de cette base seule dans la définition des connecteurs TC oblige à l'agrégation des données au niveau des centroïdes et par conséquent à masquer l'hétérogénéité des conditions d'accès aux transports en commun de la population concernée. Pour contourner cette limite, nous utilisons la nouvelle base de données carroyées. Elle décrit la population fiscale (Définition dans le § 4.1.1) dans un découpage régulier et fin. Enfin la base de données BD TOPO<sup>®</sup> de l'IGN est utilisée pour la localisation fine de cette population.

---

<sup>10</sup> Une autre échelle plus fine de 200m par 200m existe pour la diffusion des données relatives à la population fiscale (voir définition plus bas).

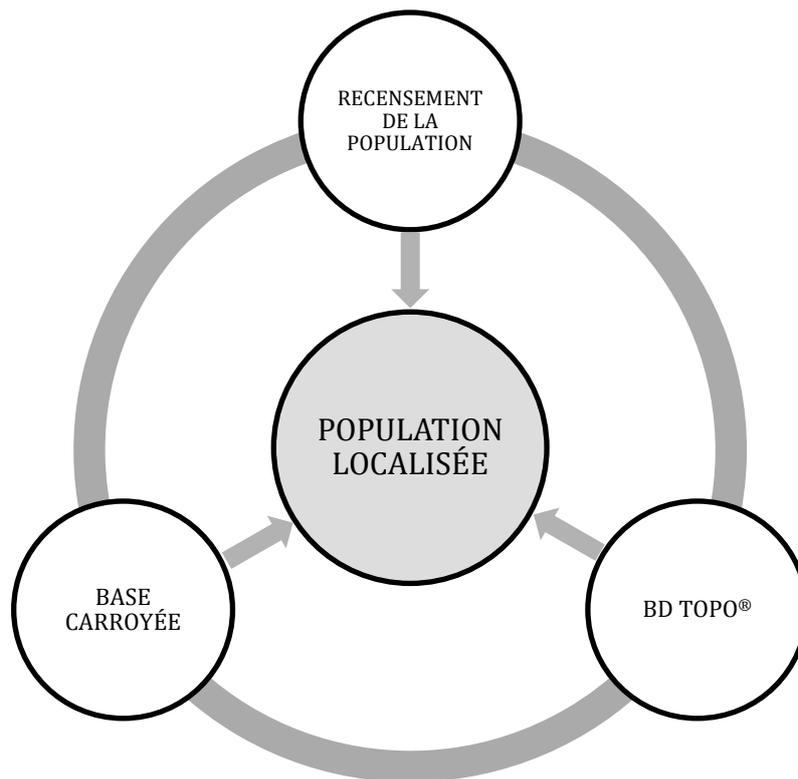


FIGURE 18 : DONNÉES D'ENTRÉE ET DE SORTIE ET LEUR INTERACTION

#### 4.1.1 PRÉSENTATION DE LA BASE CARROYÉE

En novembre 2013, l'Insee a diffusé une base de données carroyées décrivant la population fiscale de la France en 2010 (Insee, 2013). Cette base est établie à partir des fichiers des 'Revenus Fiscaux Localisés' et des déclarations de la taxe d'habitation. Elle se présente sous la forme d'une trame régulière de carreaux de 200m par 200m.

Cette base décrit la population à travers 19 variables économiques et sociodémographiques, allant du nombre de résidents par carré à la somme traitée des revenus fiscaux (Insee, 2013).

Pour respecter les prescriptions relatives à la confidentialité des données, chaque carré doit obligatoirement contenir 11 ménages, sinon il est agrégé aux carreaux adjacents pour former un rectangle avec un nombre suffisant de ménages. Ainsi, la majorité des variables décrivent les résidents du rectangle auquel le carré appartient. Seule la variable 'Ind\_c,' reprenant le nombre de résidents par carré, échappe à cette agrégation (Insee, 2013). De ce fait, l'Insee a mis en ligne les deux bases de données : carroyées contenant l'unique variable 'Ind\_c' et agrégées sous forme de rectangles contenant le reste des variables.

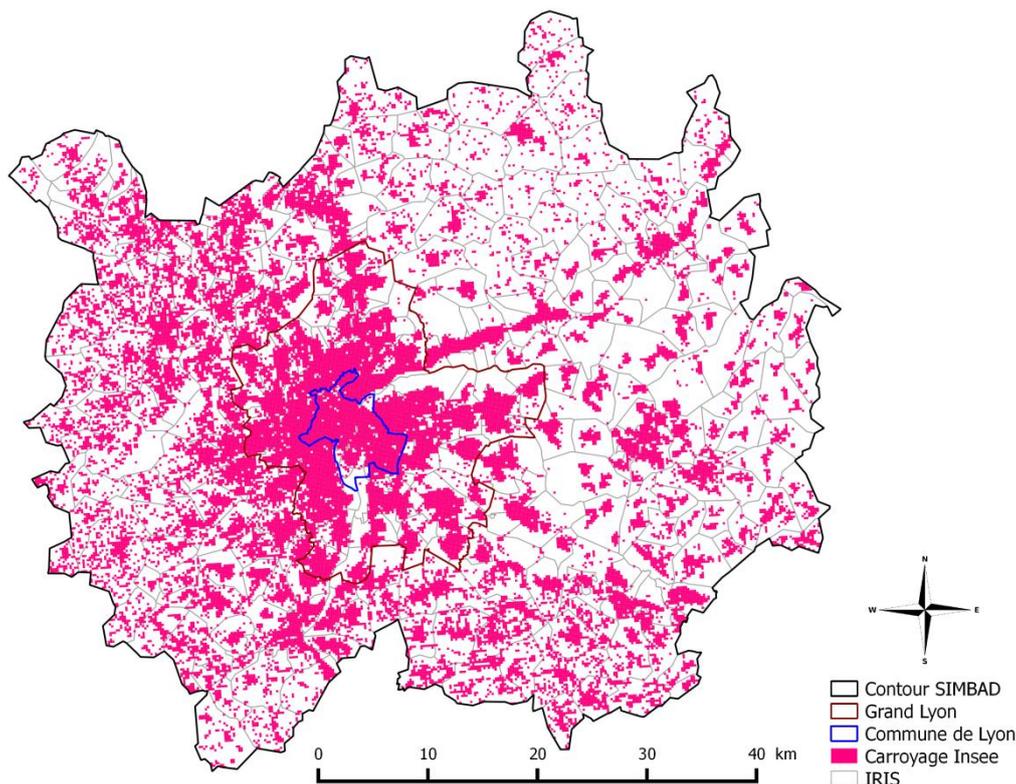


FIGURE 19 : EXTRAIT DE LA BASE CARROYÉE CORRESPONDANT À L'AIRE SIMBAD

Cette base utilise la notion de ‘ménage fiscal’. Au sens de l’Insee, *un ménage fiscal est ‘un ménage constitué par le regroupement des foyers fiscaux<sup>11</sup> répertoriés dans un même logement. Son existence, une année donnée, tient au fait que coïncident une déclaration indépendante de revenus (dite déclaration n°2042) et l’occupation d’un logement connu à la taxe d’habitation (TH)’*. Donc, la population carroyée est la population fiscale localisée par la taxe d’habitation. Par conséquent, cette base ignore certaines catégories de la population dont, par exemple, les étudiants qui sont souvent rattachés au foyer parental. Ainsi, l’utilisation de cette base doit normalement aboutir à une sous-estimation de la population.

En dépit de cette particularité, la base carroyée présente un intérêt indéniable dans le cadre de notre étude. Elle permet souvent une description infra-iris<sup>12</sup> de la population fiscale à travers plusieurs attributs. À présent, on se limite à l’exploitation de la variable ‘*Ind-c*’ reprenant le nombre de résidents au carré, sans exclure pour autant une utilisation future des autres variables : âge, revenus traités, etc.

<sup>11</sup> Selon l’Insee, un foyer fiscal est l’ensemble des personnes inscrites sur une même déclaration de revenus.

<sup>12</sup> Hormis l’hyper-centre lyonnais où le découpage IRIS est plus fin, un IRIS inclut souvent plusieurs carreaux.

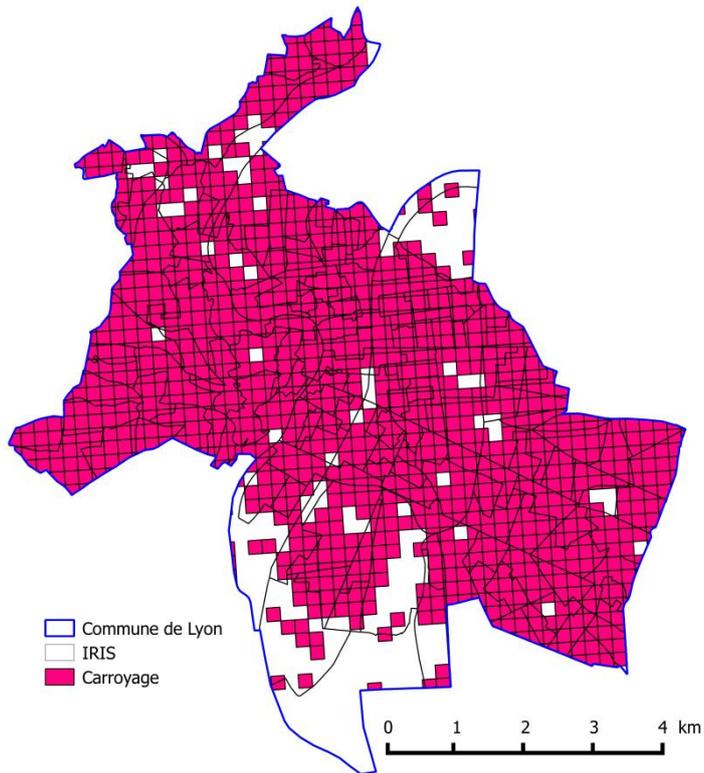


FIGURE 20 : EXTRAIT DE LA BASE CARROYÉE AU NIVEAU DE LA COMMUNE DE LYON

#### 4.1.2 PRÉSENTATION DE LA BD TOPO<sup>®</sup>

La BD TOPO<sup>®</sup> est un produit de l'IGN qui décrit, avec une précision métrique, les éléments du paysage sous forme vectorielle avec une classification thématique (voir Annexe.5). Elle représente, entre autres, le bâti. Ce dernier nous est nécessaire à la localisation de la population.

Nous utilisons le thème 'BATI' de la version 2. Ce thème est divisé en plusieurs sous-thèmes ou classes. La classe 'BATI\_INDIFERENCIE' inclut les logements, les bureaux, les bâtiments hospitaliers, les musées et plusieurs autres établissements.

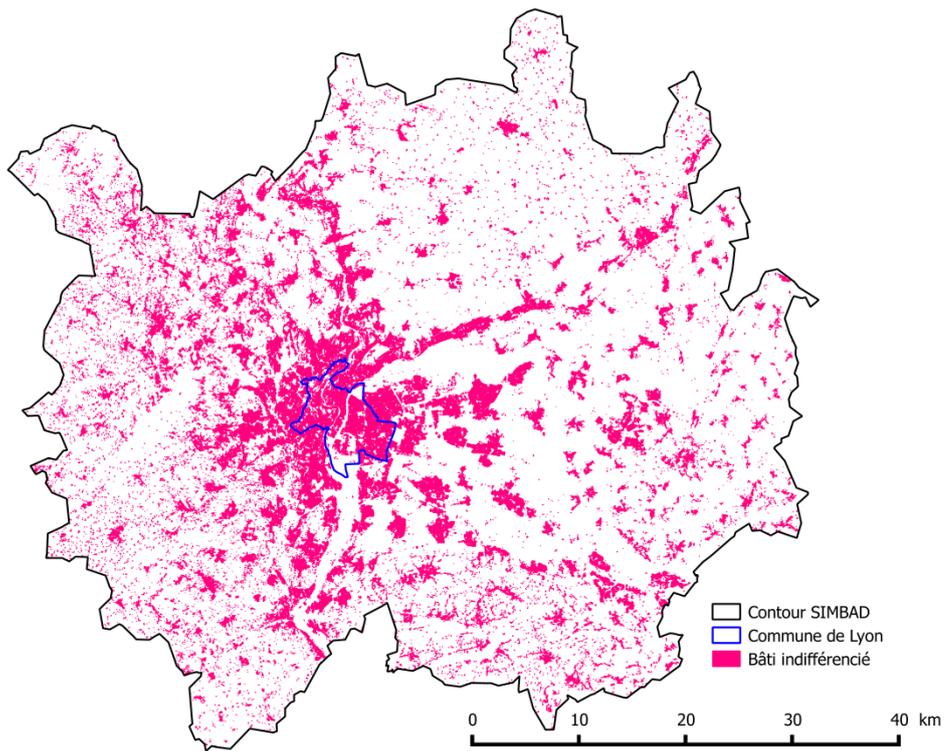


FIGURE 21 : EXTRAIT DE LA BD TOPO® CORRESPONDANT À L'AIRE SIMBAD

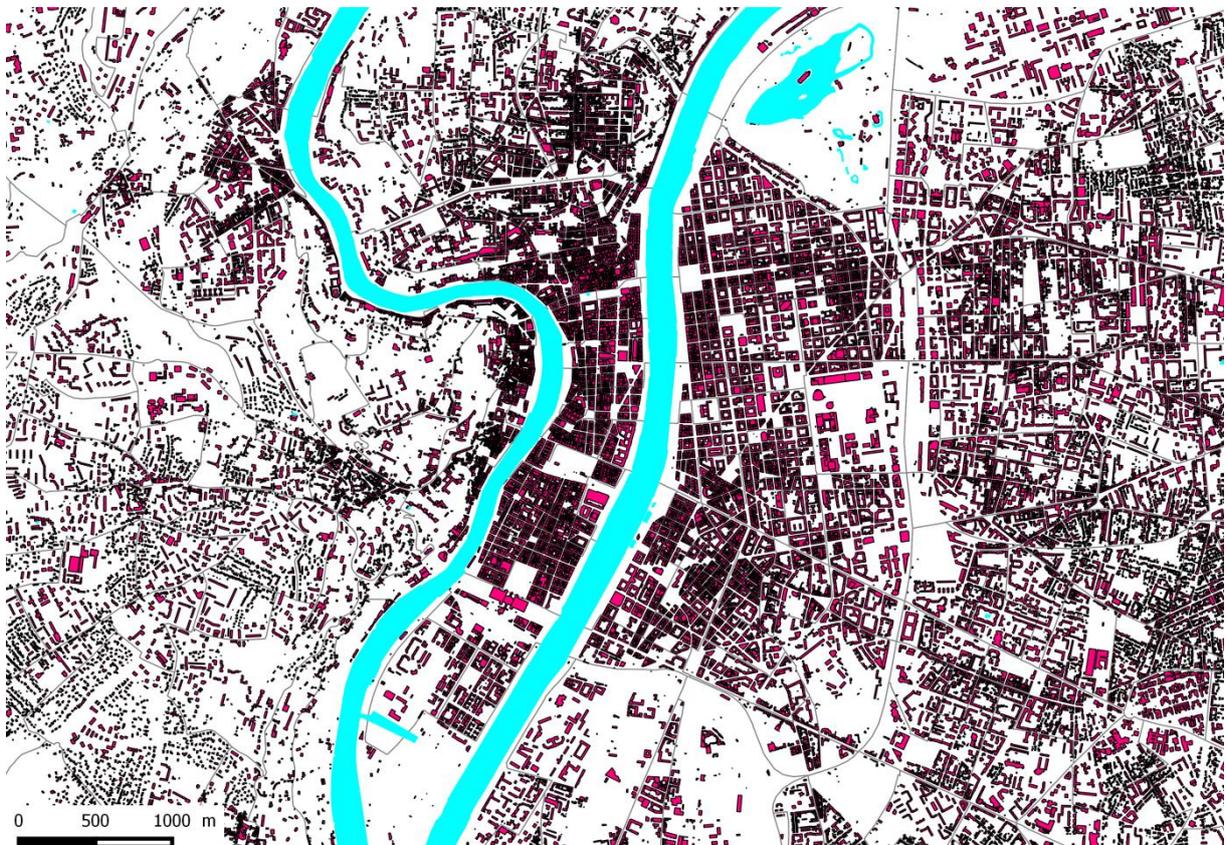


FIGURE 22 : EXTRAIT DE LA BD TOPO® AU NIVEAU DE LYON

Cette classe caractérise les constructions par un identifiant unique et leur affecte une hauteur (IGN, 2008). Elle se présente sous la forme de polygones épousant le contour des toits des bâtiments ou de plusieurs bâtiments adjacents.

<b>ID</b>	<b>PREC_PLANI</b>	<b>PREC_ALTI</b>	<b>ORIGIN_BAT</b>	<b>HAUTEUR</b>	<b>Z_MIN</b>	<b>Z_MAX</b>
BATIMENT0000000240880944	2.5	20	Cadastre	26	193	193
BATIMENT0000000240873754	2.5	20	Cadastre	22	188.7	188.7
BATIMENT0000000240873858	2.5	20	Cadastre	24	191.5	191.5
BATIMENT0000000240874183	2.5	20	Cadastre	24	191.3	191.3
BATIMENT0000000240879481	2.5	20	Cadastre	23	189.8	189.8
BATIMENT0000000240879539	2.5	20	Cadastre	23	189.5	189.5
BATIMENT0000000240879548	2.5	20	Cadastre	23	189.6	189.6
BATIMENT0000000240879586	2.5	20	Cadastre	25	192.2	192.2
BATIMENT0000000240879651	2.5	20	Cadastre	25	191.5	191.5
BATIMENT0000000240880870	2.5	20	Cadastre	21	188	188

TABLE 8 : ATTRIBUTS CONTENUS DANS LA COUCHE BÂTI INDIFFÉRENCIÉ DE LA BD TOPO® (VOIR ANNEXE.5)

Par manipulation SIG, nous calculons la superficie des bâtiments. Il s'agit de l'emprise du polygone et donc d'une superficie estimée et non réelle<sup>13</sup>. Ce champ est indispensable avec l'attribut 'HAUTEUR' pour la détermination de 'la surface développée' (voir 4.3.2.1.1).

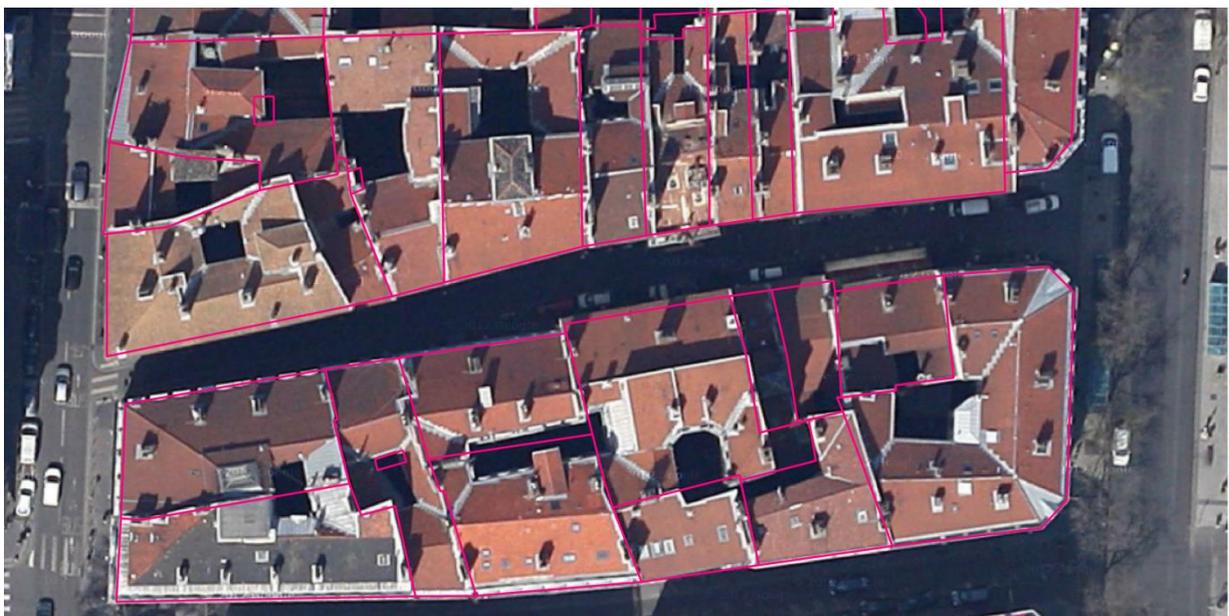


FIGURE 23 : EXEMPLE DE BATIMENTS ISSUS DE LA BD TOPO® (FOND GOOGLEMAPS)

Dans le cadre de notre étude, on s'intéresse au bâti contenant les logements d'habitation. L'utilisation de la classe 'BATI\_INDIFFERENCIE' nécessite, donc, un travail préalable de nettoyage.

<sup>13</sup> Par réelle on entend la superficie des planchers ou de l'emprise au sol du bâtiment.

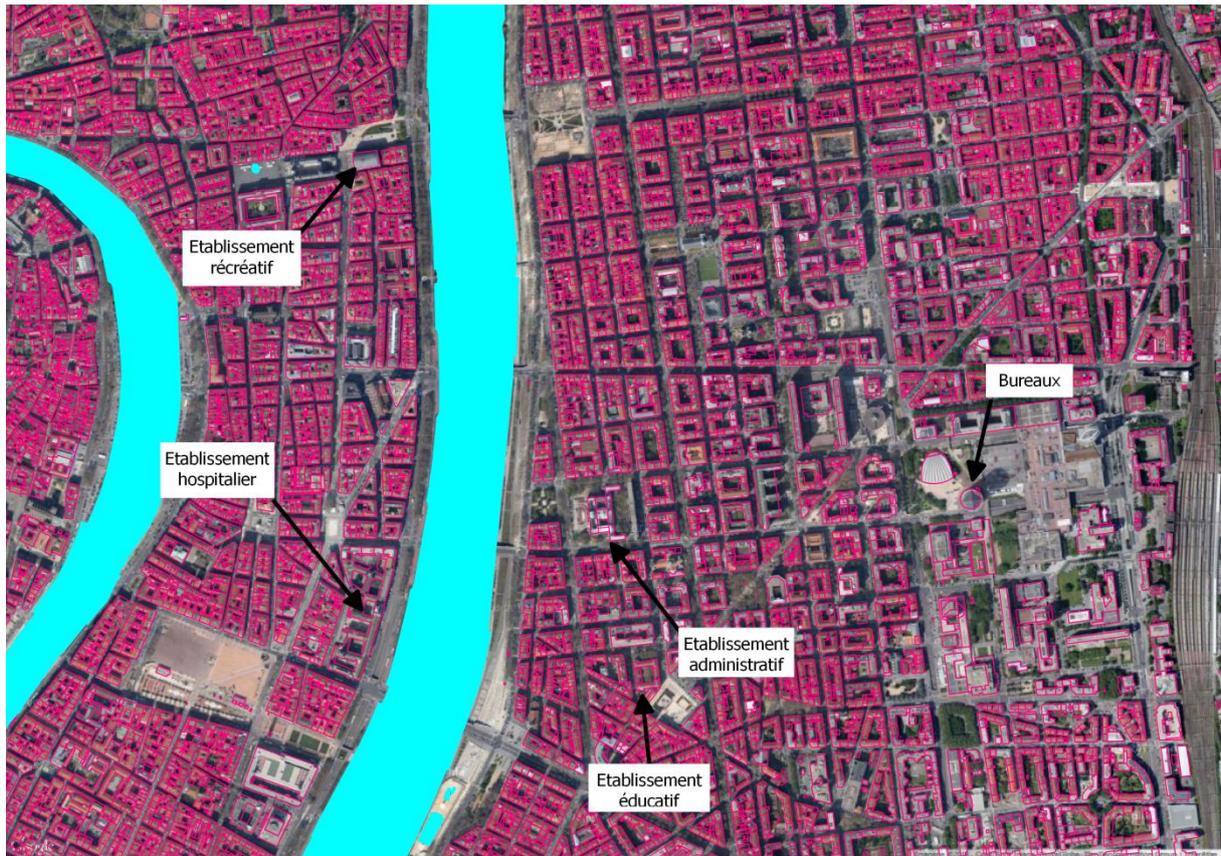


FIGURE 24 : EXEMPLES DE BÂTIMENTS NON RÉSIDENTIELS INCLUS DANS LA CLASSE BÂTI\_INDIFFÉRENCIÉ DE LA BD TOPO® (FOND GOOGLEMAPS)

Pour ne garder que les bâtiments résidentiels, une identification et une suppression du reste des bâtiments est nécessaire. Une première identification automatique, excluant la majorité des établissements scolaires, de santé, administratifs, de loisirs, de transport, militaires..., est réalisée en se basant sur d'autres thèmes compris dans la BD TOPO®.

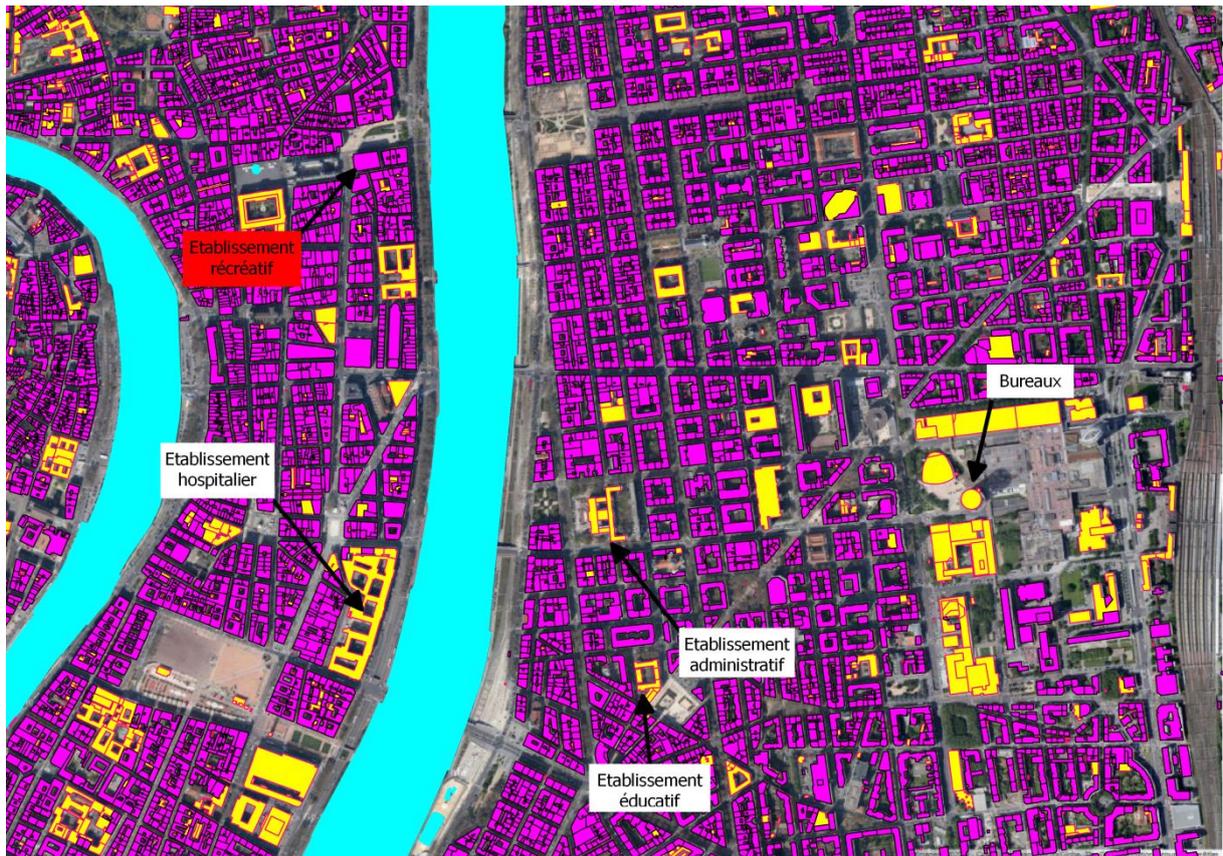


FIGURE 25 : RÉSULTAT D'IDENTIFICATION DES BÂTIMENTS NON RÉSIDENTIELS (FOND GOOGLMAPS)

Cette opération aboutit partiellement à une base avec des bâtiments à caractère résidentiel, bureautique ou mixte. Elle comporte, toutefois, quelques erreurs (voir l'exemple de l'opéra de Lyon qui n'est pas supprimée automatiquement, Figure 25).

Il est ensuite difficile de différencier entre bâtiments résidentiels et bureaux. Un premier travail non consistant d'identification à la main a été réalisé sur le secteur de la Part-Dieu pour supprimer les bâtiments réservés uniquement aux bureaux.

Compte tenu de l'incertitude et des erreurs d'estimation de la hauteur, plusieurs constructions sont codées avec une hauteur nulle. Il s'agit, pour la plupart, de petits abris de jardins ou de places de parkings, mais y figurent, aussi, des bâtiments avec une hauteur réelle non négligeable. On a procédé au traitement de ces entités en fonction de leur superficie, en supprimant les constructions à hauteur nulle avec une surface négligeable (voir Figure 26). Enfin nous avons grossièrement analysé le reste de ces constructions et on leur a affecté une hauteur estimée à partir des bâtiments environnant.



FIGURE 26 : EXEMPLE DE CONSTRUCTIONS CODÉES AVEC UNE HAUTEUR NULLE DANS LA BD TOPO® (FOND BINGMAP)

Une dernière manipulation consiste à identifier les logements avec une hauteur de 2m et à leur affecter une hauteur de 3m correspondant, en général, à la hauteur du rez-de-chaussée.

À l'issue de ces étapes nous disposons d'une base caractérisant les bâtiments à usage résidentiel et bureautique. Elle est dénommée par la suite 'BD TOPO nettoyée' et contient 339 541 bâtiments.

#### 4.1.3 PRÉSENTATION DES DONNÉES DE RECENSEMENT DE LA POPULATION DE 2010

Suite au recensement de 2010, l'Insee a mis à disposition ces données à l'échelle des IRIS ([Lien vers les données et les explications de l'Insee](#)). Ces données, à l'inverse de la base carroyée, caractérisent la population municipale<sup>14</sup> et son évolution. Nous utilisons ces données comme une référence par rapport aux données carroyées.

Cette base géolocalisée de la population, doit être complétée par les données relatives à l'emploi. Dans ce sens, nous avons profité de l'existence de la base de données SIRENE® 2011 traitée et géolocalisée par le LET.

## 4.2 PRÉSENTATION DES DONNÉES D'EMPLOI

<sup>14</sup> Selon l'Insee, 'La population municipale comprend les personnes ayant leur résidence habituelle sur le territoire de la commune, dans un logement ou une communauté, les personnes détenues dans les établissements pénitentiaires de la commune, les personnes sans-abri recensées sur le territoire de la commune et les personnes résidant habituellement dans une habitation mobile recensée sur le territoire de la commune'.

SIRENE<sup>®</sup> pour ‘Système Informatique pour le Répertoire des ENtreprises et de leurs Etablissements’ est un système de gestion du répertoire national d’identification des entreprises et des établissements implantés sur le territoire national. Il est géré par l’Insee. Il se présente sous la forme d’une base de données décrivant à travers plusieurs champs (Annexe.6) les entreprises et les établissements. Cette base est alimentée, en continu, de nouvelles entreprises (Annexe.6). Un travail précédent des équipes du LET a permis la géolocalisation des entités implantées sur l’aire SIMBAD. Il s’agit d’une géolocalisation sur la base de la BD Adresses<sup>®</sup> de l’IGN complétée par des requêtes des serveurs OpenStreetMap. L’effort de géolocalisation a permis le positionnement de 132149 entreprises soit un taux de réussite de 93%.

L’aire SIMBAD recense 143533 entreprises et établissements – les auto-entrepreneurs n’y sont pas inclus – générant 820186 emplois – Somme des emplois arrondis -. Lyon, à elle seule, accueille 312620 emplois – l’Insee affiche un chiffre de 294866 pour 2009 -, soit 38% des emplois.

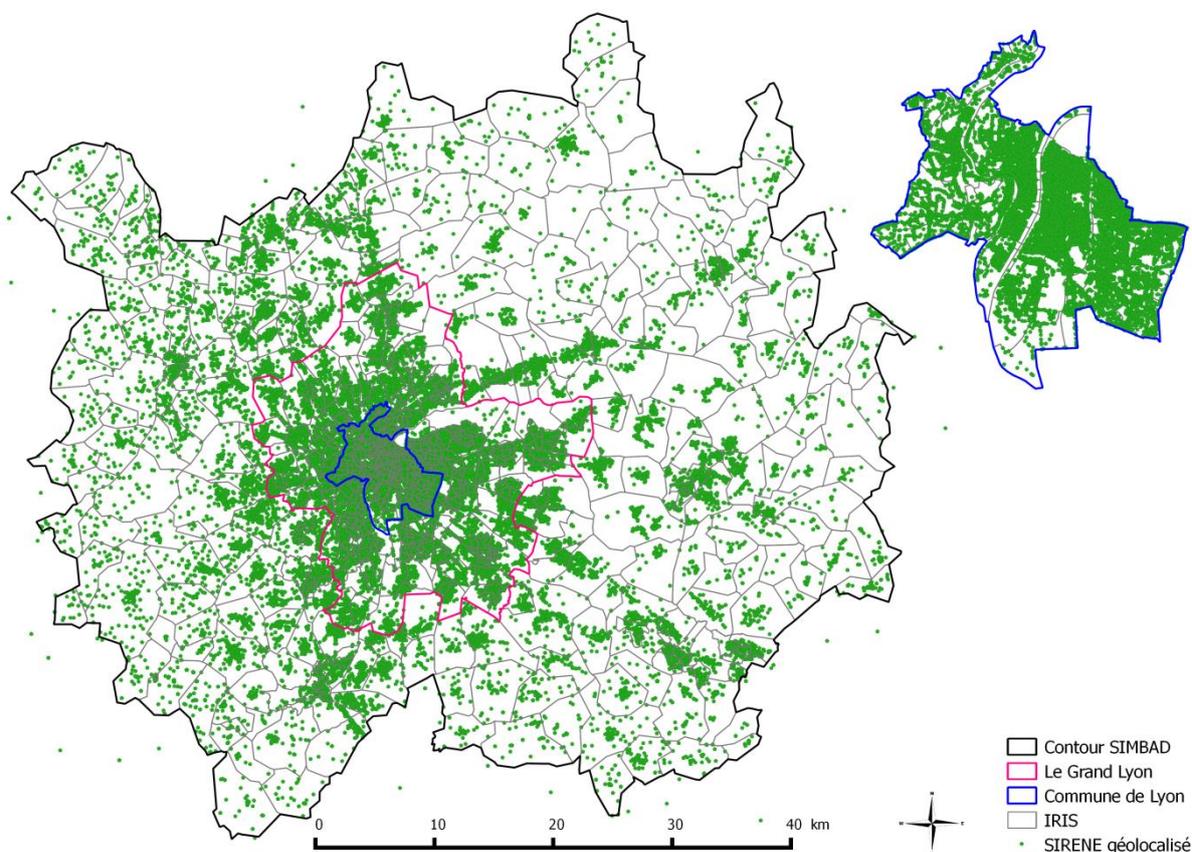
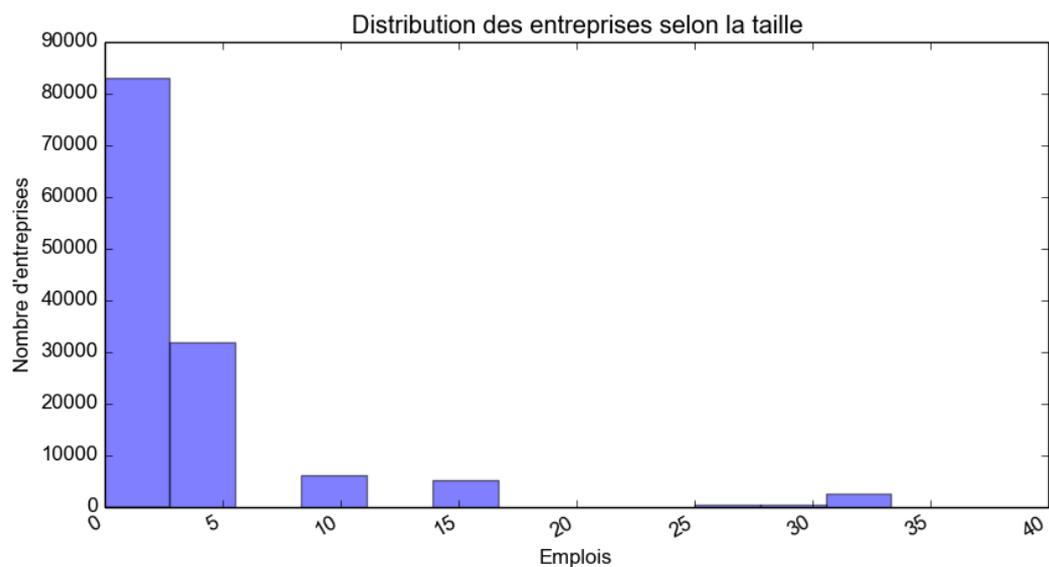


FIGURE 27 : LOCALISATION DES EMPLOIS SUR L’AIRE SIMBAD SELON LA BASE SIRENE<sup>®</sup> TRAITÉE

Les entreprises sont localisées pour 38% aux limites administratives de la commune de Lyon et pour 76% aux limites du Grand Lyon. 63% des entreprises emploient deux personnes et moins et 87% ont 5 salariés et moins.



GRAPHIQUE 11 : DISTRIBUTION DES ENTREPRISES SELON LA TAILLE

Enfin, nous signalons que la qualité de la géolocalisation des établissements SIRENE<sup>®</sup> est moins précise que celle de la population (Figure 28). Cette imprécision au niveau de la localisation peut affecter le calcul de la distance et par conséquent la longueur estimée des connecteurs.

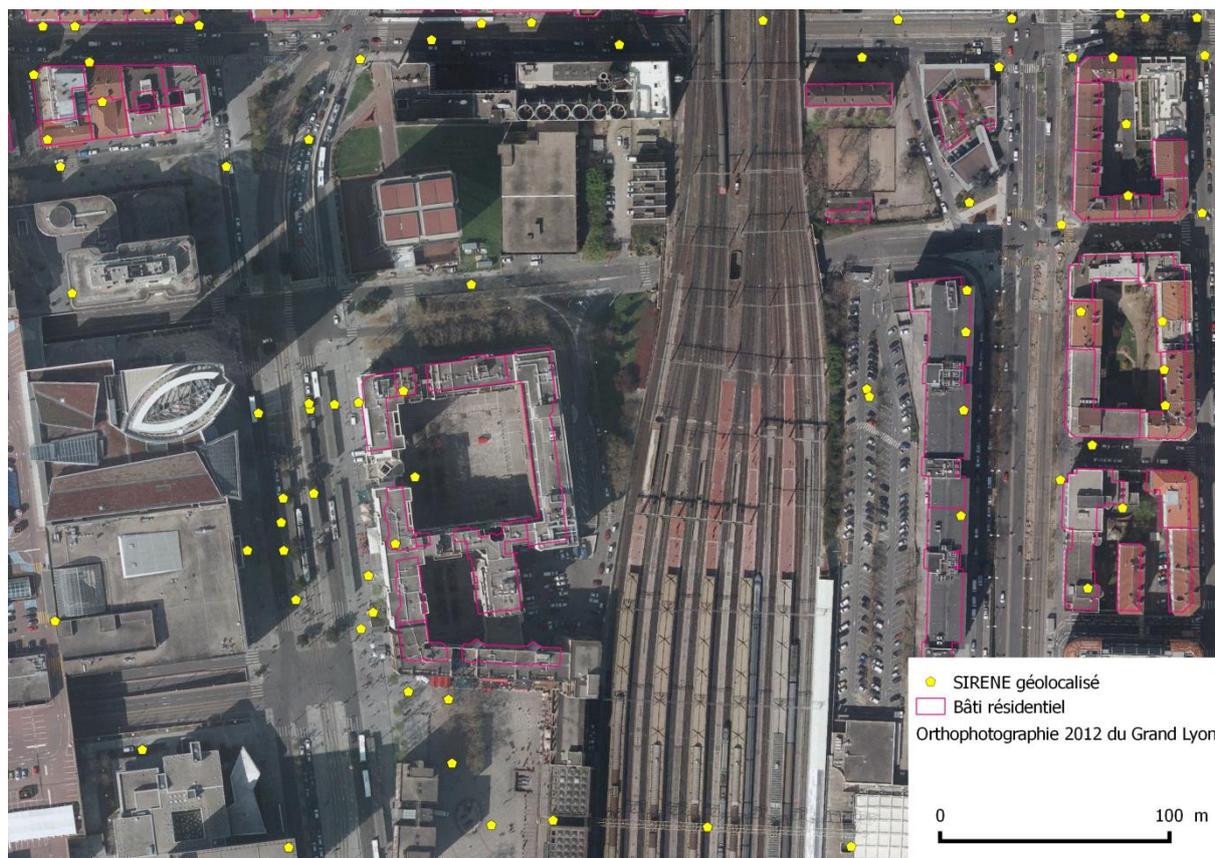


FIGURE 28 : EXEMPLE D'IMPRÉCISION DE LA GÉOLOCALISATION SIRENE<sup>®</sup> AU NIVEAU DE LA PART-DIEU (ORTHOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012)

## 4.3 EXPLOITATION DES DONNÉES D'ENTRÉE

### 4.3.1 EXPLOITATION DES DONNÉES CARROYÉES

L'aire d'étude SIMBAD comporte à-peu-près 27 101 carreaux et 1 768 140 individus. Le recensement de la population de 2010 affecte 1 806 590 individus à cette même aire, soit une sous-estimation de 2% de la base carroyée par rapport au recensement.

La base carroyée se présente sous la forme d'une trame et le zonage d'étude sous la forme du découpage IRIS. L'utilisation, donc, de la base carroyée dans la définition des connecteurs oblige à l'harmonisation de ces zonages. Toutefois et par construction la trame régulière ne tient pas compte des limites administratives. Ainsi, un carreau peut être positionné sur une frontière commune à plusieurs IRIS.

Dans ce qui suit, nous allons tester plusieurs méthodes de répartition des carreaux partagés entre IRIS. La méthode la plus adaptée est celle qui minimise la différence pondérée entre population estimée et population recensée, c'est-à-dire l'erreur pondérée par la population. Soit  $pop_{estim\_iris}$  et  $pop_{10\_iris}$  les populations estimée et recensée respectivement au niveau de l'IRIS,  $Er\_iris_{Méthode}$  et  $Er\_pond_{Méthode}$  l'erreur relative au niveau de l'iris et l'erreur pondérée au niveau de l'aire d'étude respectivement.

$$ER\_IRIS_{Méthode} = \frac{(pop_{estim\_iris} - pop_{10\_iris})}{pop_{10\_iris}}$$

$$Er\_pond_{Méthode} = \sum_{iris} \frac{|Er\_iris_{Méthode}| * pop_{10\_iris}}{\sum_{IRIS} pop_{10\_iris}} = \sum_{iris} \frac{|pop_{estim\_iris} - pop_{10\_iris}|}{\sum_{IRIS} pop_{10\_iris}}$$

#### 4.3.1.1 RÉPARTITION PONCTUELLE AUX CENTROÏDES DES CARREAUX

La plus simple des méthodes consiste en l'agrégation des données des carrés en leur centroïde. Ainsi un carré n'est plus représenté par son étendue mais par la position de son centre. Dans ce cas il n'y a plus de carreaux partagés et un carré appartient désormais à un seul IRIS.

Cette méthode suppose que toute la population d'un carré est localisée en son centre, ce qui constitue une hypothèse simpliste.

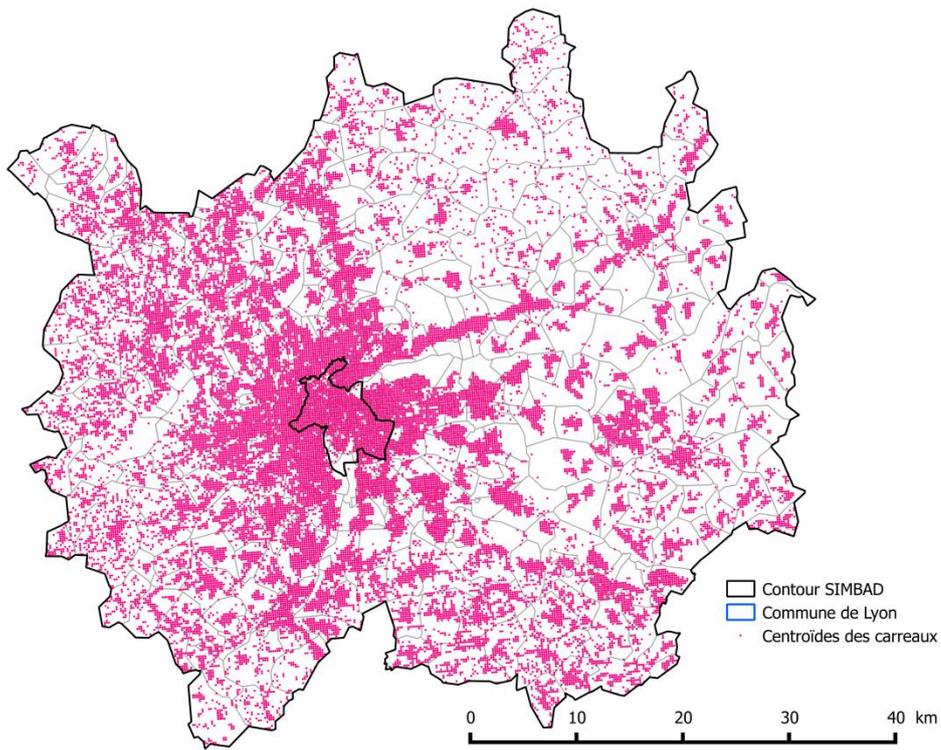


FIGURE 29 : CENTROÏDES DES CARREAUX SUR L'ARE SIMBAD

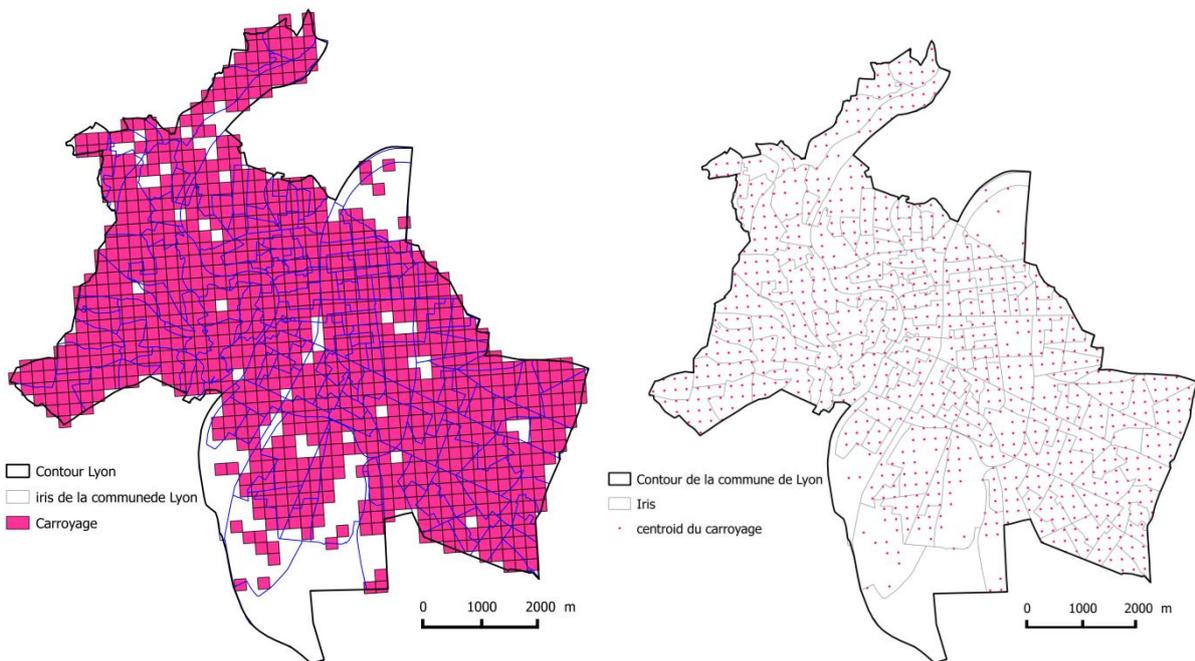


FIGURE 30 : CARROYAGE ET CENTROÏDES AU NIVEAU DE LA COMMUNE DE LYON

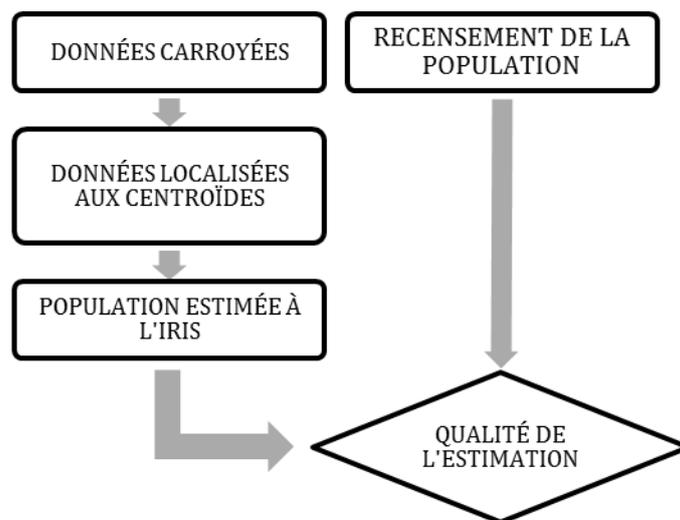


FIGURE 31 : PROCESSUS D'ÉLABORATION DE LA MÉTHODE 1

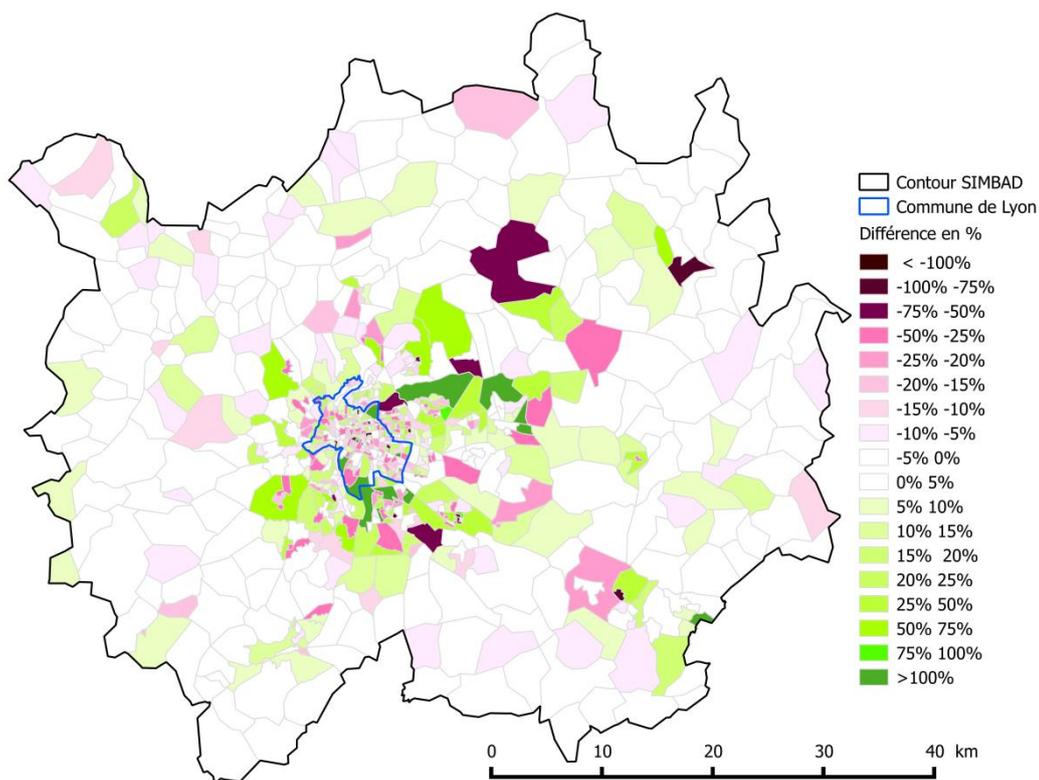
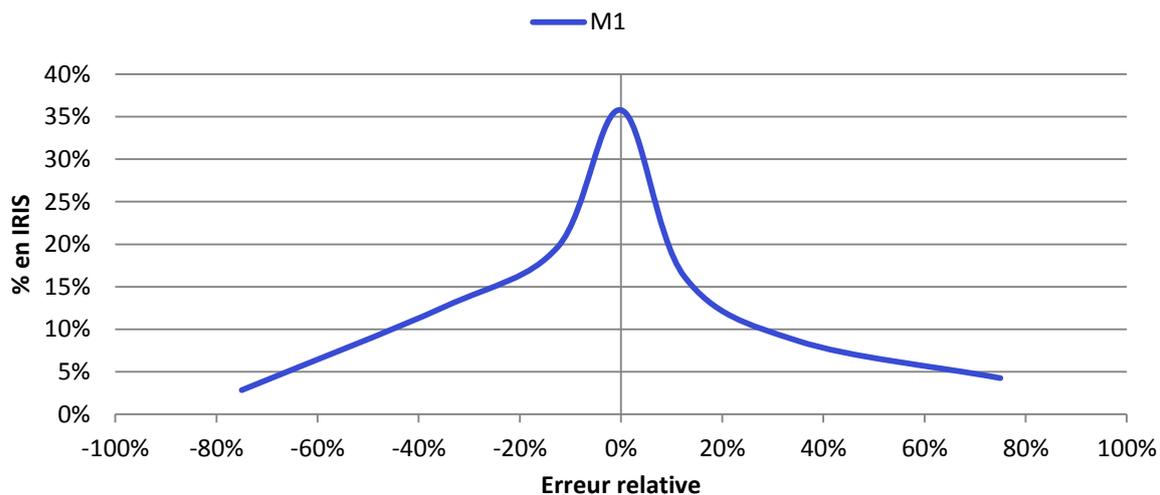
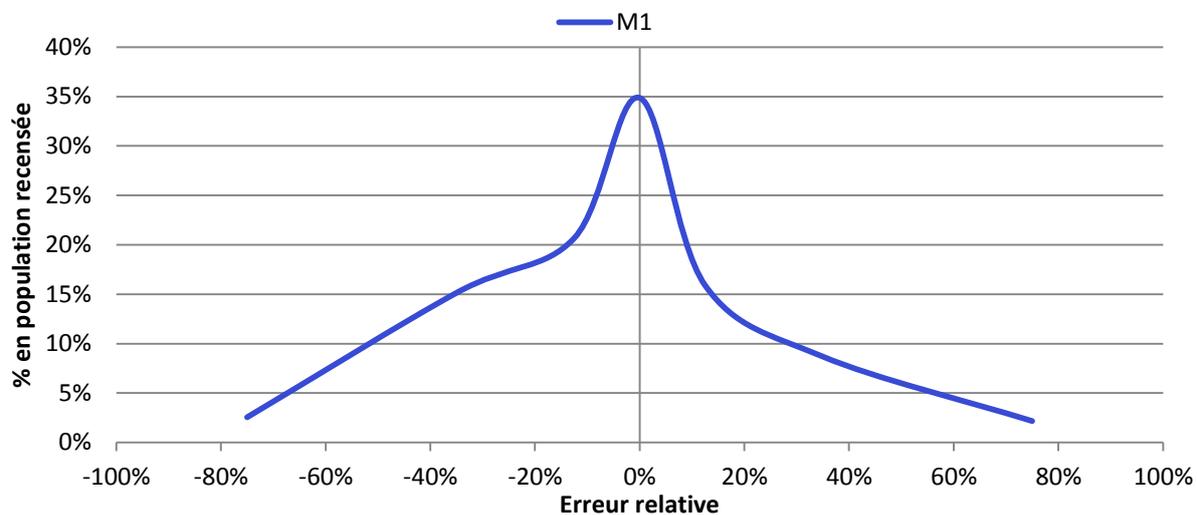


FIGURE 32 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA MÉTHODE 1



GRAPHIQUE 12 : COURBE DE FRÉQUENCE DE L'ERREUR RELATIVE SUR L'AIRE SIMBAD



GRAPHIQUE 13 : POURCENTAGE DE LA POPULATION RECENSEE CONCERNEE PAR L'ERREUR RELATIVE SUR L'AIRE SIMBAD

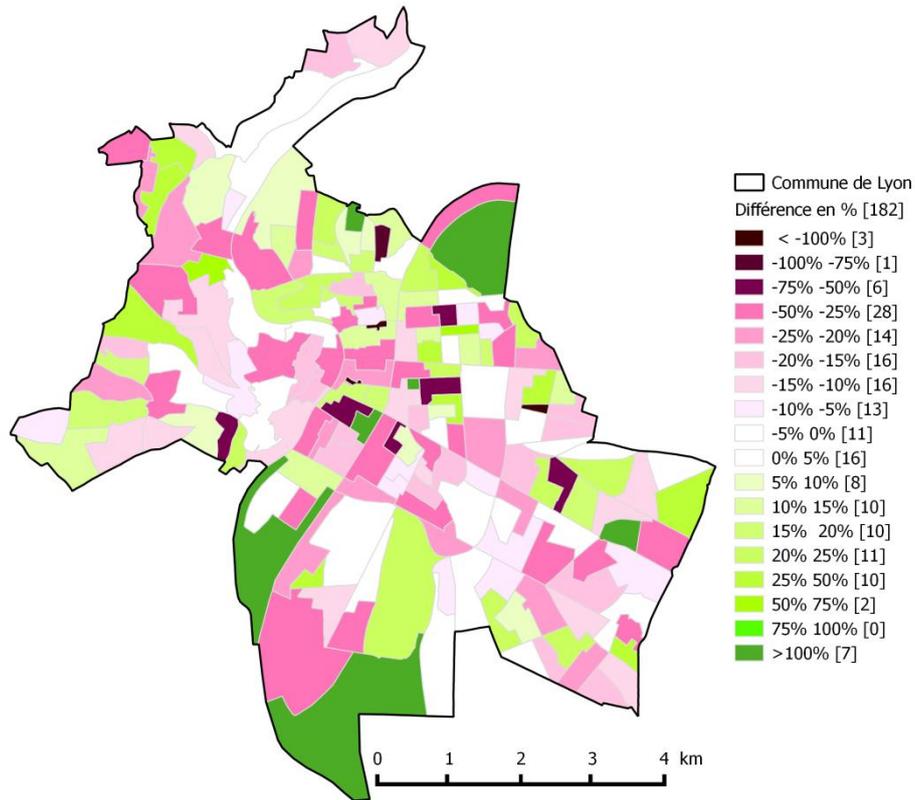
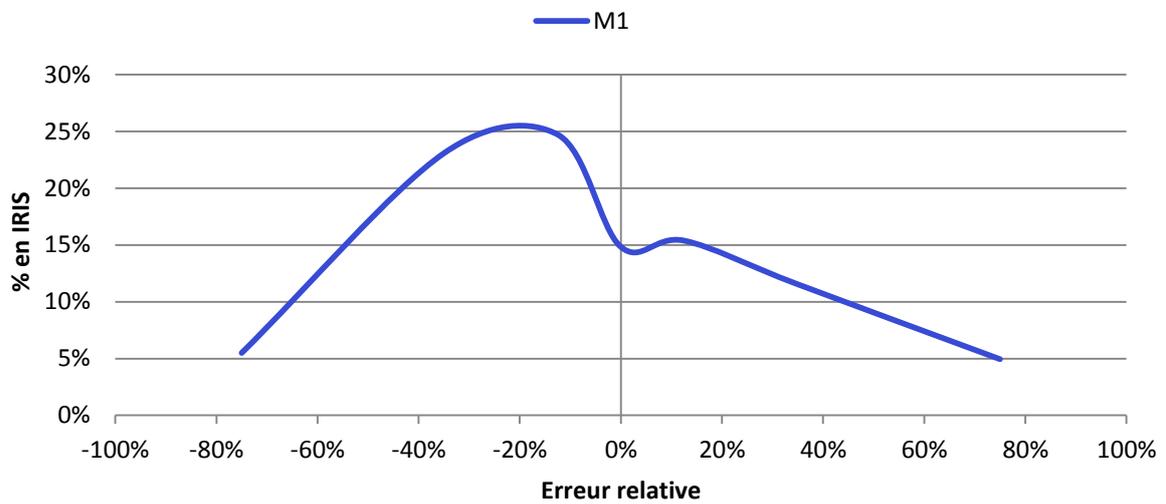
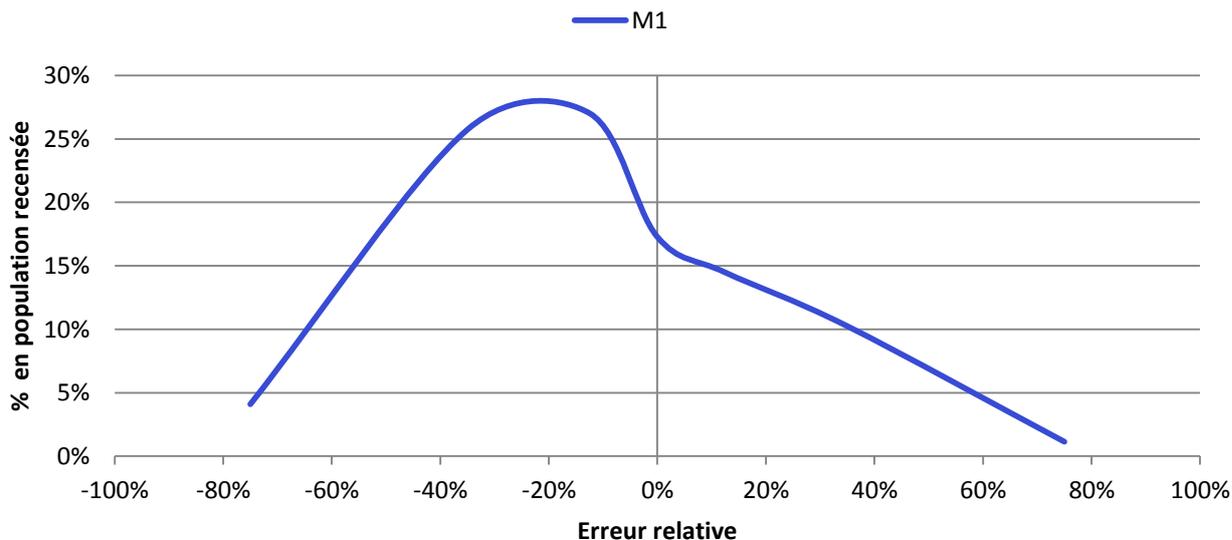


FIGURE 33 : ERREUR RELATIVE AU NIVEAU DE LA COMMUNE DE LYON PAR LA MÉTHODE 1



GRAPHIQUE 14 : COURBE DE FRÉQUENCE DE L'ERREUR RELATIVE SUR LYON



GRAPHIQUE 15 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON

Par rapport à l'ensemble de l'aire d'étude, la méthode 1 semble être moins performante au niveau de Lyon. En effet, seulement 55% des IRIS de Lyon ont une erreur relative entre -20% et 20% contre 72% pour l'aire totale. De même, seulement 59% de la population de Lyon a une erreur relative entre -20% et 20% contre 71% pour l'aire SIMBAD.

Ce constat est, à la fois, dû à la finesse du découpage IRIS au niveau de Lyon et au poids de la population y résidant. En effet, au niveau de Lyon le découpage IRIS s'affine et sa superficie diminue pour n'inclure que des portions de carreaux. Ce résultat peut être amélioré en introduisant une répartition proportionnelle à la surface.

#### 4.3.1.2 RÉPARTITION PROPORTIONNELLE À LA SURFACE DES CARREAUX

La dernière méthode recourt à l'hypothèse de concentration de la population aux centroïdes et aboutit à des résultats moins pertinents au niveau de Lyon.

Cette seconde méthode lève l'hypothèse de concentration de la population et introduit l'hypothèse d'uniformité de la densité au niveau du carreau. Ainsi chaque carreau de 200m par 200m est supposé avoir une densité surfacique uniforme de la population. Par conséquent, à toute entité appartenant au carreau, on peut affecter une population en fonction de sa surface. En particulier, les entités partagées entre plusieurs IRIS.

La mise en place de cette méthode passe par la construction d'une nouvelle base dite par la suite 'Base Carreaux répartis'. Cette dernière est la répartition du carroyage sur les IRIS. Cette base passe donc de 27 101 carreaux à 35 127 entités.

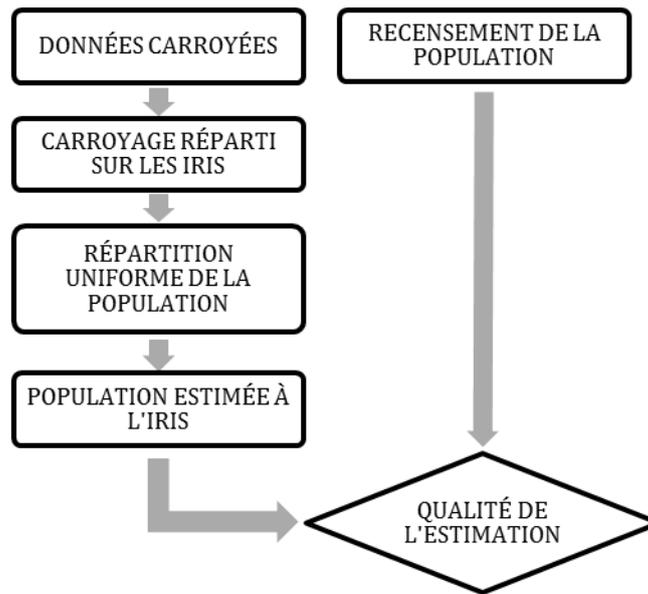


FIGURE 34 : PROCESSUS D'ÉLABORATION DE LA MÉTHODE 2

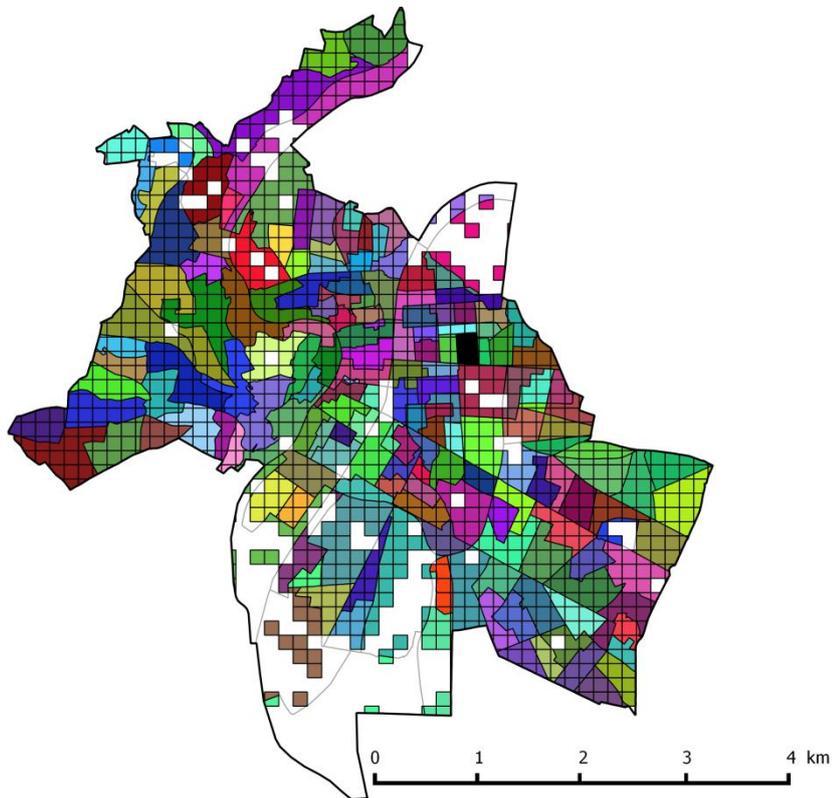


FIGURE 35 : EXEMPLE D'ENTITÉS COLORIÉES EN FONCTION DE L'IRIS D'APPARTENANCE

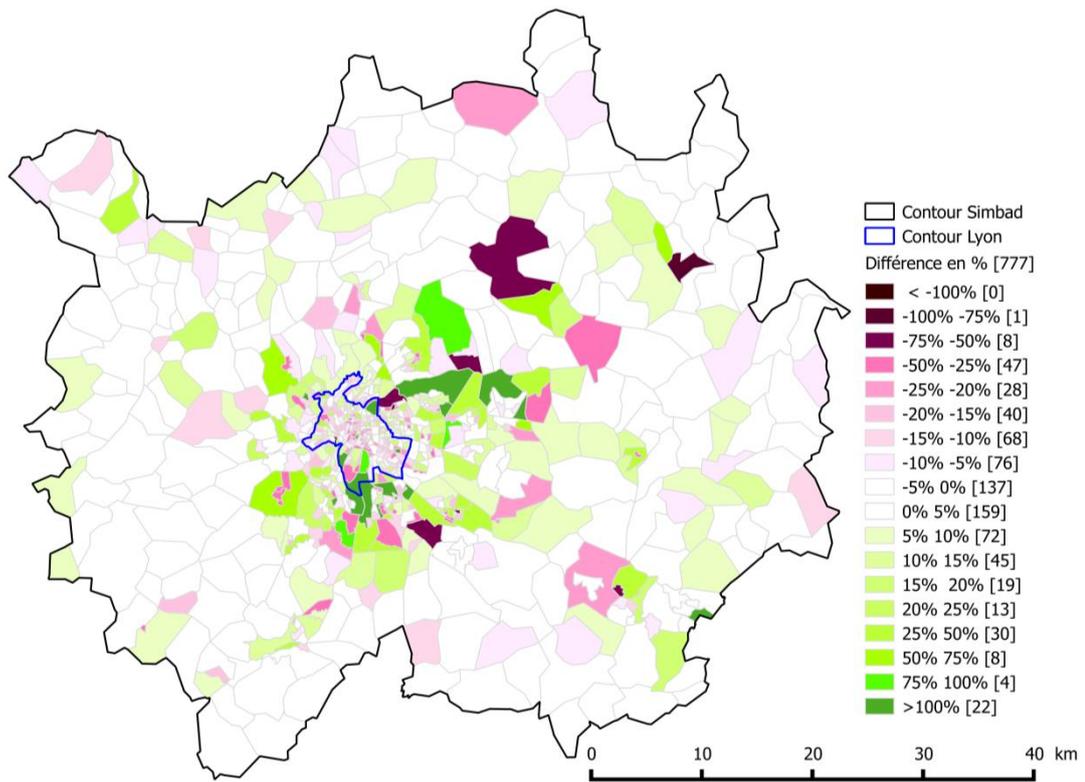
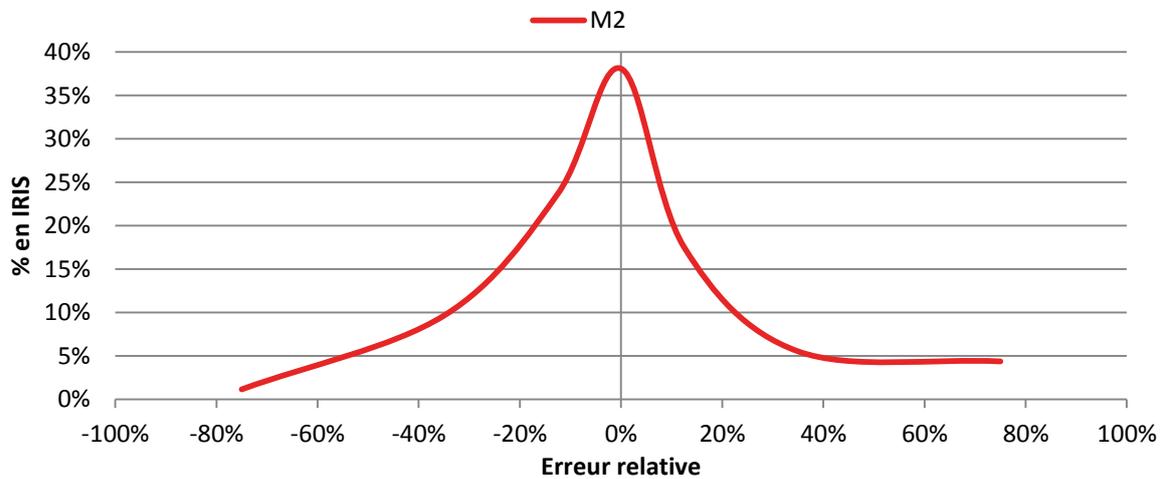
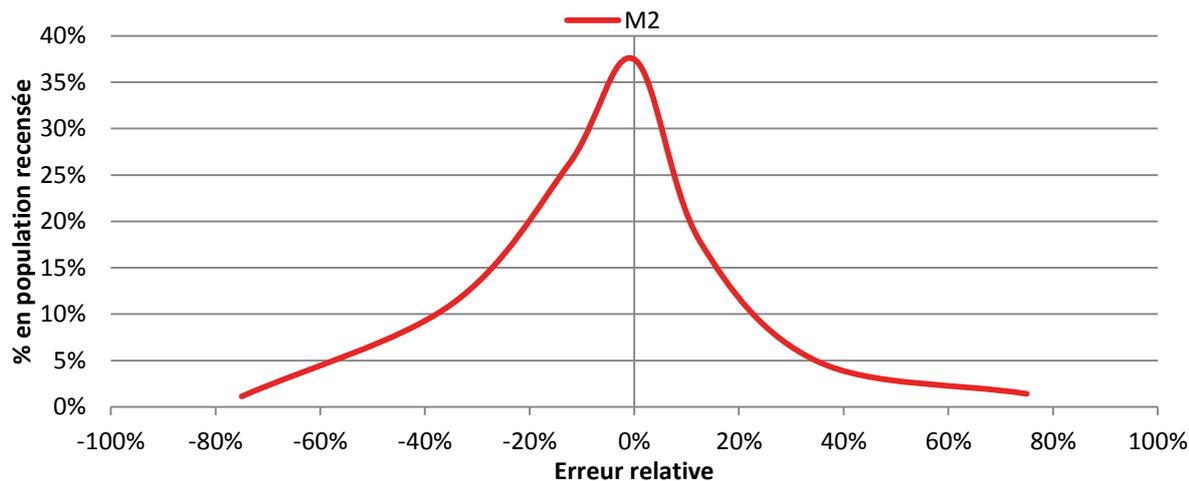


FIGURE 36 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA MÉTHODE 2



GRAPHIQUE 16 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR L'AIRE SIMBAD



GRAPHIQUE 17 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD

Cette méthode apporte une légère amélioration au niveau de l'aire SIMBAD. La population concernée par la fourchette d'erreur relative de +/-20% passe ainsi de 71% à 82%.

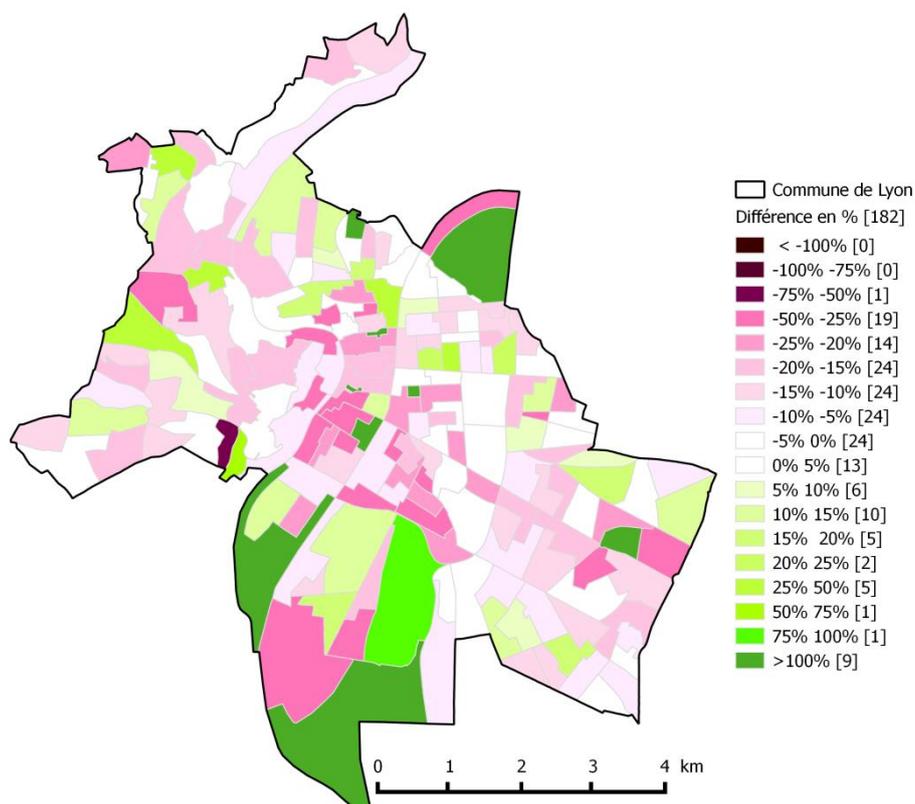
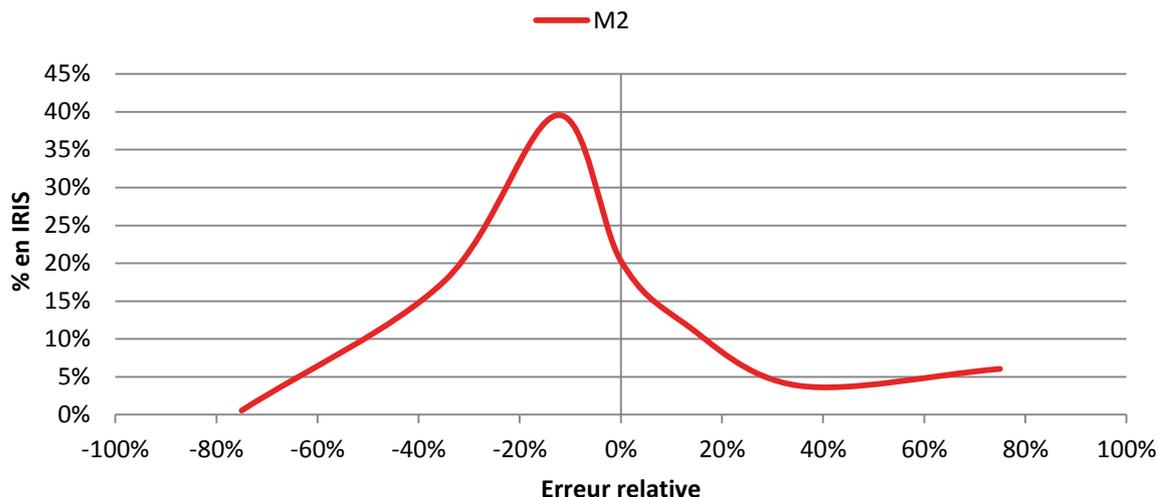
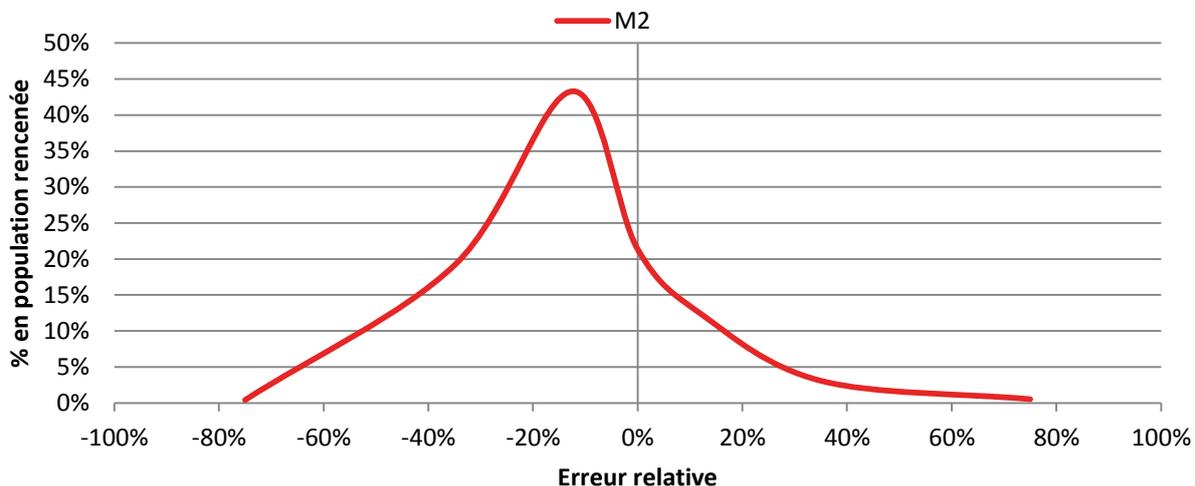


FIGURE 37 : ERREUR RELATIVE AU NIVEAU DE LA COMMUNE DE LYON PAR LA MÉTHODE 2



GRAPHIQUE 18 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR LYON



GRAPHIQUE 19 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON

Comme prévu cette méthode améliore la qualité de l'estimation de la population au niveau de Lyon. Cette amélioration est plus notable sur Lyon dont la population concernée par une erreur relative de -20% à 20% passe de 59% avec la méthode précédente à 77%. En parallèle 72% des IRIS de Lyon affichent cette erreur relative contre 55% avec la méthode 1.

L'hypothèse d'uniformité de la densité au sein du carreau semble être plus réaliste que celle de la concentration de la population, toutefois elle est inadaptée aux zones denses à usage mixte du sol.

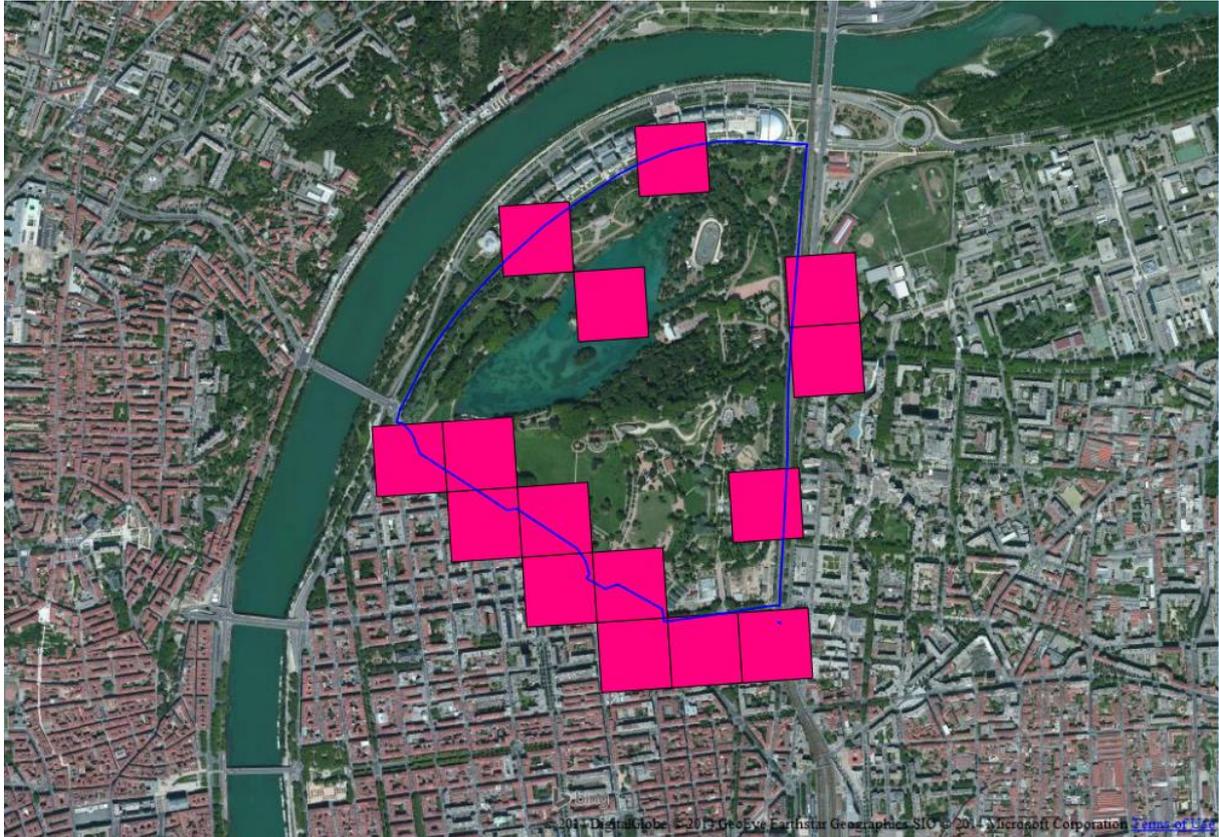


FIGURE 38 : EXEMPLE DE RÉPARTITION DE LA POPULATION AU NIVEAU DE L'IRIS DU PARC (FOND BINGMAP)

Dans cet exemple le recensement de 2010 n'identifie aucune population à l'intérieur de l'IRIS du Parc, toutefois l'utilisation de la méthode 2, ainsi que de la méthode 1, génère une population de presque 600 individus. Cette inconsistance des méthodes est principalement due au traitement indifférencié des données au sein des carreaux. Pour contourner cette limite préjudiciable à la qualité de la répartition de la population, nous avons choisi de faire introduire la base IGN BD TOPO®.

Grâce à cette base on peut lever l'hypothèse d'uniformité de la densité de population à l'intérieur d'un carreau et travailler à l'échelle plus fine des bâtiments.

#### 4.3.2 EXPLOITATION DES DONNÉES CARROYÉES ET DE LA BD TOPO®

##### 4.3.2.1 RÉPARTITION PROPORTIONNELLE AU VOLUME BÂTI

Cette méthode consiste en la répartition de la population des carreaux sur les bâtiments résidentiels à l'intérieur du même carreau. Cette répartition se fait au prorata de la 'Surface Développée (SD)'.

$$SD_{\text{bâtiment}} = (\text{Superficie}_{\text{bâtiment}}) * (\text{Nombre}_{\text{étages}}) = S_{\text{bâtiment}} * Nb_{\text{étages}}$$

Ainsi la population estimée affectée à un bâtiment est :

$$pop_{\text{estim}_{\text{bâtiment}}} = pop_{\text{carreau}} * \frac{S_{\text{bâtiment}}}{\sum_{\text{carreau}} S_{\text{bâtiment}}} \quad (iv)$$

#### 4.3.2.1.1 CALCUL DU VOLUME BÂTI

Pour le calcul de  $pop_{estim-bâtiment}$  il n'est pas nécessaire de définir sur toute l'aire d'étude une valeur forfaitaire de la hauteur par niveau. En effet, l'hypothèse suivante : 'A l'intérieur d'un carreau, la hauteur d'un étage est constante pour tous les bâtiments' couplée à l'équation (iv) nous évite de définir une valeur forfaitaire. Soit  $h_{carré}$  cette hauteur, dans ce cas :

$$SD_{bâtiment} = S_{bâtiment} * \frac{Hauteur_{bâtiment}}{h_{carré}}$$

$$pop_{estim\_bâtiment} = pop_{carreau} * \frac{S_{bâtiment} * Hauteur_{bâtiment} / h_{carré}}{\sum_{carreau} S_{bâtiment} * Hauteur_{bâtiment} / h_{carré}}$$

$$pop_{estim\_bâtiment} = pop_{carreau} * \frac{S_{bâtiment} * Hauteur_{bâtiment}}{\sum_{carreau} S_{bâtiment} * Hauteur_{bâtiment}} \quad (v)$$

Soit  $VB_{bâtiment} = S_{bâtiment} * Hauteur_{bâtiment}$  le volume bâti.

Dans un premier temps, la hauteur totale, y compris celle des rez-de-chaussée, est prise en compte dans le calcul des volumes bâtis, Autrement dit, les bâtiments à usage mixte - résidence avec commerces en rez-de-chaussée par exemple - sont considérés comme étant à usage résidentiel uniquement. Cette hypothèse peut être levée par la suite.

On affecte ensuite à chaque bâtiment de la base BD TOPO nettoyée l'attribut  $VB_{bâtiment}$ . Suite à une manipulation sous SIG on affecte à chaque élément de la 'Base Carreaux répartis' le volume bâti total à l'intérieur de cette entité ainsi que le volume bâti total du carreau auquel l'entité appartient. Ainsi la répartition de la population du carreau sur les entités est réalisée au prorata de la surface développée de cette dernière.

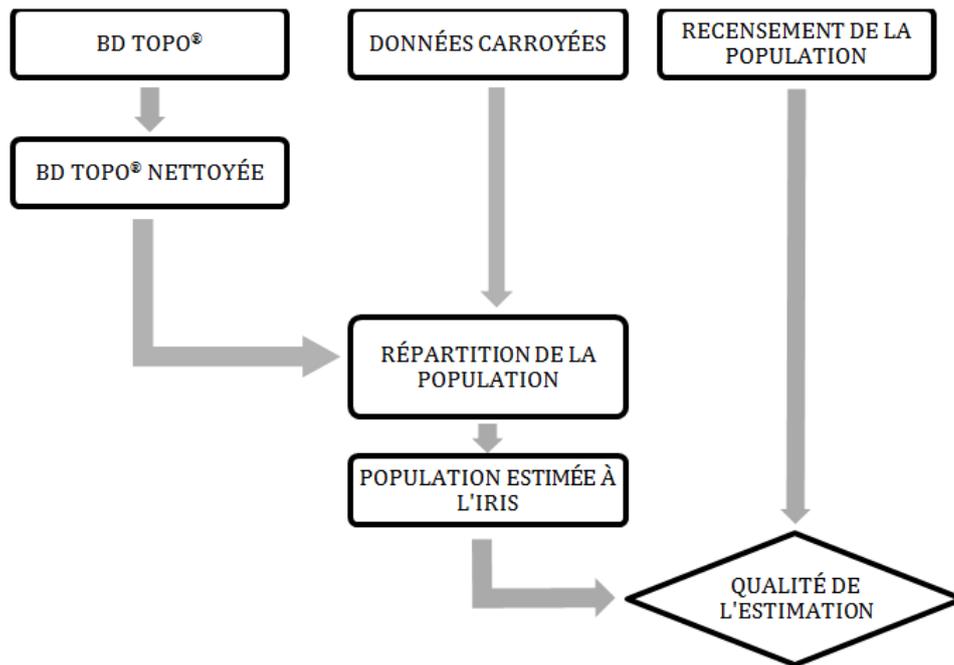


FIGURE 39: PROCESSUS D'ÉLABORATION DE LA MÉTHODE 3

Remarque :

- De nouveau se pose le problème des bâtiments chevauchant deux ou plusieurs entités de la 'Base Carreaux répartis'. L'approche de résolution de ce problème aboutit à deux variantes.
- Une inconsistance entre la base carroyée et la BD TOPO® apparaît avec quelques carreaux qui ne contiennent pas de bâtiments. On a choisi de garder ces carreaux et de répartir leur population fiscale par la méthode 2. De même quelques bâtiments ne sont pas couverts par des carreaux. Ces derniers sont supprimés de la base.

#### 4.3.2.2 VARIANTE 1 : DÉCOUPAGE DES BÂTIMENTS EN PORTIONS

Cette variante découpe un bâtiment chevauchant plusieurs entités en plusieurs portions et affecte à chaque entité les portions de bâtiments qui sont incluses dans ses limites. Cette variante aboutit à 387 322 bâtiments ou portions de bâtiments.

Ensuite, comme expliqué plus haut, une population est affectée à chaque entité de la 'Base Carreaux répartis' en fonction de la surface totale bâtie à l'intérieur.

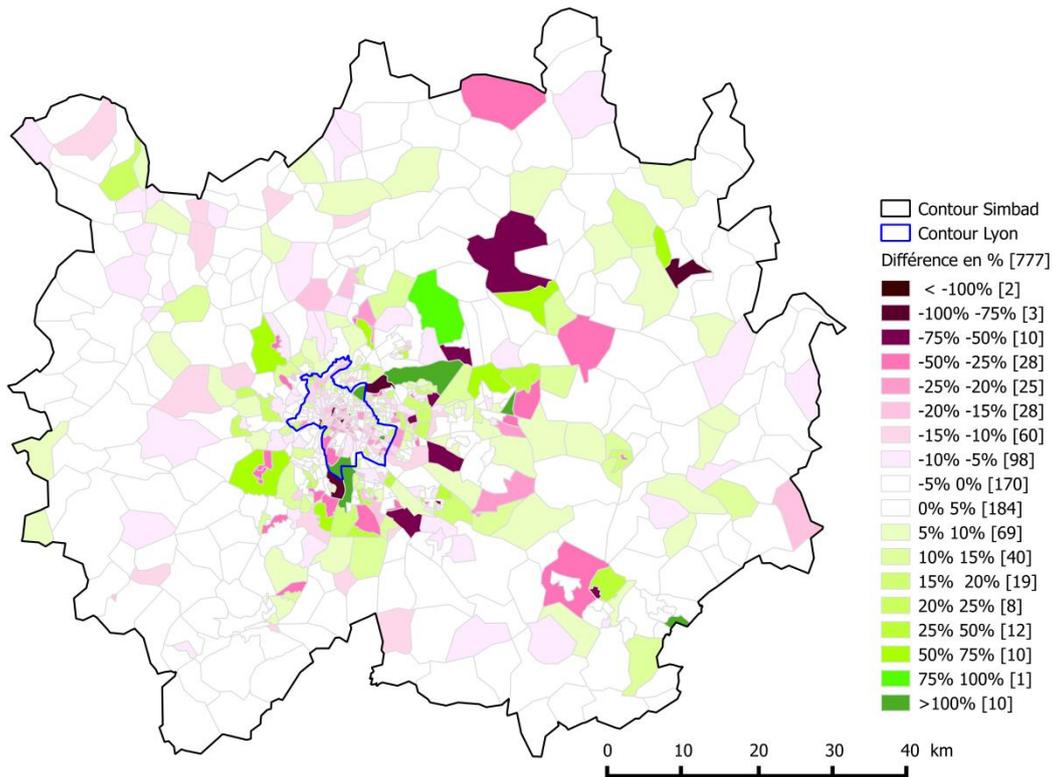
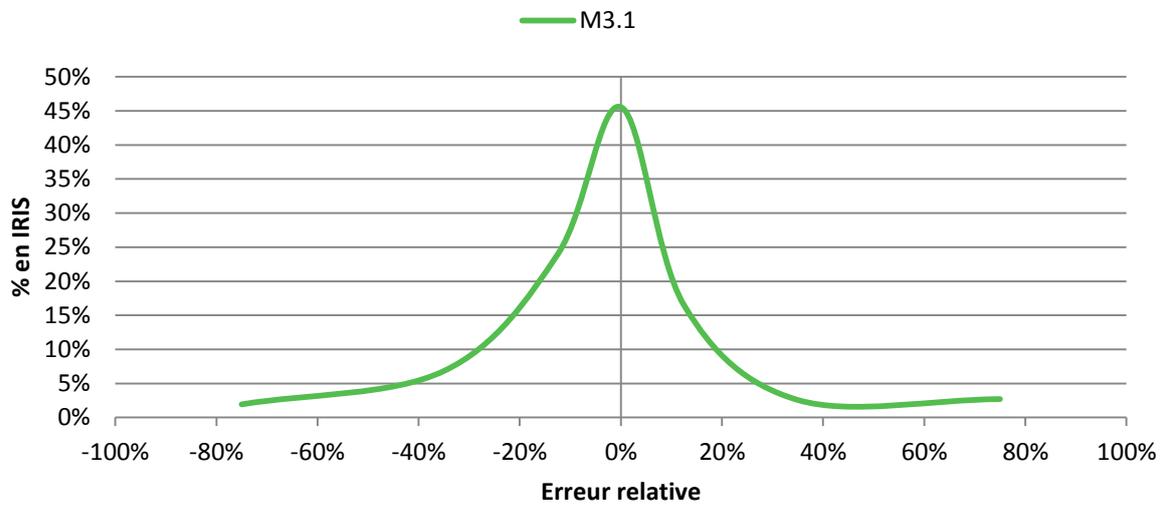
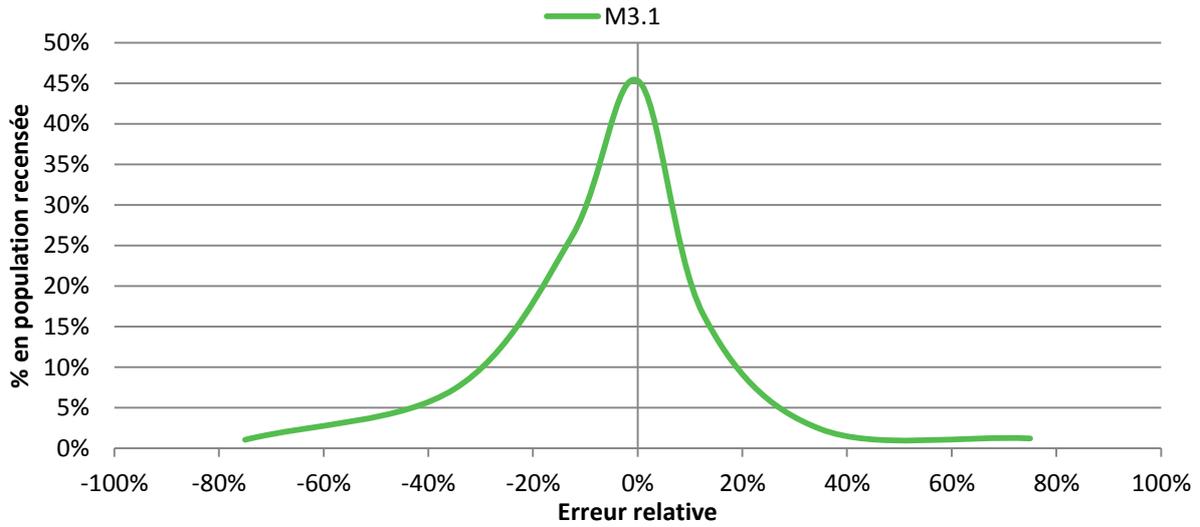


FIGURE 40 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA METHODE 3.1



GRAPHIQUE 20 : COURBE DE FRÉQUENCE DE L'ERREUR RELATIVE SUR L'AIRE SIMBAD



GRAPHIQUE 21 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD

L'amélioration apportée par cette méthode est notable par rapport aux dernières méthodes. Ainsi 88% de la population de totale a une erreur relative entre -20% et 20%.

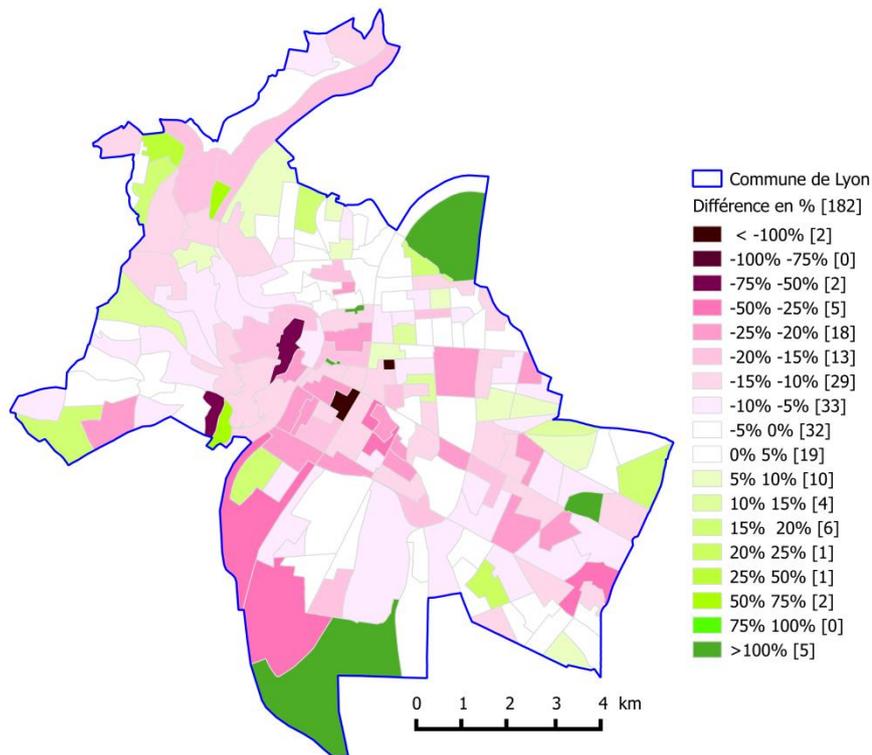
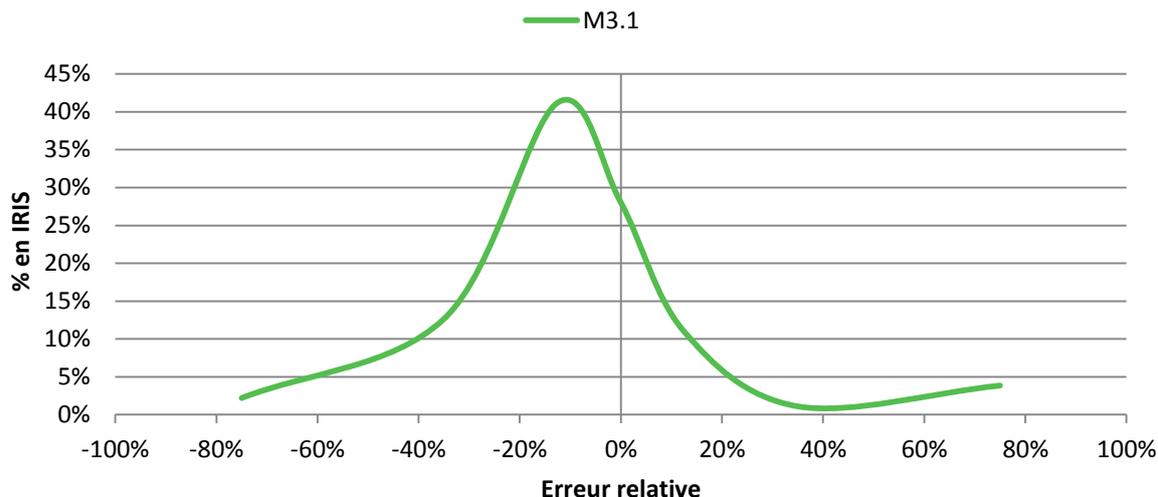
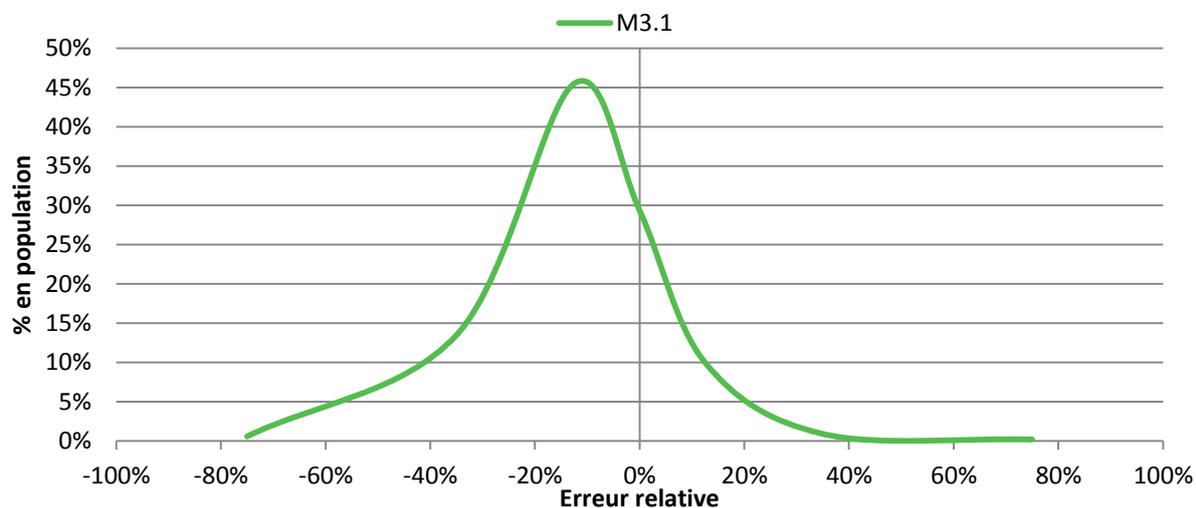


FIGURE 41 : ERREUR RELATIVE AU NIVEAU DE LYON PAR LA METHODE 3.1



GRAPHIQUE 22 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR LYON



GRAPHIQUE 23 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON

Cette méthode est plus performante que les précédentes. Elle améliore l'estimation de la population pour Lyon ainsi que sur toute l'aire d'étude. La population lyonnaise concernée par une erreur entre -20% et 20% représente 85% de la population communale.

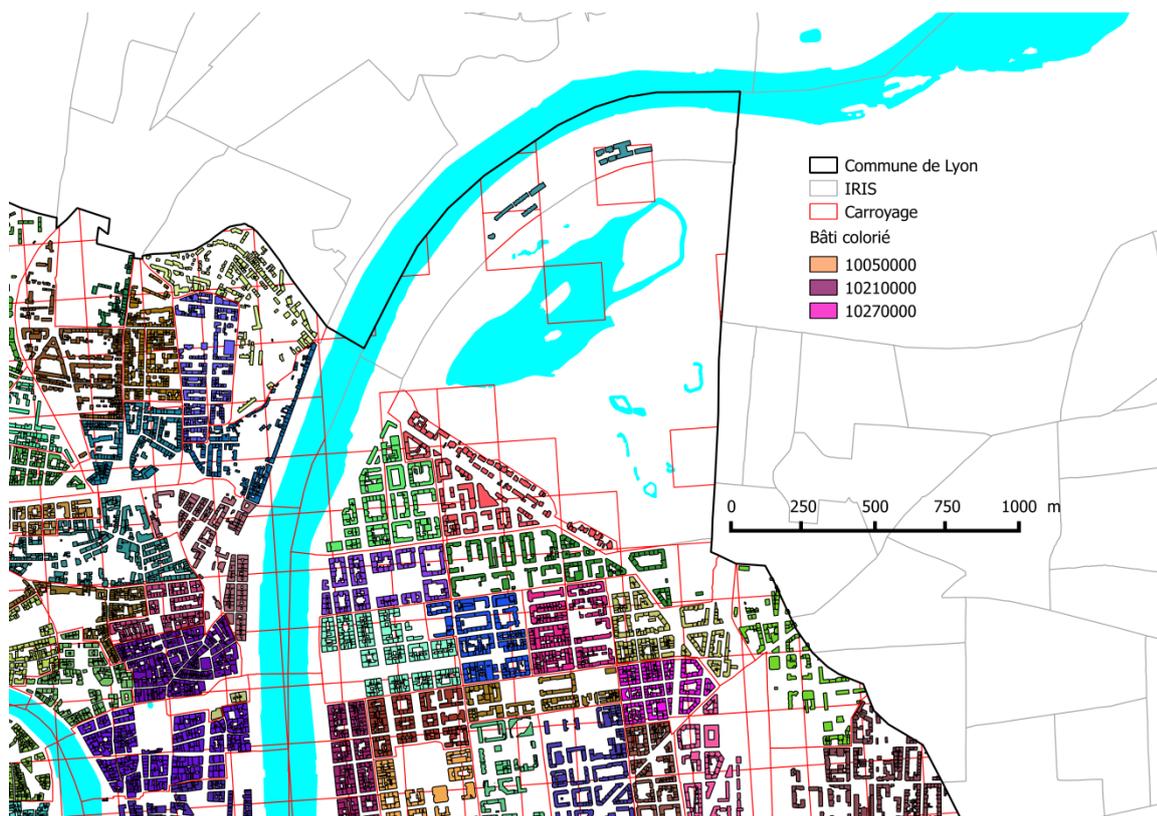


FIGURE 42 : EXEMPLE DE RÉPARTITION DU BÂTI EN COULEUR SUR LES DIFFÉRENTES ENTITÉS

Remarquons que les entités appartenant à l'IRIS du Parc ne contiennent quasiment pas de bâtiments résidentiels (voir Figure 42). Ainsi la méthode 3.1 estime à 12 le nombre de résidents à cet IRIS.

#### 4.3.2.3 VARIANTE 2 : RÉPARTITION DES BÂTIMENTS PAR LA MÉTHODE DES CENTROÏDES

Cette variante est l'adaptation de la méthode 1 appliquée aux bâtiments. Au lieu de découper les éléments de la BD TOPO nettoyée, on représente chaque bâtiment par son centroïde qui n'appartient qu'à une seule entité. Cette méthode peut, a priori, sembler simpliste, toutefois, elle correspond à la logique de construction des données Insee carroyées qui utilise le ménage comme unité et non la superficie bâtie. En outre, des portions de bâtiments générées par la méthode 3.1 peuvent ne pas avoir de correspondance réelle sur le terrain. Enfin pour le travail de construction des connecteurs, la localisation exacte des entrées d'immeubles est nécessaire. Cette dernière a plus de chance d'appartenir au carreau qui contient la majorité de l'immeuble.

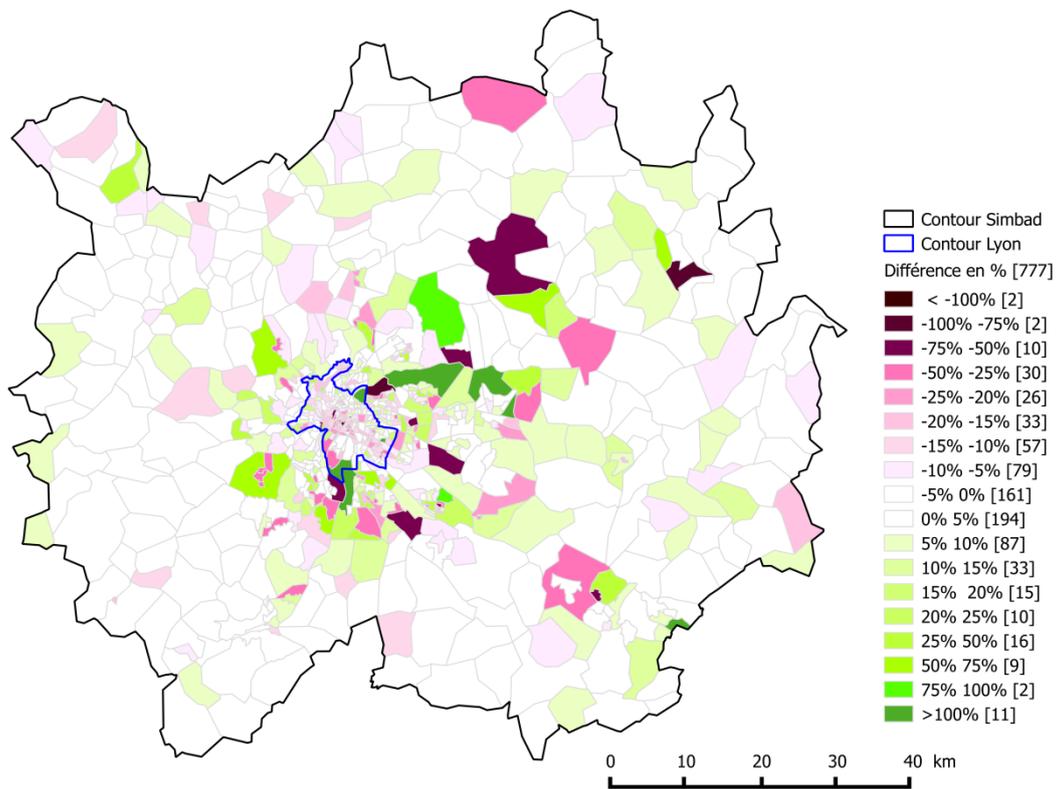
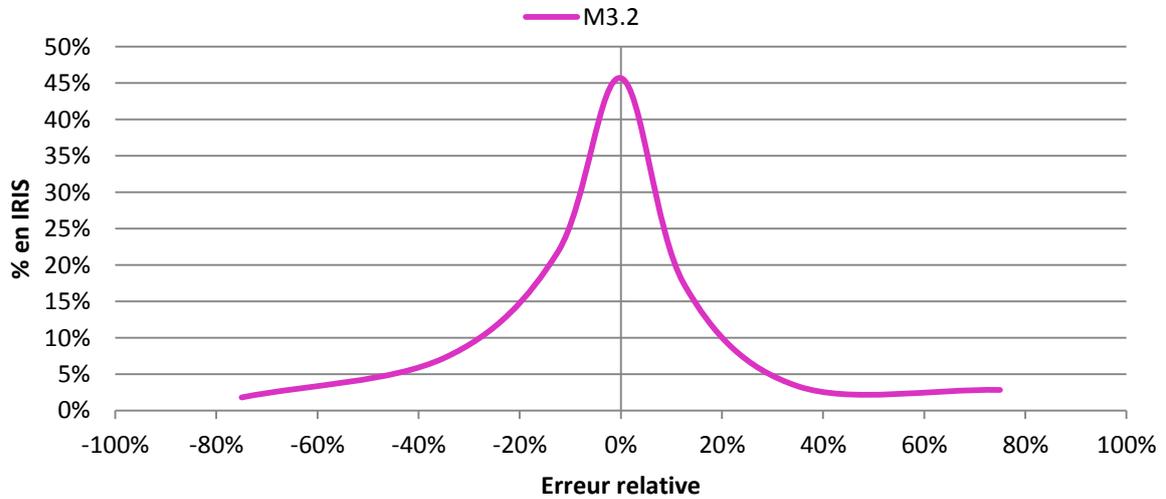
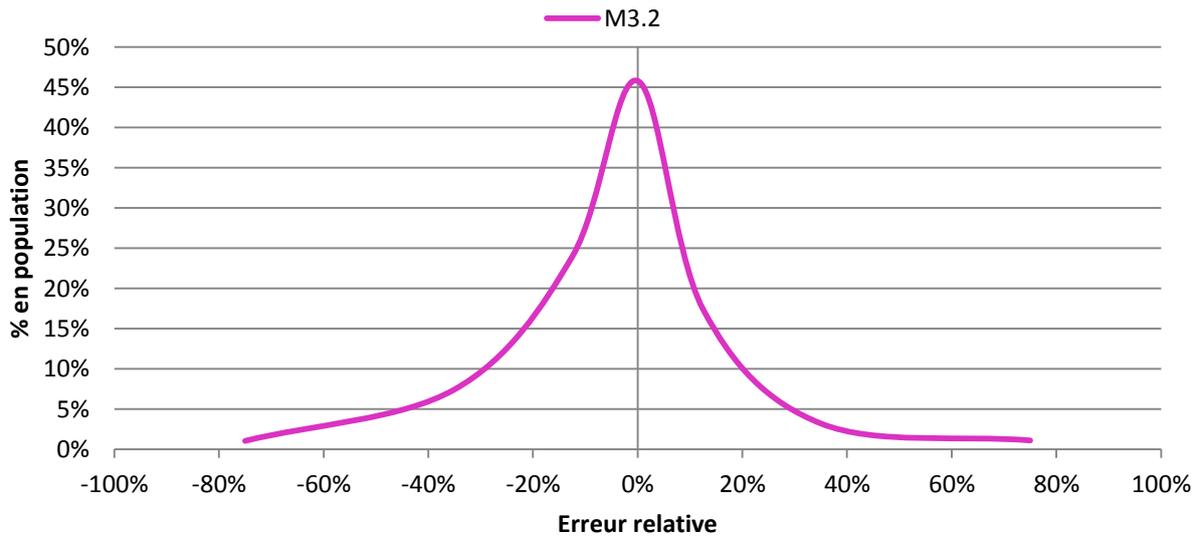


FIGURE 43 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA METHODE 3.2



GRAPHIQUE 24 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR L'AIRE SIMBAD



GRAPHIQUE 25 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD

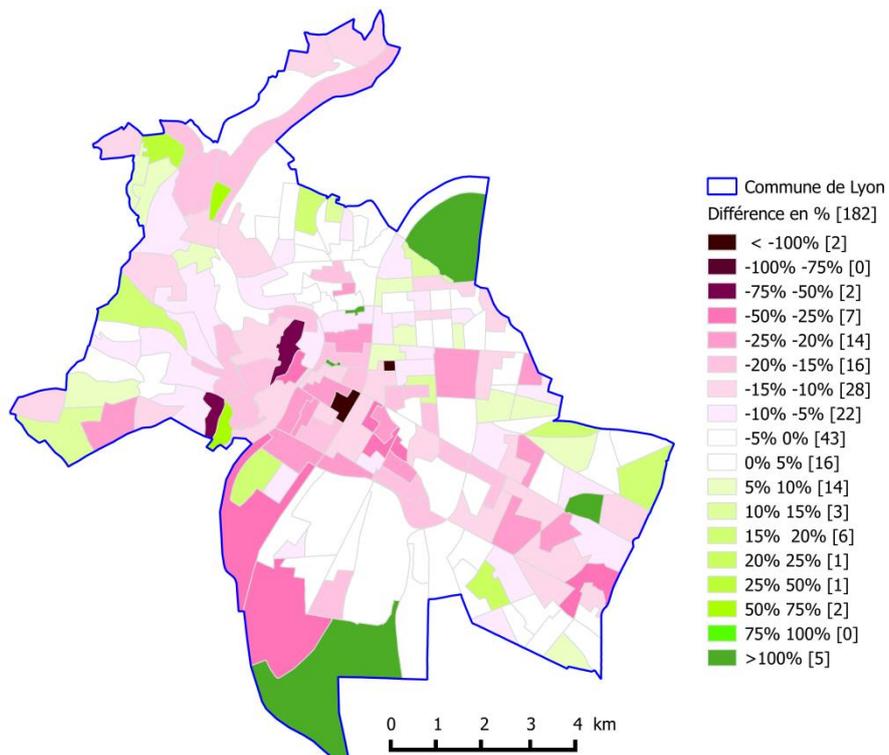
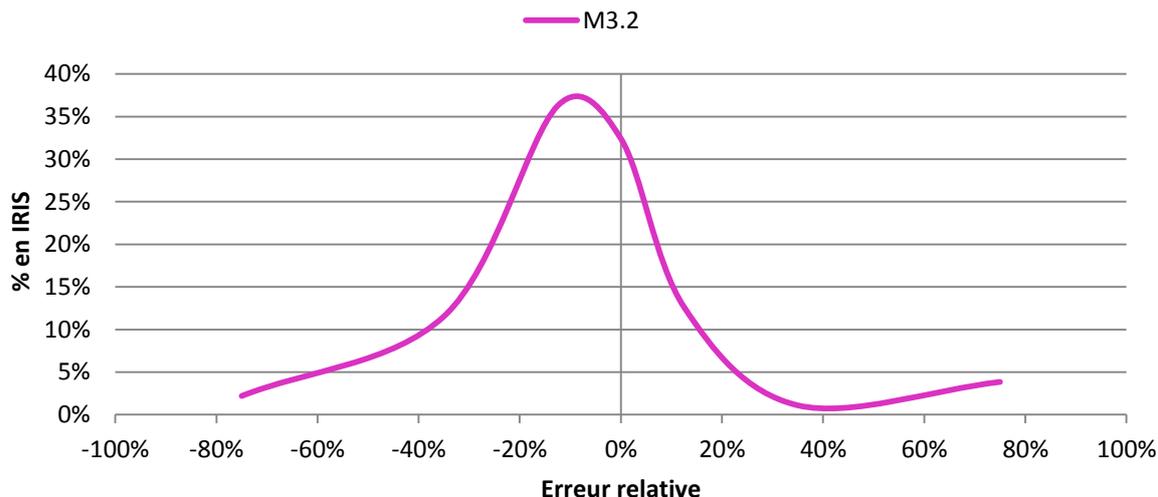
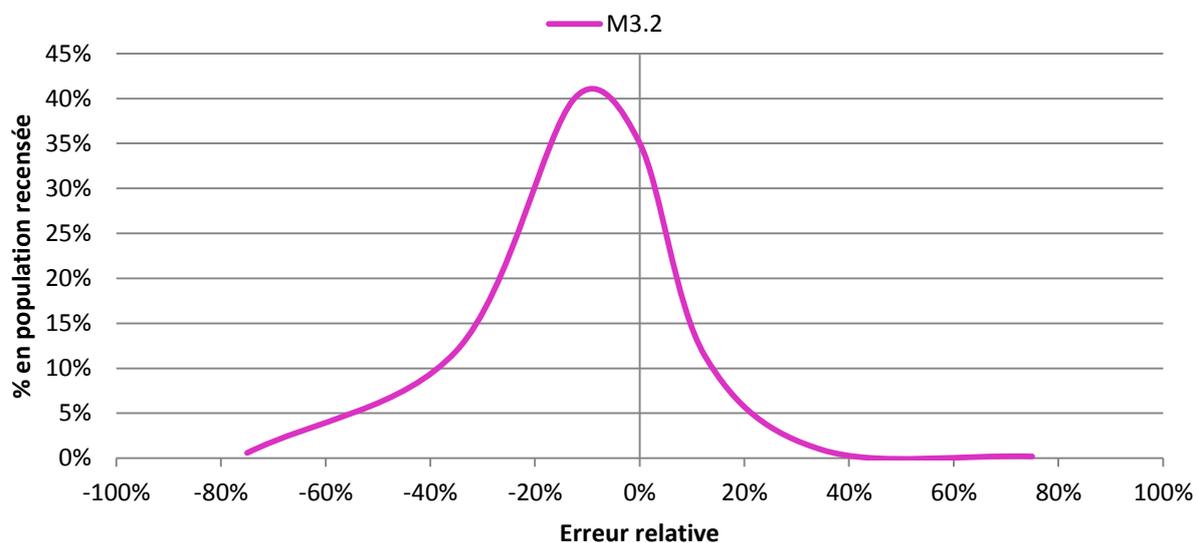


FIGURE 44 : ERREUR RELATIVE AU NIVEAU DE LYON PAR LA METHODE 3.2



GRAPHIQUE 26 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR LYON

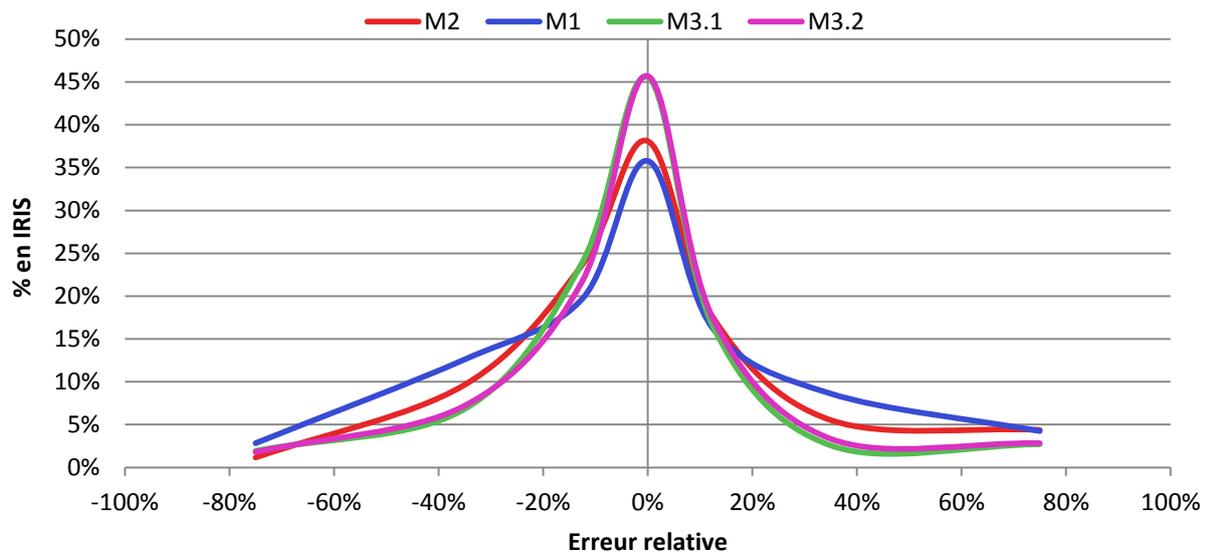


GRAPHIQUE 27 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON

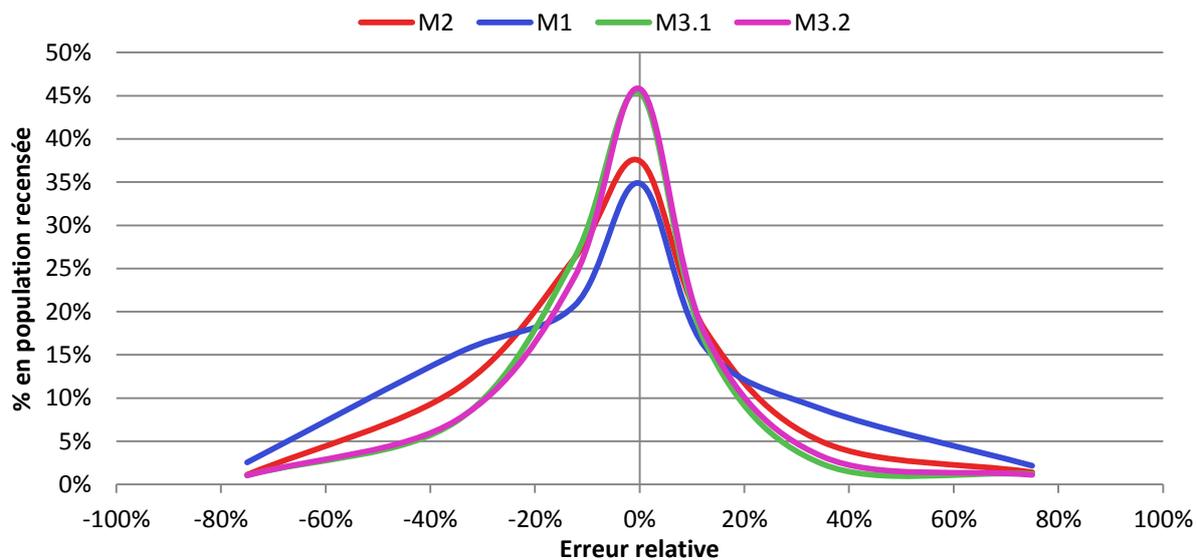
Cette dernière méthode aboutit à des résultats comparables à ceux de la 3.1 avec une légère différence au niveau de Lyon.

#### 4.4 COMPARAISON DES MÉTHODES D'ESTIMATION DE LA POPULATION

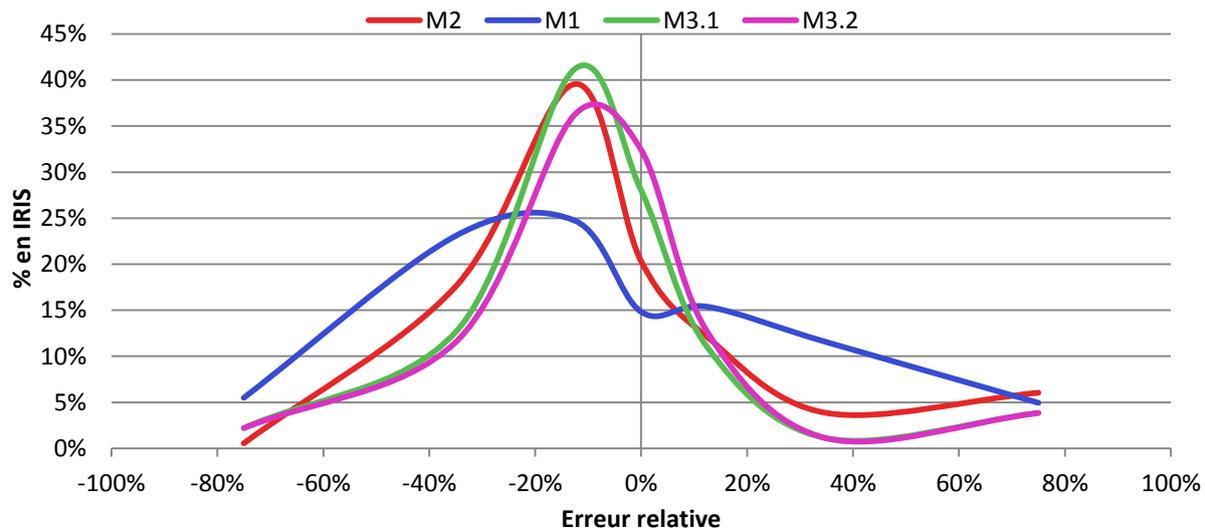
À travers l'analyse précédente, il s'avère que l'utilisation de la BD TOPO<sup>®</sup> améliore la qualité de l'estimation de la population. Sur la base de cette population et de son écart par rapport à la population recensée, une comparaison est dressée ci-après.



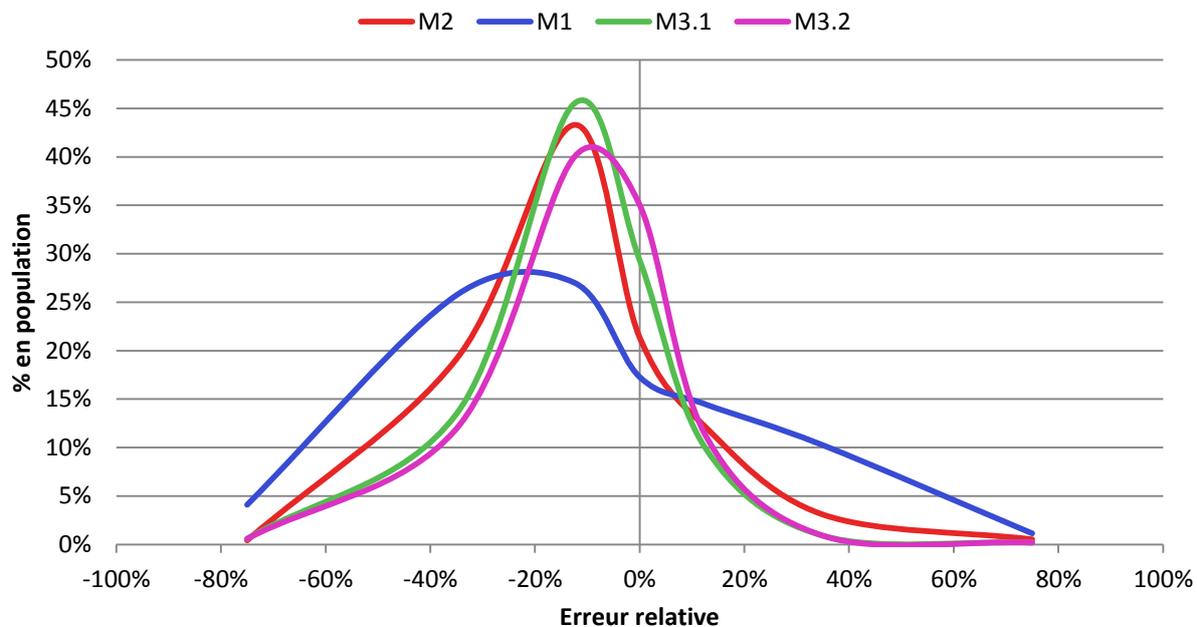
GRAPHIQUE 28 : COMPARAISON DES ERREURS REATIVES GÉNÉRÉES PAR LES 4 MÉTHODES SUR L'AIRE SIMBAD



GRAPHIQUE 29 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRE SIMBAD

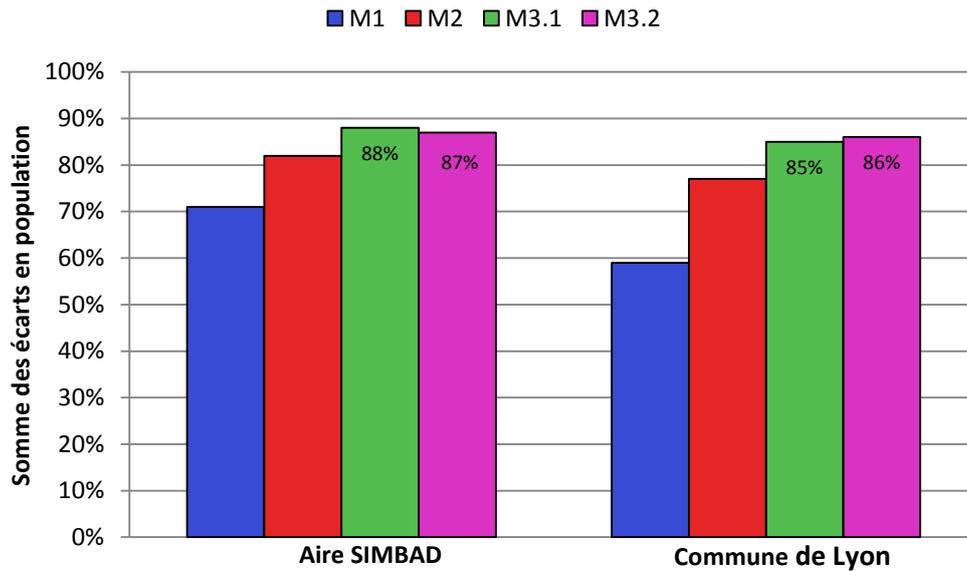


GRAPHIQUE 30 : COMPARAISON DES ERREURS RELATIVES GÉNÉRÉES PAR LES 4 MÉTHODES SUR LYON

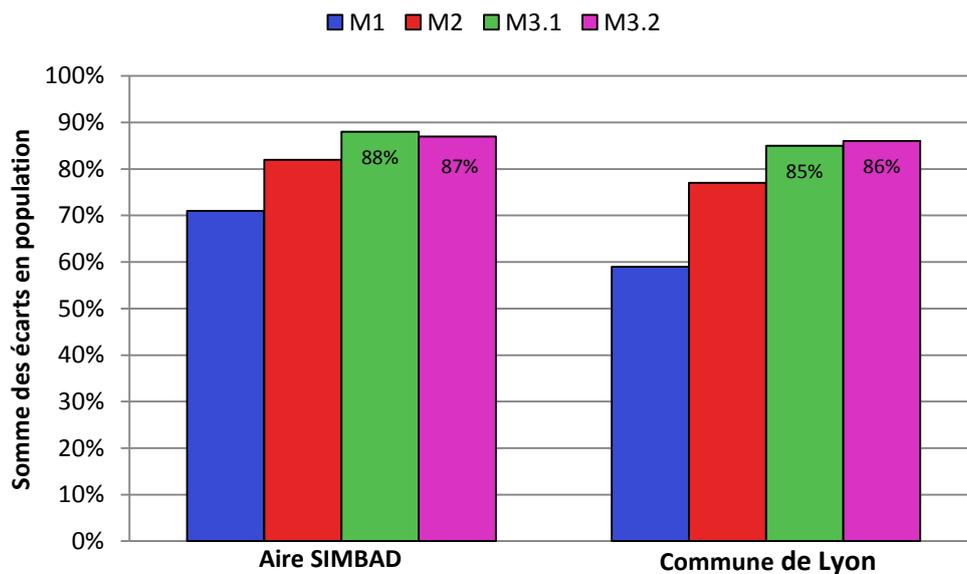


GRAPHIQUE 31 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON

Cette comparaison fait sortir les deux dernières méthodes ([3.1](#) et [3.2](#)) comme étant les plus adaptées à l'estimation de la population. Pour différencier entre ces deux méthodes nous allons recourir à l'erreur globale pondérée par la population ( $Er\_ponder$ ).



GRAPHIQUE 32 : SOMME DES ÉCARTS ENTRE POPULATIONS ESTIMÉE ET RECENSÉE SELON L'AIRE D'ÉTUDE



GRAPHIQUE 33 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR UNE ERREUR RELATIVE DE -20% À 20%

Sur l'aire SIMBAD, ce dernier critère montre que la méthode 3.1 est la plus performante. Par contre sur la commune de Lyon, la méthode 3.2 génère moins d'écart par rapport au recensement de la population (Graphique 33). Ceci dit, la méthode 3.2 est plus facile d'application et moins onéreuse en capacité de calcul.

Dans ce qui suit nous allons travailler avec cette dernière méthode pour caler la population estimée sur la population recensée.

## 4.5 CALAGE DE LA POPULATION ESTIMÉE SUR LA POPULATION RECENSÉE

À l'issue des étapes précédentes, nous disposons d'une base de bâtis géolocalisés, occupés par une population fiscale.



FIGURE 45 : EXEMPLE D'ESTIMATION DE LA POPULATION DES BÂTIMENTS (FOND GOOGLEMAPS)

Cette base, même si elle utilise la notion de 'ménage fiscal', permet de donner une idée sur la répartition spatiale à l'intérieur des IRIS. Par la suite cette population est 'corrigée' par les données du recensement de 2010 afin de caler la population estimée à l'IRIS à celle recensée à cette échelle. Cette population 'corrigée', notée  $pop^*_{estim\_bâtiment}$  est telle que :

$$pop^*_{estim\_bâtiment} = \frac{pop_{10\_iris}}{pop_{estim\_iris}} * pop_{estim\_bâtiment}$$

À ce stade, chaque bâtiment de la BD TOPO nettoyée abrite une population estimée, calée et arrondie (voir Figure 45).

## 4.6 CONCLUSION

À travers l'élaboration des données de la population et des emplois, il s'avère que la base désagrégée est celle qui donne des résultats satisfaisants. En outre, cette partition permettra, comme on l'a vu, de réduire l'erreur de calcul des connecteurs. Nous allons, donc, utiliser cette base pour la mise en application de la méthodologie présentée dans le chapitre précédent.

# CHAPITRE 5

---

## 5 MISE EN APPLICATION DE LA MÉTHODOLOGIE DE CODIFICATION DES CONNECTEURS

On rappelle ici l'algorithme synthétisé de l'implémentation de la nouvelle méthodologie.

1. Déterminer tous les points de desserte de l'aire d'étude.
2. Affecter à chaque point de desserte la distance d'accessibilité correspondant à la catégorie de la ligne le desservant.
3. Déterminer les points de desserte accessibles depuis chaque IRIS  $I$ .
4. Dédire les lignes accessibles depuis  $I$ .
5. Grouper les points de desserte en fonction des lignes accessibles desservant ces points.
6. Pour chaque ligne accessible et chaque IRIS, partager la population et les emplois entre les différents points de desserte appartenant à cette même ligne.
7. Pour chaque ligne accessible et chaque IRIS, calculer la distance entre les points de desserte de cette ligne et la population et les emplois de l'IRIS. Il s'agit des connecteurs des points de desserte.
8. Pour chaque ligne accessible et chaque IRIS, calculer la longueur moyenne pondérée des différentes longueurs de connecteurs des points de desserte. Il s'agit des connecteurs moyens des points de desserte.
9. Pour un groupe de points de desserte, attribuer la longueur moyenne récemment calculée au connecteur d'un point de desserte de ce groupe. De préférence, ce point de desserte doit appartenir au point d'arrêt qui contient le moins de points de desserte.
10. Pour chaque point d'arrêt, calculer la longueur moyenne pondérée des différentes longueurs des connecteurs moyens des points de desserte y appartenant.

### 5.1 PREMIER CAS ÉLÉMENTAIRE : UN IRIS, UNE LIGNE DE TRANSPORT ET UN SEUL POINT DE DESSERTE

Nous allons commencer par le cas le plus simple qui constitue, dans notre méthodologie, une brique élémentaire de calcul. Ce cas se présente souvent pour les lignes ferroviaires. À l'inverse des bus ou des métros, les gares ferroviaires sont suffisamment espacées pour desservir, sans concurrence ni redondance, un IRIS. À présent, le système ferroviaire n'est pas encore pris en compte dans SIMBAD, toutefois, nous allons présenter l'exemple de l'IRIS 13330000 dit Saint-André-de-Corcy, identifié dans le réseau par le code 46.

Saint-André-de-Corcy a une population estimée à 2991 et accueille presque 774 emplois. Cette population est répartie sur 874 bâtiments résidentiels, tandis que les emplois sont

répartis sur 178 sites, localisés autour de la gare ferroviaire, identifiée dans le réseau par le code 3120 (Figure 46).

L'IRIS 46 accueille une gare ferroviaire située au centre. Cette dernière est la seule à être accessible depuis l'IRIS. Les deux gares environnantes sont situées, respectivement, à 1.5km et 1.8km.

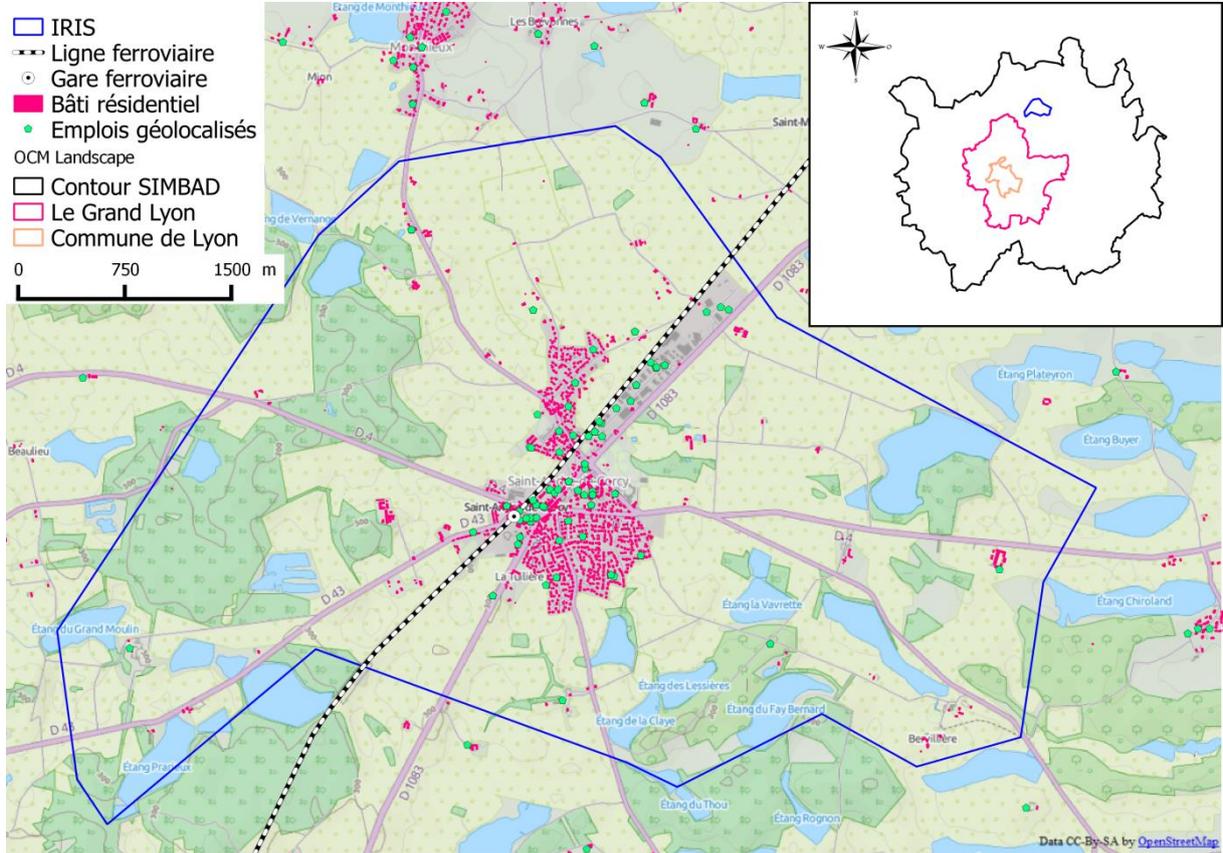


FIGURE 46 : CAS DE L'IRIS NUMÉRO 13330000 DE SAINT-ANDRÉ-DE-CORCY (FOND DE CARTE OPENSTREETMAP)

Le calcul de la longueur du connecteur reliant cet IRIS à la gare ferroviaire inclut 1052 sous-zones (874 bâtiments résidentiels + 178 sites d'activités). Selon le modèle (5), cette longueur peut être formulée telle que :

$$\|\tilde{c}_{46,3120}\| = \sum_{i=1}^{1052} d_{(z_i,3120)} \left[ \frac{\tau * f_{pop}(3120; d_{(z_i,3120)}) * Pop_{(z_i,3120)} + f_{emp}(3120; d_{(z_i,3120)}) * Emp_{(z_i,3120)}}{\sum_1^{1052} (\tau * f_{pop}(3120; d_{(z_i,3120)}) * Pop_{(z_i,3120)} + f_{emp}(3120; d_{(z_i,3120)}) * Emp_{(z_i,3120)})} \right]$$

Où 46 est le numéro de l'IRIS et 3120 le numéro du point de desserte.

### 5.1.1 CALCUL PAR LA FONCTION 'TOUT OU RIEN'

L'application de cette fonction oblige au choix d'une aire de service. Nous allons tester plusieurs configurations de cette aire. Une aire de service, par exemple, de 800m - équivalente à une distance rectilinéaire de 1111m<sup>15</sup> - aboutit à la sélection de 575 bâtiments abritant 2147 résidents et 109 sites de travail accueillant 472 emplois.

<sup>15</sup> Cette distance est convertie selon la formule de Gallez (Gallez d'après Godinot, 2008).

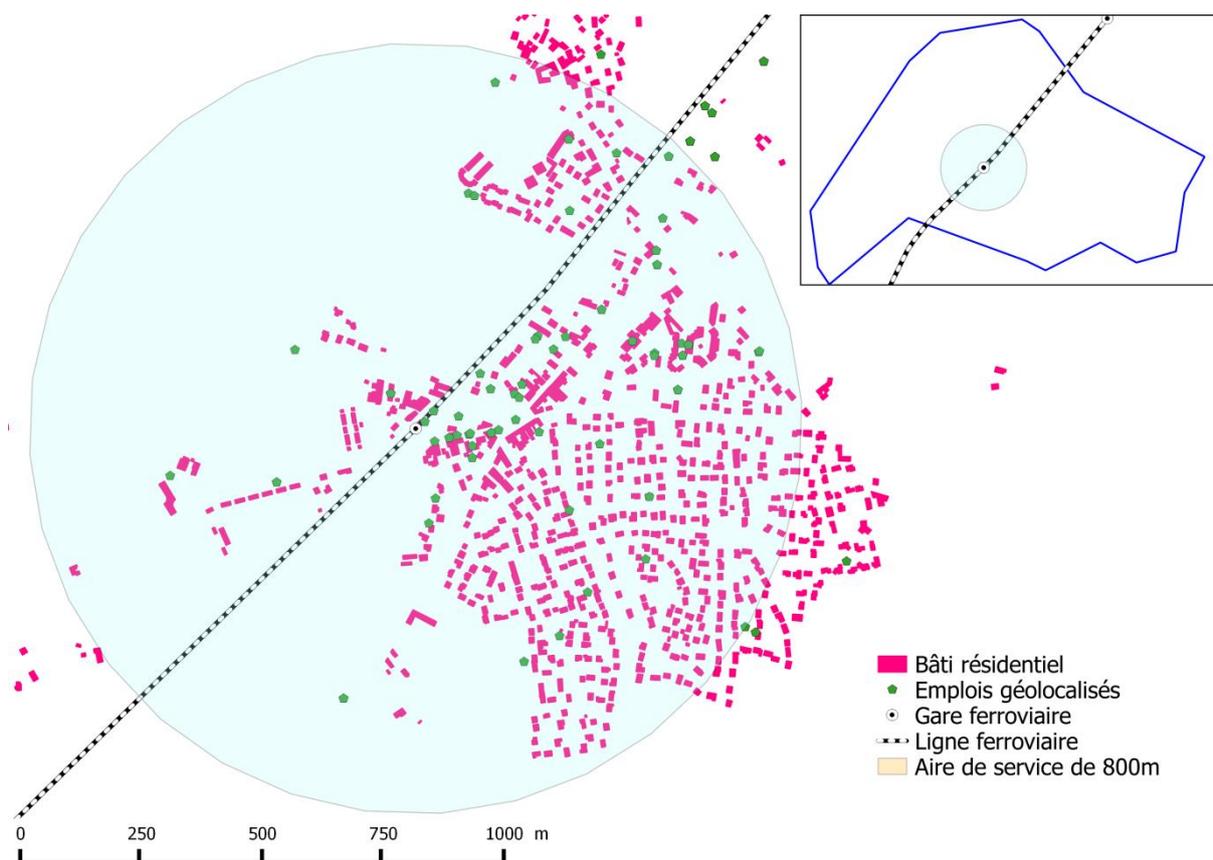


FIGURE 47 : APPLICATION DE LA FONCTION 'TOUT OU RIEN' AVEC UNE DISTANCE DE 800M

$$\|\tilde{C}_{46,3120}\| = \sum_{i=1}^{684} d_{(z_i,3120)} * \left[ \frac{\tau * Pop_{(z_i,3120)} + Emp_{(z_i,3120)}}{\sum_1^m (\tau * Pop_{(z_i,3120)} + Emp_{(z_i,3120)})} \right]$$

Avec la distance euclidienne et un ratio  $\tau = 1^{16}$ , on obtient une longueur moyenne de connecteur de :

<i>Aire de service</i>	<i>Longueur du connecteur</i>
800m	489m
1000m	546m
1200m	577m
1400m	617m

TABLE 9 : LONGUEUR DU CONNECTEUR DE L'IRIS 46 EN FONCTION DE L'AIRE DE SERVICE CHOISIE (VOIR ANNEXE.7 POUR LE CALCUL)

Il apparaît donc que la longueur du connecteur calculée par la fonction 'Tout ou rien' est sensible au rayon de l'aire de service choisi.

L'estimation de cette longueur par la méthode SIMBAD, reposant sur la position du centroïde, aboutit à un connecteur de 423m. Cette valeur est proche de celle calculée

<sup>16</sup> On a fixé ce coefficient à 1 pour toutes les fonctions, car il est, implicitement, déjà pris en considération dans le calage des fonctions de pondération. Cette hypothèse n'affecte, donc, que la fonction 'Tout ou rien'.

avec une aire de service de 800m. À notre avis, ceci est principalement dû à la nature du tissu urbain centré autour de la gare. On constatera par la suite des différences marquées entre les deux méthodes dans le cas d'un urbanisme polarisé.

### 5.1.2 CALCUL PAR LA FONCTION 'DÉGRESSIVE'

En guise d'exemple, nous allons implémenter les deux fonctions dégressives, l'agrégée et la désagrégée :

$$f_{pop/emp}(3120; d_{(Z_i,3120)}) = 1.0126 * \exp(-0.004265 * d_{(Z_i,3120)}) \quad \text{sur } Z_i \text{ de l'iris 13330000}$$

$$\begin{cases} f_{pop}(s; d_{(Z_i,s)}) = -0.0004 * d_{(Z_i,s)} + 0.5339 \\ f_{emp}(s; d_{(Z_i,s)}) = 1.3626 * \exp(-0.0026 * d_{(Z_i,s)}) \end{cases} \quad \text{sur } Z_i \text{ de l'iris 13330000}$$

L'estimation de la longueur du connecteur repose sur toutes les unités comprises dans l'IRIS. Elle se calcule, pour la première fonction, par la formule suivante :

$$\|\tilde{c}_{46,3120}\| = \sum_{i=1}^{1052} d_{(Z_i,3120)} \left[ \frac{1.0126 * \exp(-0.004265 * d_{(Z_i,3120)}) * (\tau * Pop_{(Z_i,3120)} + Emp_{(Z_i,3120)})}{\sum_1^{1052} [1.0126 * \exp(-0.004265 * d_{(Z_i,3120)}) * (\tau * Pop_{(Z_i,3120)} + Emp_{(Z_i,3120)})]} \right]$$

En utilisant la fonction agrégée, la distance euclidienne et un ratio  $\tau = 1$ , on obtient une longueur moyenne de connecteur de 297m. Sous les mêmes hypothèses et en appliquant la fonction désagrégée, la longueur est de 449m. Cet écart entre les résultats des deux fonctions peut être expliqué par la différence des poids appliqués. En effet, outre la différence des poids des emplois pour les deux fonctions, la première pondère la population par des poids dégressifs en exponentiel avec la distance, alors que la deuxième adopte des poids linéairement dégressifs.

Dans les deux cas, cette longueur est inférieure à celle calculée avec la fonction 'Tout ou rien'.

Pour faciliter l'implémentation de cette fonction, nous allons la combiner avec la fonction 'Tout ou rien'.

### 5.1.3 CALCUL PAR LA FONCTION 'COMBINÉE'

En combinant la logique des deux fonctions précédentes, la pondération 'Combinée' est plus pertinente que la fonction 'Tout ou rien' et est plus simple à implémenter que la fonction 'dégressive'.

La longueur du connecteur calculée par la fonction 'Combinée' avec une distance tampon variable est telle que :

<i>Distance tampon</i>	<i>Longueur du connecteur par la fonction agrégée</i>	<i>Longueur du connecteur par la fonction désagrégée</i>
800m	280m	398m
1000m	292m	433m
1200m	294m	443m
1400m	296m	447m

TABLE 10 : LONGUEUR DU CONNECTEUR EN FONCTION DE LA DISTANCE TAMPON CHOISIE POUR LA FONCTION COMBINÉE

Pour chaque fonction, on constate que la longueur du connecteur varie légèrement en fonction du rayon du tampon. En outre, les longueurs convergent vers une valeur proche des résultats issus de la fonction ‘Dégressive’, d’où l’intérêt de cette fonction. La connaissance de ce seuil de convergence permettra de réduire davantage les temps de calcul.

#### 5.1.4 CONCLUSION

Comme attendu, pour une même aire de service, la longueur du connecteur est estimée à la baisse en passant de la fonction ‘Tout ou rien’ à la fonction ‘Combinée’. Ce résultat est dû à la pondération différenciée de la population et des emplois pris en compte dans le calcul en fonction de la distance.

Comme rappelé plus haut, le cas d’un IRIS avec une seule ligne et un seul point de desserte est, à côté de l’exemple suivant, la brique de calcul de la présente méthodologie. Nous allons présenter par la suite d’autres cas de figure plus complexes, certes, mais dont le principe de calcul repose sur ces premiers cas simples.

## 5.2 SECOND CAS ÉLÉMENTAIRE : UN IRIS, UNE LIGNE DE TRANSPORT ET DEUX POINTS DE DESSERTE

Dans la majorité des cas, plusieurs points de desserte sont accessibles depuis un même IRIS. Ainsi, la demande se répartit entre les différents arrêts et le calcul des distances ne peut pas être approché par le précédent cas. Le cas présent illustre cette situation.

Pour les mêmes raisons présentées précédemment, nous illustrons ce cas par l’exemple d’un IRIS desservi par une ligne ferroviaire via deux gares. Il s’agit de l’IRIS de Beynost (code 5) dont le numéro est 10430000. La population de cet IRIS est estimée à 4475 et est répartie sur 1519 bâtiments. Les emplois sont au nombre de 2678 et sont répartis sur 364 structures. Il accueille deux gares ferroviaires, celle de Beynost (code 3146) et celle de Saint Maurice de Beynost (code 10029).

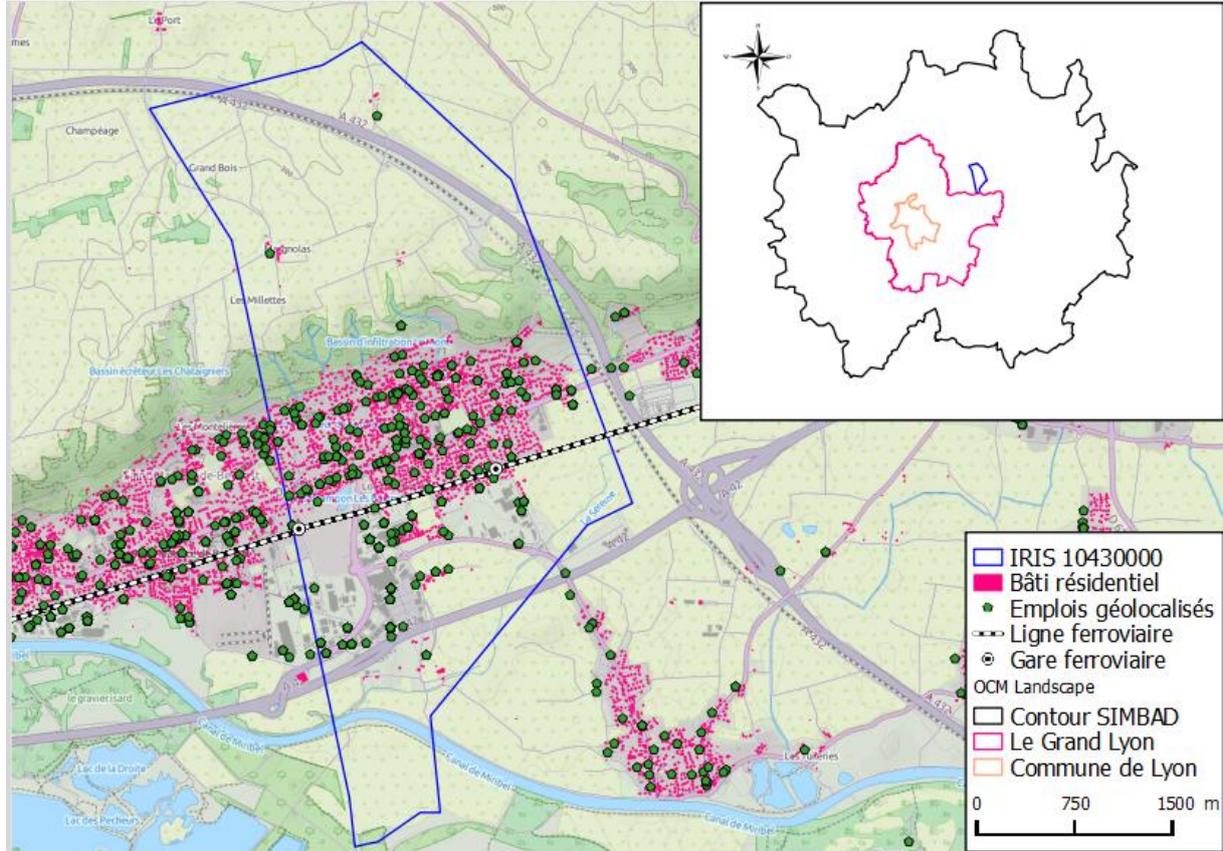


FIGURE 48 : CAS DE L'IRIS 10430000 DE BEYNOST (FOND DE CARTE OPENSTREET MAP)

Dans le cas d'un IRIS parcouru par une seule ligne de transport et desservi par plusieurs points de desserte accessibles, les usagers se répartissent entre ces différents arrêts. Le choix d'un point de desserte donné dépend de plusieurs facteurs : la distance, les conditions de déroulement de la marche à pied, la sécurité, etc. Dans notre cas, nous allons prendre la distance comme seul facteur déterminant de ce choix.

Toutes choses étant égales par ailleurs, les individus choisissent l'arrêt le plus proche à leur origine ou leur destination. Autrement dit, la population et les emplois de l'IRIS se répartissent entre ces différents arrêts en fonction de la distance.

On rappelle que  $Pnt_{desserte,I} = \cup_i \{Pnt_{desserte_i} \mid d_{(I,Pnt_{desserte_i})} \leq d_{cat\_ligne}\}$  est l'ensemble des points de desserte accessibles depuis la zone  $I$ . Soit  $Z^s_I = \{Z^{s_j}_I \mid s_j \in Pnt_{desserte,I}\}$  la répartition de l'ensemble de la population et des emplois sur ces différents points de desserte.  $Z^{s_j}_I$  est l'ensemble de la population et des emplois captifs pour la station  $j$ . On a :

$$Z^{s_j}_I = \bigcup_1^{m_j} Z_i^{s_j} = \{Z_i \in Z_I \mid d(Z_i, s_j) \leq d(Z_i, s_k); \forall k \in \llbracket 1; n \rrbracket\}$$

Avec  $d(Z_i, s_j)$  la distance entre  $Z_i$  et  $s_j$ . La partition  $Z^{s_j}_I$  dépend donc de la méthodologie de calcul de cette distance. Cette répartition forme un diagramme de Voronoï de l'IRIS dont les centres sont les points de desserte.

Pour chaque partition  $Z_I^{s_j}$ , on procède à un calcul similaire à celui détaillé au premier cas.

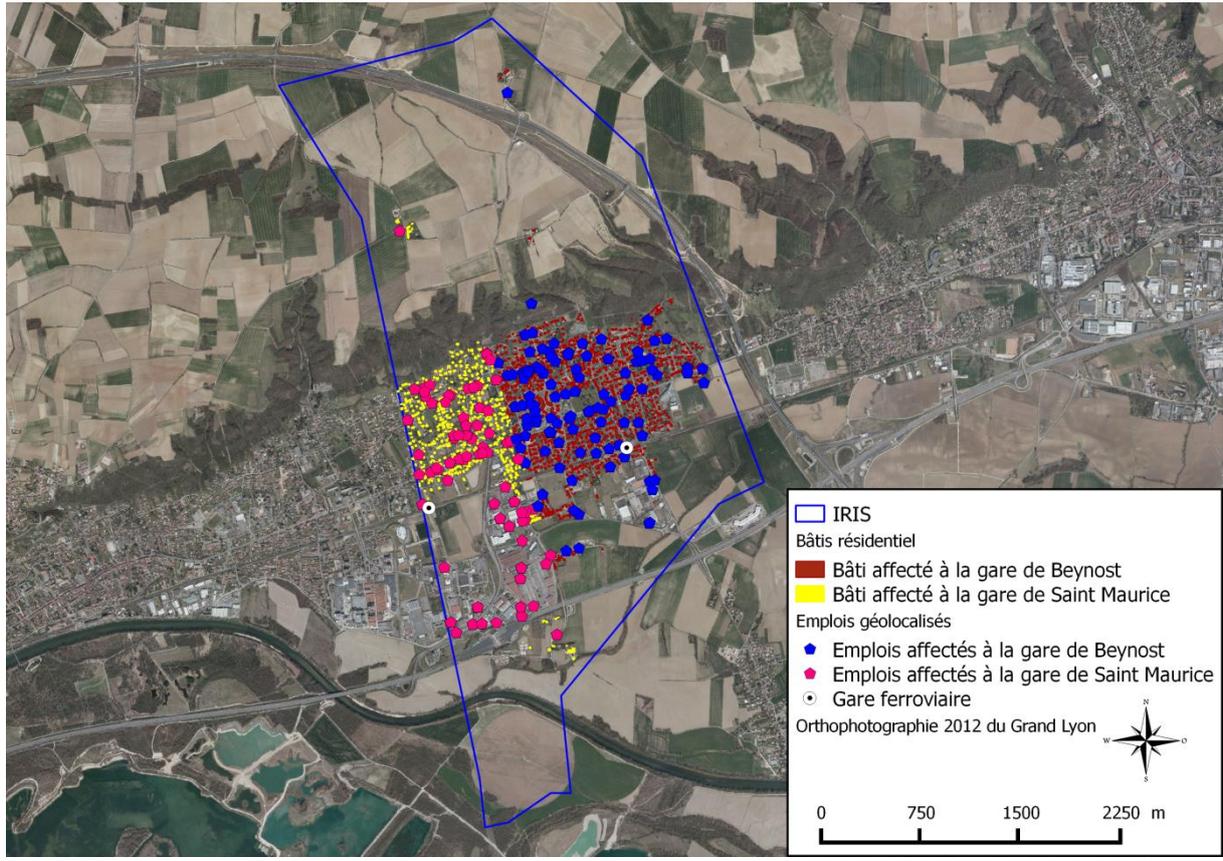


FIGURE 49 : RÉPARTITION DES ÉLÉMENTS DE L'IRIS 5 SUR 2 STATIONS (ORTHOPHOTOGRAPHIE ISSUE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012)

Dans ce cas, on applique le modèle (5) pour chaque station  $s_j$  en se limitant aux zones  $Z_i^{s_j}$  affectées à  $s_j$  :

$$\|\tilde{c}_{I,s_j}\| = \sum_{i=1}^{m_j} d_{(Z_i^{s_j}, s_j)} \left[ \frac{\tau * f_{pop}(s_j; d_{(Z_i^{s_j}, s_j)}) * Pop_{(Z_i^{s_j}, s_j)} + f_{emp}(s_j; d_{(Z_i^{s_j}, s_j)}) * Emp_{(Z_i^{s_j}, s_j)}}{\sum_1^m (\tau * f_{pop}(s_j; d_{(Z_i^{s_j}, s_j)}) * Pop_{(Z_i^{s_j}, s_j)} + f_{emp}(s_j; d_{(Z_i^{s_j}, s_j)}) * Emp_{(Z_i^{s_j}, s_j)})} \right]$$

Où  $s_j$  est un point de desserte.  $Z_I^{s_j} = \cup_1^{m_j} Z_i^{s_j}$  est l'ensemble des éléments de  $Z_I$  affectés à la station la plus proche  $s_j$ .

Comme signalé plus haut, dans le cas de plusieurs points de desserte accessibles depuis une même zone, un seul connecteur suffira pour rendre compte de l'accessibilité. La longueur moyenne du connecteur produit est alors la somme des longueurs correspondant à chaque station pondérées par les poids correspondants.

Pour simplifier l'expression du modèle (5), nous posons :

$$U_{(Z_i^{s_j}, s_j)} = \tau * f_{pop}(s_j; d_{(Z_i^{s_j}, s_j)}) * Pop_{(Z_i^{s_j}, s_j)} + f_{emp}(s_j; d_{(Z_i^{s_j}, s_j)}) * Emp_{(Z_i^{s_j}, s_j)}$$

Ainsi le modèle (5) devient :

$$\|\tilde{C}_{I,L(s)}\| = \frac{1}{\sum_{j=1}^n \sum_{i=1}^{m_j} U_{(Z_i^{s_j}, s_j)}} * \sum_{j=1}^n \left[ \left[ \sum_{i=1}^{m_j} U_{(Z_i^{s_j}, s_j)} \right] * \left[ \sum_{i=1}^{m_j} d_{(Z_i^{s_j}, s_j)} * \frac{U_{(Z_i^{s_j}, s_j)}}{\sum_{i=1}^{m_j} U_{(Z_i^{s_j}, s_j)}} \right] \right] \quad (5i)$$

Où  $n$  est le nombre de stations appartenant à la même ligne de transport et accessibles depuis le même IRIS  $I$ .  $s_j$  est une des  $n$  stations. Et  $m_j = \text{card}(Z_I^{s_j})$  est le nombre des éléments – Population et Emplois - de  $Z_I^{s_j}$ .

Dans le cas de l'IRIS de Beynost, le modèle (5i) s'écrit :

$$\|\tilde{C}_{5,L(3146)}\| = \frac{1}{\sum_{j=1}^2 \sum_{i=1}^{m_j} U_{(Z_i^{s_j}, s_j)}} * \sum_{j=1}^2 \left[ \left[ \sum_{i=1}^{m_j} U_{(Z_i^{s_j}, s_j)} \right] * \left[ \sum_{i=1}^{m_j} d_{(Z_i^{s_j}, s_j)} * \frac{U_{(Z_i^{s_j}, s_j)}}{\sum_{i=1}^{m_j} U_{(Z_i^{s_j}, s_j)}} \right] \right]$$

Avec  $s_1 = 3146$  est la gare de Beynost et  $s_2 = 10029$  est la gare de Saint Maurice de Beynost.  $Z_I^{3146}$  est constitué de  $m_1 = 1198$  unités et  $Z_I^{10029}$  de  $m_2 = 685$  sous-zones (Figure 49).

### 5.2.1 CALCUL PAR LA FONCTION 'TOUT OU RIEN'

Sous l'hypothèse d'un calcul euclidien de la distance et d'un ratio  $\tau = 1$ , on obtient une longueur moyenne de connecteur de :

<i>Distance tampon</i>	<i>Gare de Beynost</i>	<i>Gare de Saint Maurice</i>	<i>Longueur pondérée du connecteur</i>
800m	520m	584m	541m
1000m	619m	710m	656m
1200m	645m	767m	698m
1400m	649m	774m	703m

TABLE 11 : LONGUEUR DU CONNECTEUR SELON LA DISTANCE TAMPON CHOISIE

On note que la longueur du connecteur varie en fonction de l'aire de service choisie.

### 5.2.2 CALCUL PAR LA FONCTION 'DEGRESSIVE'

En utilisant la fonction agrégée comme poids et sous l'hypothèse d'un calcul euclidien de la distance et d'un ratio  $\tau = 1$ , on obtient une longueur de 405m. La fonction désagrégée aboutit à une longueur de 577m.

### 5.2.3 CALCUL PAR LA FONCTION 'COMBINÉE'

En utilisant les fonctions agrégée et désagrégée comme poids, et sous l'hypothèse d'un calcul euclidien de la distance et d'un ratio  $\tau = 1$ , on obtient :

<i>Distance tampon</i>	<i>Gare de Beynost</i>	<i>Gare de Saint Maurice</i>	<i>Longueur pondérée du connecteur</i>
800m	332m	406m	352m
1000m	361m	480m	396m
1200m	365m	498m	404m
1400m	365m	499m	405m

TABLE 12 : LONGUEUR DU CONNECTEUR PAR LA FONCTION COMBINÉE AGRÉGÉE SELON LA DISTANCE TAMPON CHOISIE

<i>Distance tampon</i>	<i>Gare de Beynost</i>	<i>Gare de Saint Maurice</i>	<i>Longueur pondérée du connecteur</i>
800m	454m	517m	473m
1000m	520m	620m	556m
1200m	531m	652m	576m
1400m	532m	654m	577m

TABLE 13 : LONGUEUR DU CONNECTEUR PAR LA FONCTION COMBINÉE DÉSAGRÉGÉE SELON LA DISTANCE TAMPON CHOISIE

L'estimation par la méthode SIMBAD aboutit au choix de la gare de Beynost comme station de connexion - Elle est la plus proche du centroïde -. Elle génère un connecteur unique d'une longueur de 664m. On constate, ainsi, les écarts entre les différentes longueurs estimées. Ces écarts auraient éventuellement un impact sur le reste de la chaîne de modélisation.

#### 5.2.4 CONCLUSION

Cet exemple illustre la méthodologie de calcul de la longueur des connecteurs dans le cas où une ligne de transport a plusieurs points de desserte accessibles depuis une même zone.

Pour étudier le cas général d'un IRIS avec plusieurs lignes de transport et plusieurs points de desserte, nous avons adopté, comme indiqué plus haut (voir § 3.1.3), l'hypothèse de non concurrence entre lignes de transport. Cette concurrence, si elle est prise en compte partiellement sous la forme de zones de service différentes et par la suite de poids différents, elle ne l'est pas pour des zones de chalandise qui se chevauchent partiellement ou complètement. Dans ce cas, chaque ligne est prise en compte indépendamment des autres.

### 5.3 CAS GÉNÉRAL

#### 5.3.1 IMPLÉMENTATION DE L'ALGORITHME

Pour illustrer le cas le général, nous avons choisi l'IRIS 693860303, dit Vitton, localisé à Lyon 6<sup>ème</sup>. Cet IRIS a une population de 2781 répartie sur 194 bâtiments et accueille 1307 emplois répartis sur 391 sites.



FIGURE 50 : CAS DE L'IRIS 693860303 DE VITTON (FOND DE CARTE OPENSTREETMAP)

Nous allons, tout d'abord, déterminer les points de desserte accessibles depuis l'IRIS. Pour ce faire, nous utilisons délibérément les valeurs suivantes de la distance d'accessibilité.

<i>Mode</i>	<i>Distance d'accessibilité</i>
Bus	300m
Tramway	400m
Métro	800m

TABLE 14 : EXEMPLE D'AIRES DE SERVICE APPLIQUÉES À L'IRIS VITTON

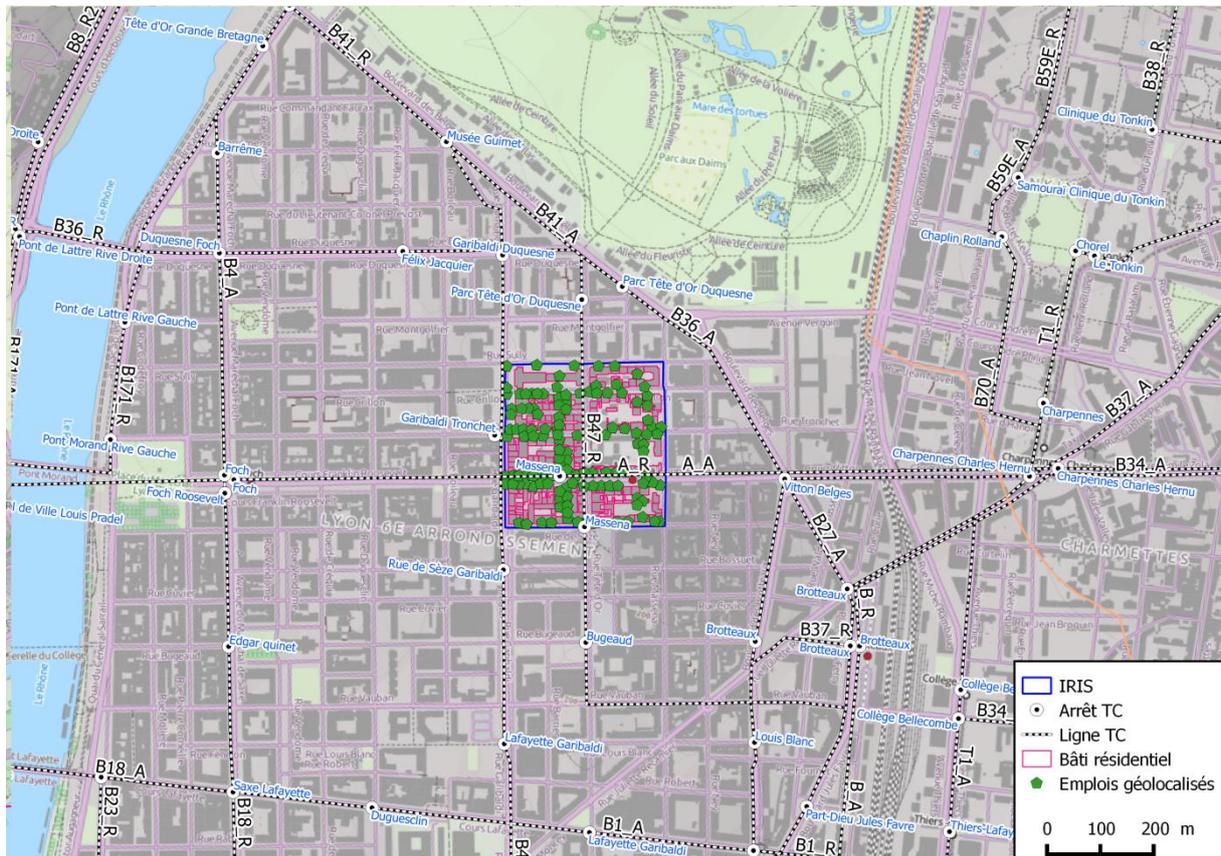


FIGURE 51 : CARTE DE SITUATION DE L'IRIS 693860303 ET DES LIGNES ET ARRÊTS À PROXIMITÉ (FOND DE CARTE OPENSTREETMAP)

L'IRIS Vitton accueille, dans ses limites, trois points de desserte : deux relatifs aux deux sens de la ligne A du métro et l'autre correspond à la ligne de bus 47\_R. Ceci dit, cette zone a accès à d'autres points de desserte, hors ses limites administratives. Compte tenu des distances d'accessibilité, les points d'arrêt : Masséna (1184), Charpennes (270) et Foch (551), desservis par la ligne A du métro, sont accessibles depuis Vitton. Le point d'arrêt Brotteaux (173) du métro B est aussi accessible. De même pour les points d'arrêt : Garibaldi Duquesne (numéro 608), Garibaldi Tronchet (611) et Rue de Sèze Garibaldi (1571), desservis par le bus 47\_A. Les points d'arrêt : Parc tête d'Or Duquesne (1328), Masséna (1185) et Bugeaud (182), desservis par la ligne de bus 47\_R, sont accessibles. Enfin, les points d'arrêt : Parc tête d'Or Duquesne (1327), Vitton Belges (1852), Brotteaux (172) et Félix Jaquier (536) desservis, dans les deux sens, par les lignes de bus : 27, 36 et 41 sont, aussi, considérés comme accessibles.

Il y aura, donc, cinq connecteurs pour cet IRIS. Chacun pour un groupe de points d'arrêts desservis par les mêmes lignes. Le calcul est similaire à celui effectué aux deux exemples élémentaires dessus.

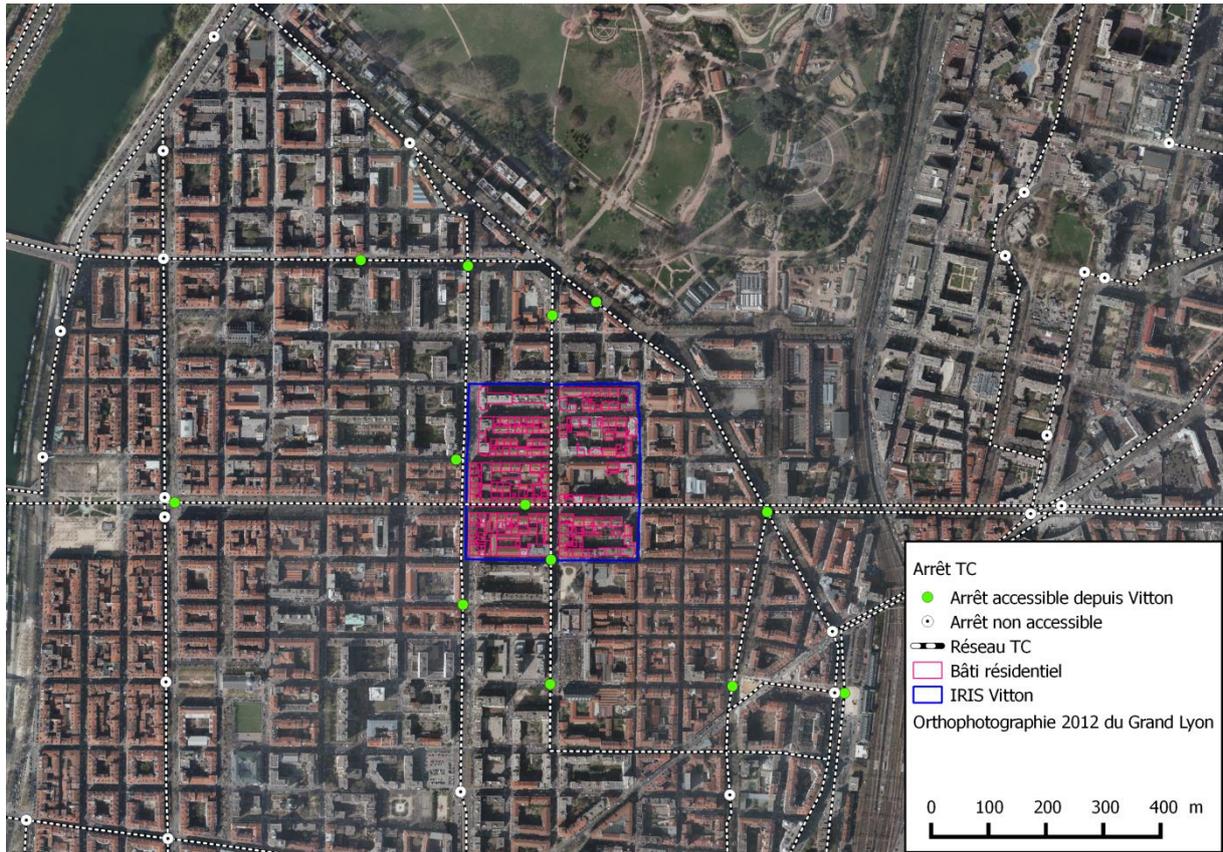


FIGURE 52 : DÉTERMINATION DES ARRÊTS ACCESSIBLES (ORTHOPHOTOGRAPHIE ISSUE DE LA PLATEFORME D'INTEGRATION DE DONNEES GEOGRAPHIQUES DU GRAND LYON, 2012)

	Connecteur ligne métro A	Connecteur ligne métro B	Connecteur ligne bus 47 A	Connecteur ligne bus 47 R	Connecteur lignes bus 27, 36 et 41
<b>Fonction 'Tout ou rien'</b>	119m	646m	162m	146m	239m
<b>Fonction 'Dégressive agrégée'</b>	105m	616m	143m	132m	290m
<b>Fonction 'Combinée agrégée'</b>	105m	615m	139m	132m	233m

TABLE 15 : LONGUEUR DE CONNECTEURS POUR DIFFÉRENTES LIGNES

La méthodologie actuellement adoptée par SIMBAD<sup>17</sup> génère, dans le cas de l'IRIS Vitton, deux connecteurs. Un reliant le centroïde à la gare Masséna (1184) et l'autre à l'arrêt Masséna du bus 47\_R. Dans ces deux cas, peu d'écart existe entre les longueurs calculées, vue l'étendue limitée de l'IRIS et son tissu urbain continu (Figure 50).

<sup>17</sup> Il s'agit de la version 2006 du réseau.

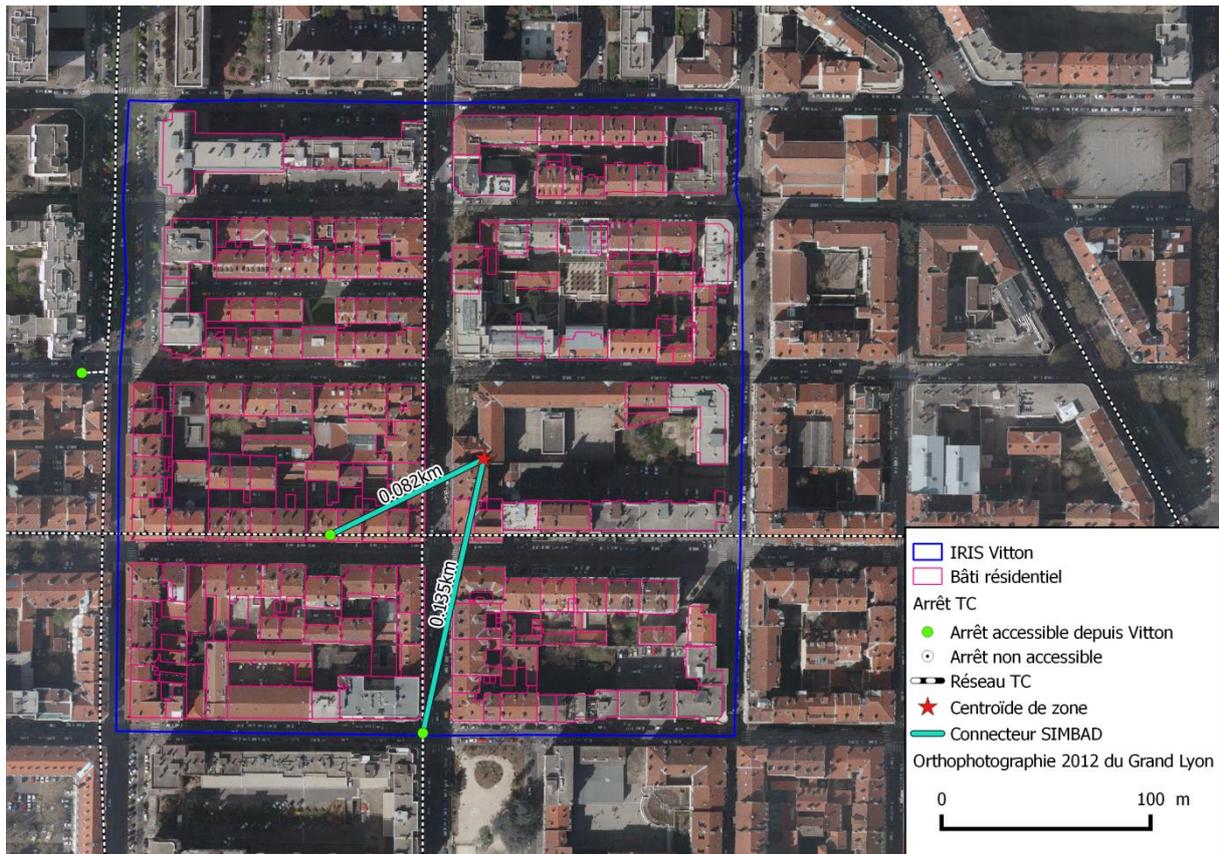


FIGURE 53 : CONNECTEURS GÉNÉRÉS PAR SIMBAD DANS LE CAS DE L'IRIS VITTON (ORTHOGRAPHIE ISSUE DE LA PLATEFORME D'INTEGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012)

### 5.3.2 CONCLUSION

Ce dernier exemple nous a permis de mettre en application la nouvelle approche dans son cas le plus généralisé. Au moment où l'ancienne méthodologie ne retient que deux connecteurs, ce qui constitue une altération de l'offre<sup>18</sup>, la nouvelle définition permet de connecter l'IRIS Vitton à cinq points d'arrêt avec des longueurs qui sont censées refléter les distances parcourues par les usagers potentiels des transports collectifs.

En guise de conclusion, on peut dire que la nouvelle codification s'affranchit de la plupart des hypothèses spatiales. En effet, elle réduit la portée de l'hypothèse d'uniformité de la densité car elle travaille à un niveau quasi-désagrégé<sup>19</sup>. Les IRIS présentant des typologies de bâti différenciées sont mieux pris en compte dans le calcul (voir 2.2.4.1). Cette méthode permet, aussi, de différencier la population de l'activité, vu la différence entre les comportements de mobilité correspondants. Cette distinction n'est pas procédée dans l'ancienne méthodologie (voir 2.2.4.1). Enfin, à l'inverse de la codification SIMBAD, le calcul désagrégé des connecteurs est totalement indépendant de la configuration spatiale et de la nature des tissus urbains – continus ou discontinus – (voir 2.2.4.2). Ainsi, la longueur des connecteurs est faiblement sensible au système de zonage adopté.

<sup>18</sup> L'algorithme d'implémentation de la méthodologie SIMBAD doit normalement connecter plus que deux points d'arrêt. Nous soupçonnons, donc, un problème d'implémentation lors de l'automatisation sous VISUM.

<sup>19</sup> On a eu besoin, dans une faible mesure, de cette hypothèse pour construire la nouvelle base désagrégée de la population localisée (voir § 4.3.2.3).

# CONCLUSION

---

Dans ce mémoire, nous avons traité la codification des connecteurs des transports collectifs et leurs impacts sur les sorties de la chaîne de modélisation. Dans ce sens, nous nous sommes particulièrement penchés sur la méthodologie utilisée par SIMBAD. Il faut dire qu'il s'agit d'une approche générique, adoptée par la plupart des modélisateurs pour sa simplicité d'implémentation. Elle repose sur le positionnement géographique des centroïdes de zones, à partir desquels des connecteurs relient les arrêts des transports en commun. De prime abord, cette pratique semble être en cohérence avec l'approximation de l'individu moyen induit par l'agrégation spatiale, toutefois, elle n'interprète pas exactement la définition des connecteurs des transports collectifs. En effet, ces derniers sont censés rendre compte de la distance moyenne parcourue par les usagers de ce système et non de la distance moyenne de toute la population de la zone. De ce fait, la longueur des connecteurs codifiés est biaisée.

Cette inexactitude au niveau des connecteurs n'est pas sans incidence sur le reste de la chaîne de modélisation. À travers un test de sensibilité conduit sur le réseau SIMBAD, il s'est avéré que les résultats de modélisation et en particulier les charges estimées lors de l'affectation sont sensibles à cette méthodologie. Dans ce sens, une variation moyenne de 5% des charges des lignes de transports collectifs a été relevée. Pour cette raison, nous avons proposé une nouvelle codification qui ne recourt pas à la notion de centroïde.

Pour s'affranchir du biais d'agrégation, cette nouvelle méthode fait appel à des données désagrégées. L'unité d'analyse n'est plus l'individu moyen mais l'individu propre. Cette approche désagrégée de la codification permet de mieux rendre compte de la distance parcourue en marche à pied depuis et vers les arrêts. Pour reconstituer la distance moyenne parcourue par les usagers potentiels, l'implémentation de cette méthode intègre une régression pondérée en fonction de l'éloignement des arrêts.

Une première application de cette nouvelle codification a été procédée dans le cadre de SIMBAD et les premiers résultats obtenus sont encourageants. Il apparaît que cette codification s'affranchit des hypothèses sous-jacentes à la méthodologie SIMBAD et est par conséquent plus pertinente. On souligne que la validation de cette nouvelle méthode nécessite une implémentation automatisée de son algorithme. Cette opération peut constituer une suite à ce travail.

Ceci dit, nous nous permettons d'énoncer quelques pistes d'amélioration pouvant faire l'objet des études à venir. En premier lieu, il nous semble intéressant de pouvoir généraliser la relation de régression en dehors de la population et des emplois pour intégrer des variables relatives à la nature de l'arrêt (arrêt de correspondance, arrêt terminal, multimodalité, etc.) et à son environnement (présence d'aménités comme les universités, les centres commerciaux, etc.). À présent, aucune distinction n'est procédée à l'intérieur de la population et des emplois. Compte tenu des données disponibles, il serait convenable d'introduire des variables de ségrégation comme l'âge, le revenu ou la nature de l'emploi. Ensuite, il serait opportun de caler des fonctions de pondération propres à l'aire d'étude et au contexte français. Ce calage serait possible grâce à une mise à disposition des données géocodées des utilisateurs des transports en commun. En dernier lieu et au vu de l'importance accordée par les usagers à l'accessibilité des transports en

commun, une implémentation de la méthodologie précédente avec la distance viaire serait intéressante.

## Annexe.1. Affectation à l'horaire

Cette définition est extraite des 'Bases théoriques' de PTV VISUM<sup>®</sup> 12.5 édité par la société PTV AG. © PTV AG.

*'Une procédure de recherche est appelée **selon les horaires** si tous les services des lignes TC sont considérés avec leurs heures exactes de départ et d'arrivée.*

*Les procédures selon les horaires conviennent pour l'affectation et le calcul d'indicateurs lorsqu'il existe un plan du réseau de lignes et un horaire détaillé pour l'offre TC étudiée. Ces procédures prennent en considération la coordination des horaires des lignes et garantissent de cette façon des résultats très précis lors du calcul des indicateurs.*

*La procédure d'affectation selon les horaires détermine des liaisons pour chaque relation O-D et suppose lors de la recherche que les voyageurs disposent de certaines informations sur les horaires et qu'ils choisissent leurs heures d'arrivée à l'arrêt de départ en concordance avec le départ de la première ligne TC. Pendant la recherche, vous pouvez influencer de façon différenciée le type de liaisons déterminées à l'aide d'un temps généralisé de recherche. Pour la recherche de liaisons, deux variantes sont proposées (algorithme Branch & Bound et recherche du plus court chemin), qui constituent chacune un compromis différent entre le nombre d'alternatives d'une part et l'espace mémoire et le temps de calcul requis d'autre part.*

*Les liaisons déterminées par l'algorithme de recherche sont une nouvelle fois examinées selon des critères généraux lors de la présélection de liaisons afin de déterminer si certaines liaisons sont de qualité moindre et peuvent être supprimées.*

*Lors du choix de liaisons, la demande de déplacements est répartie sur les alternatives retenues selon un des modèles cités auparavant. Au choix, on peut tenir compte de l'indépendance des liaisons.'*

## Annexe.2. Enquête à bord de la Floride du sud

Ci-après le questionnaire administré aux usagers des transports collectifs lors de l'enquête à bord en l'an 2000. Notons que Zhao et al. (Zhao et al., 2003) ont géocodé les origines et les destinations pour aboutir au calage de la fonction de pondération dégressive agrégée.

Les questionnaires sont extraits du rapport synthétique 'Southeast Florida Regional Travel Characteristics Study' préparé par CARR SMITH CORRADINO et al. (Carr Smith Corradino et al., 2000)

**ID#**

The FDOT is planning for the future. To do that we need to learn more about your experience on public transportation in southeast Florida. We would appreciate it if you would take a couple of minutes right now to complete this survey and return it to the surveyor as you leave the bus. Please complete as many questions as your time allows. If you don't have time to complete the survey, you can take it home and mail it in (no stamp required). We don't need your name and all information is confidential.

A. How many people live in your household? (Please circle one.)  
                   1        2        3        4        5+

B. How many children under 18 are in your household? (Please circle one.)  
                   0        1        2        3        4+

C. How many people in your household have a job outside the home during the weekday? (Please circle one.)  
                   0        1        2+

D. How many vehicles (cars, pick up trucks, motorcycles, etc.) are available to the drivers in your household? (Please circle one.)  
                   0        1        2        3+

E. Where were you before you got on the first bus or train for this trip?

I WAS AT:

(Circle One)

- Home
- Work
- Shopping
- Social/Recreational
- School
- Other

*Address of the place that you circled*

---

*(street address [or corner or building], city, zip) or  
nearest major intersection*

---

*name of place or nearby landmark*

**VERY  
IMPORTANT**

F. What time was it when you left the place you circled in Question E above?  
                   \_\_\_\_\_ a.m.    \_\_\_\_\_ p.m.

G. Approximately how far did you walk or drive to get to the bus stop or train station (Please circle one)?  
                   Walked (in blocks): 3 or less    4 - 8    9+    or    Drove (in miles): 1 or less    1 - 3    3+

H. How did you get to the bus or train from the place you circled in Question E above? (Please circle one.)  
                   Walked                    Drove Auto                    Dropped off by Auto                    Other

I. How many minutes did you wait for the bus or train (Please circle one)?  
                   0 - 5                    6 - 10                    11 - 20    20+ minutes

**OVER**

J. What fare did you pay to ride the bus or train? (Please check one.)

Full Cash Fare \_\_\_\_\_ Discounted Cash Fare \_\_\_\_\_ Discounted Pass \_\_\_\_\_ Token \_\_\_\_\_

K. Where are you going after you complete this trip?

**VERY  
IMPORTANT**

**I WILL BE AT:**

(Circle One)

Home

Work

Shopping

Social/Recreational

School

Other

Address of the place that you circled

\_\_\_\_\_  
(street address [or corner or building], city, zip) or  
nearest major intersection

\_\_\_\_\_  
name of place or nearby landmark

L. How will you get from your final stop to the place you are going, that is, the place you circled in Question K, above? (Please circle one.)

Walk

Drive Auto

Dropped off by Auto

Other

M. Approximately how far will you travel from your final stop to the place you are going (Please circle one)?

Walked (in blocks): 3 or less 4 – 8 9+ or Drove (in miles): 1 or less 1 – 3 3+

N. Does the trip you are now on require you to transfer between buses or trains? (Please circle one.)

No

Once

More than Once

O. On this trip what kind of transit vehicles did you use? (Please circle one.)

Buses only

Trains only

Both buses and trains

P. If a bus or train were not available, how would you make this trip? (Please circle one.)

Taxi

Auto

Walk

Trip would not be made

Q. What would you say that your total household income is (Please circle one)?

\$15,000 or less

\$15,000 to \$30,000

\$30,000 to \$50,000

\$50,000 or more

## Annexe.3. Enquête téléphonique de Montréal

Cette enquête Origine-Destination a été menée sur la grande région de Montréal en 2003 pour comprendre les pratiques d'usage des transports en commun (Agence Métropolitaine de Transport, 2012; le secrétariat à l'enquête Origine-Destination, 2003). © 2003 AMT.

La figure suivante résume la structure logique du questionnaire utilisé pour la présente enquête O-D.

<b>QUESTIONNAIRE SCHÉMATISÉ</b> <b>Enquête Origine-Destination 2003</b>															
<b>Introduction</b>	Bonsoir, Monsieur (Madame) ... Mon nom est ... je vous appelle au nom de Léger Marketing qui effectue présentement une étude sur les habitudes de transport des ménages de la région de Montréal pour le compte de l'AMT et du ministère des Transports du Québec. <u>Avez-vous entendu parler de cette enquête par le biais de votre courrier, des journaux, ... ? (Sinon, informer)</u>														
<b>Questionnaire</b>	<p>Votre adresse est bien le ... ?</p> <p>Combien y a-t-il de personnes qui habitent de façon permanente dans votre logis, y compris vous-même ?</p> <p>Combien y a-t-il d'automobiles à la disposition des membres de votre logis ?</p> <p>Depuis quand habitez-vous à cet endroit ?</p> <p>Où habitez-vous auparavant ? (Si déménagement dans les cinq dernières années)</p> <div style="border: 1px solid black; padding: 5px;"> <p><b>I</b></p> <p>Quel âge avez-vous ? Est-ce un homme ou une femme (garçon ou fille) ? (pour les autres personnes)</p> <p>Détenez-vous un permis de conduire ?</p> <p>Quelle est votre occupation principale ? (Travailleur, étudiant, retraité ou autre) (Si travailleur ou étudiant) Quel est votre lieu d'emploi / d'étude habituel ?</p> </div> <p><b>II</b></p> <p>Avez-vous effectué une sortie au cours de la journée d'hier ? (Si non) Pourquoi ne vous êtes-vous pas déplacé pour votre travail / étude ?</p> <p>Pour chacune des autres personnes du logis, demander la séquence de questions précédentes à partir de I</p> <p>Poser les questions suivantes pour toutes les personnes qui ont répondu « Oui » à la question II, en commençant par le répondant.</p> <p>Concernant la première sortie que vous avez effectuée hier, à quelle heure êtes-vous parti ?</p> <div style="border: 1px solid black; padding: 5px;"> <p><b>III</b></p> <p>Pour quel motif avez-vous effectué cette sortie ?</p> <p>De quel endroit êtes-vous parti et à quel endroit êtes-vous allé lors de cette sortie ? (Place, adresse, intersection, municipalité)</p> <p>Par quel moyen de transport avez-vous effectué ce déplacement ?</p> <table border="0" style="width: 100%;"> <thead> <tr> <th style="text-align: left;"><u>Réponse</u></th> <th style="text-align: left;"><u>Question</u></th> </tr> </thead> <tbody> <tr> <td>Automobile</td> <td>Comme conducteur ou comme passager ? Qui était le conducteur ? Pont franchi ? Autoroute(s) utilisée(s) ? Type de stationnement à destination ?</td> </tr> <tr> <td>Autobus</td> <td>Quelle(s) ligne(s) d'autobus avez-vous utilisée(s) ?</td> </tr> <tr> <td>Méto</td> <td>À quelles stations êtes-vous monté et descendu ? Y a-t-il eu correspondance ?</td> </tr> <tr> <td>Train</td> <td>À quelles gares êtes-vous monté et descendu ?</td> </tr> <tr> <td>Alternance modes public et privé</td> <td>À quel endroit avez-vous changé de mode de transport ?</td> </tr> <tr> <td>Si transport en commun</td> <td>De quelle façon avez-vous payé votre passage ?</td> </tr> </tbody> </table> <p>Avez-vous effectué une autre sortie au cours de la journée d'hier ?</p> <p>Oui : À quelle heure ? Retourner à la question III</p> <p>Non : Passer à la personne suivante s'étant déplacée (question III, à la 3<sup>e</sup> personne)</p> </div>	<u>Réponse</u>	<u>Question</u>	Automobile	Comme conducteur ou comme passager ? Qui était le conducteur ? Pont franchi ? Autoroute(s) utilisée(s) ? Type de stationnement à destination ?	Autobus	Quelle(s) ligne(s) d'autobus avez-vous utilisée(s) ?	Méto	À quelles stations êtes-vous monté et descendu ? Y a-t-il eu correspondance ?	Train	À quelles gares êtes-vous monté et descendu ?	Alternance modes public et privé	À quel endroit avez-vous changé de mode de transport ?	Si transport en commun	De quelle façon avez-vous payé votre passage ?
<u>Réponse</u>	<u>Question</u>														
Automobile	Comme conducteur ou comme passager ? Qui était le conducteur ? Pont franchi ? Autoroute(s) utilisée(s) ? Type de stationnement à destination ?														
Autobus	Quelle(s) ligne(s) d'autobus avez-vous utilisée(s) ?														
Méto	À quelles stations êtes-vous monté et descendu ? Y a-t-il eu correspondance ?														
Train	À quelles gares êtes-vous monté et descendu ?														
Alternance modes public et privé	À quel endroit avez-vous changé de mode de transport ?														
Si transport en commun	De quelle façon avez-vous payé votre passage ?														
<b>Conclusion</b>	S'il s'agit de la dernière personne du logis s'étant déplacée : Pour des fins statistiques, pourriez-vous m'indiquer le revenu global de votre ménage ? Merci de votre précieuse collaboration.														



## Annexe.4. Base des données carroyées

Dans cette annexe, on intègre fidèlement le descriptif synthétique de la base carroyée éditée par l'Insee. © 2013 Insee.



### Diffusion de données carroyées - novembre 2013

#### Documentation synthétique sur les données carroyées à 200 mètres

Les données carroyées (carreaux de 200 mètres) sont issues de la source Revenus Fiscaux Localisés 2010. Elles sont utiles pour disposer d'information à des niveaux infra-communaux selon des découpages propres aux différents utilisateurs. La diffusion de ces données doit donc respecter la règle d'au moins 11 ménages par unité géographique d'observation.

La méthode mise en œuvre afin de respecter cette règle consiste à regrouper les carreaux de trop faible effectif, en rectangles de taille plus importante. Seule la population totale n'est pas soumise à cette règle et peut donc être diffusée sur l'ensemble des carreaux.

Par ailleurs, les variables considérées comme sensibles, en particulier le revenu fiscal des ménages, sont traitées afin de garantir le respect de la confidentialité.

Pour les DOM, seules la Réunion et la Martinique sont concernées par cette diffusion de données carroyées.

#### Constitution des rectangles

Le territoire métropolitain<sup>1</sup> a préalablement été divisé en 36 grands carreaux de taille semblable.

Les rectangles sont constitués par subdivision progressive de ces grands carreaux, le processus itératif s'arrêtant dès qu'il conduit à un rectangle de moins de 11 ménages.

Pour la métropole, 698 659 rectangles d'au moins 11 ménages ont ainsi été constitués.

Sur les 2 278 213 carreaux de 200 mètres habités, 20,3 % comptent plus de 10 ménages (12 % sont diffusés comme tels et 8,3 % se retrouvent regroupés avec d'autres carreaux dans des rectangles).

#### Traitement préalable des revenus fiscaux par unité de consommation des individus

Afin de garantir la confidentialité en évitant de faire apparaître des zones constituées exclusivement de hauts ou de bas revenus, les revenus fiscaux des ménages par u.c. ont été préalablement traités pour les valeurs extrêmes (procédure dite de « winsorisation »).

Si le revenu fiscal par u.c. du ménage est supérieur au 8<sup>e</sup> décile de la distribution (29 336 euros en métropole, 25 098 euros en Martinique et 23 131 euros à la Réunion), alors ce revenu est rabaisé à ce seuil.

Si le revenu fiscal par u.c. du ménage est inférieur à 40 % de la médiane de la distribution (7 499,6 euros en métropole, 5 209,2 euros en Martinique et 4 136 euros à la Réunion), alors ce revenu est remonté à ce seuil.

---

<sup>1</sup> Sauf la Corse.

## Traitement des statistiques au rectangle pour les variables sensibles

Quelques variables, considérées comme sensibles, sont traitées afin que les informations diffusées ne correspondent pas à une proportion dépassant 80 % (et inférieure à 20 % pour le nombre de ménages propriétaires) :

- nombre d'individus âgés de 65 ans et plus
- nombre de ménages d'une personne
- nombre de ménages propriétaires
- nombre de ménages dont le revenu fiscal par u.c. se situe en dessous du seuil de bas revenu<sup>2</sup>.

Pour respecter ces règles, dans certains cas, l'effectif réel n'a donc pu être diffusé. Ceci est signalé à l'utilisateur de la manière suivante : à chacune de ces variables sensibles est associée dans le fichier diffusé une variable indicatrice, valant 1 si le nombre « n » diffusé doit être compris comme « n ou plus », et 0 sinon.

Dans le cas particulier du nombre de ménages propriétaires, l'indicatrice de traitement prend la valeur 2 si le nombre « n » de ménages propriétaires diffusé doit être compris comme « n ou moins ».

## Contenu des fichiers diffusés

Pour chacun des trois territoires (Métropole, Martinique, Réunion), 2 fichiers sont mis à disposition : un fichier de carreaux et un fichier de rectangles ([voir dessin des fichiers dans la documentation complète](#)).

Les grilles de carreaux sont fournies au niveau départemental.

**Variables diffusées dans les fichiers de carreaux :**

Pour chaque carreau habité :

- Les identifiants géographiques du carreau
- Le nombre de carreaux habités du rectangle auquel appartient le carreau
- Les identifiants géographiques du rectangle auquel appartient le carreau
- Le nombre d'individus du carreau

**Variables non traitées diffusées dans les fichiers de rectangles :**

- Le nombre total d'individus
- Le nombre total de ménages (i.e. de résidences principales)
- Le nombre total d'individus âgés de 0 à 3 ans
- Le nombre total d'individus âgés de 4 à 5 ans
- Le nombre total d'individus âgés de 6 à 10 ans
- Le nombre total d'individus âgés de 11 à 14 ans
- Le nombre total d'individus âgés de 15 à 17 ans
- Le nombre total d'individus âgés de 25 ans et plus
- Le nombre total de ménages de 5 personnes et plus
- Le nombre total de ménages présents depuis 5 ans ou plus dans leur logement actuel
- Le nombre total de ménages en logement collectif
- La surface cumulée des résidences principales, en mètres carrés
- La somme des revenus fiscaux par u.c. winsorisés des individus (le revenu fiscal par u.c. d'un individu est le revenu par u.c. du ménage auquel il appartient)<sup>3</sup>

<sup>2</sup> Le seuil de bas revenu est égal à 60% de la médiane de la distribution des revenus par u.c. soit 11 249,4 euros en métropole, 7 813,8 euros en Martinique et 6 204 euros à La Réunion

<sup>3</sup> Cette somme doit être divisée par le nombre total d'individus du rectangle pour obtenir le revenu fiscal par UC moyen des individus.

## **Variables traitées diffusées dans les fichiers de rectangles :**

- Le nombre total d'individus âgés de 65 ans et plus + la variable indicatrice correspondante
- Le nombre total d'individus âgés de 75 ans et plus + la variable indicatrice correspondante ;
- Le nombre total de ménages d'une personne + la variable indicatrice correspondante
- Le nombre total de ménages propriétaires + la variable indicatrice correspondante
- Le nombre total de ménages dont le revenu fiscal par u.c. se situe en dessous du seuil de bas revenu (60 % de la médiane de la distribution) + la variable indicatrice correspondante

## **Utilisation des fichiers**

**Le fichier de rectangles est un fichier intermédiaire : il ne doit pas être utilisé en tant que tel, en particulier pour faire des cartes.** Les rectangles étant de superficie variable, des cartes d'effectifs seraient en effet erronées. Par ailleurs, pour les rectangles de taille relativement importante, la valeur globale masque d'éventuelles fortes disparités spatiales internes. Il convient donc de travailler au niveau des carreaux.

**Pour cartographier les effectifs totaux d'individus,** le fichier de carreaux diffusé peut être utilisé tel quel.

**Pour cartographier les autres variables,** il faut construire, à partir des deux fichiers livrés (rectangles avec les variables et carreaux sans les variables), un nouveau fichier de carreaux en répartissant, pour chaque variable, l'effectif total du rectangle dans chacun de ses carreaux habités, au prorata de la population totale ([voir modes opératoires Mapinfo et Qgis dans la documentation complète](#)).

## Annexe.5. Présentation de la BD TOPO®

Cette annexe est tirée sans modification du descriptif IGN de la de la BD Topo® Version 2. © IGN.

<b>Définition</b>	Bâtiment de plus de 20 m <sup>2</sup> , ne possédant pas de fonction particulière pouvant être décrit dans les autres classes de bâtiments surfaciques (voir 8.2, 8.3, 8.4) : bâtiments d'habitation, d'enseignement... (voir détails dans § sélection)	
<b>Topologie</b>	Simple	
<b>Genre</b>	Surface 3D	
<b>Attributs</b>	ID PREC_PLANI PREC_ALTI ORIGIN_BAT HAUTEUR Z_MIN Z_MAX	Identifiant du bâtiment Précision planimétrique Précision altimétrique Source du bâtiment Hauteur du bâtiment Altitude minimale du bâtiment Altitude maximale du bâtiment

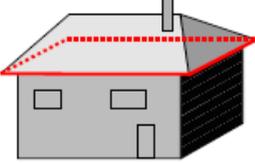
**Sélection** : Bâtiments d'habitation, bergerie, borie, bungalow, bureaux, chalet, bâtiments d'enseignement, garage individuel, bâtiments hospitaliers, immeubles collectifs, lavoirs couverts, musée, prison, refuge, villages de vacances.

Taille du bâtiment	Sélection
Moins de 20 m <sup>2</sup>	Les bâtiments de moins de 20 m <sup>2</sup> sont exclus au moment de la saisie photogrammétrique. S'ils sont très hauts, ou s'ils sont spécifiquement désignés sur la carte au 1 : 25 000 en cours (ex. monument, antenne,...) : ils sont dans ce cas représentés par un objet de classe CONSTRUCTION_PONCTUELLE. S'ils sont présents dans la BDParcelleaire®, ils sont alors inclus lors de l'appariement de la BDTopo® avec la BDParcelleaire®. Les petits bâtiments d'aspect précaire (cabanes de chantier, petits abris pour animaux,...) sont exclus ici : ils sont regroupés dans la classe CONSTRUCTION_LEGERE.
Entre 20 et 50 m <sup>2</sup>	Les bâtiments faisant entre 20 et 50 m <sup>2</sup> sont sélectionnés en fonction de leur environnement et de leur aspect : Les petits bâtiments isolés (plus de 100 m d'une habitation) de plus de 20 m <sup>2</sup> sont inclus, alors que les petits bâtiments situés entre 20 et 50 m <sup>2</sup> en ville ne le sont pas (ex : petit garage individuel, petit atelier, annexes diverses).

**Modélisation géométrique** : La modélisation géométrique peut être de deux types suivant que le bâtiment est issu initialement de la BD TOPO® (c'est-à-dire principalement obtenu par restitution photogrammétrique à partir d'une prise de vue aérienne), ou que celui-ci est obtenu après intégration des données du cadastre. Les deux possibilités coexistent actuellement dans la BD TOPO®, jusqu'à intégration complète des bâtiments du cadastre.

**Remarque concernant la géométrie** : La géométrie des bâtiments de la BD TOPO® n'est pas exactement superposable avec la géométrie des bâtiments de la BD PARCELLAIRE®, pour laquelle la source de données est différente (directement issue du cadastre).

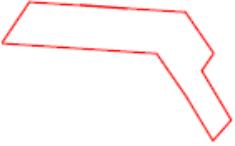
**Modélisation géométrique AVANT intégration du bâti du cadastre :** Contour extérieur du bâtiment tel qu'il apparaît vu d'avion (le plus souvent, ce contour correspond à celui du toit). L'altitude correspondant à ce contour est l'altitude de l'arête supérieure en cas de façade verticale (généralement l'altitude des gouttières). Seules les cours intérieures de plus de 10 m de large sont représentées par un trou dans la surface bâtie.

Description	Monde réel et modélisation	Modélisation géométrique
<u>Modélisation d'une maison</u>		

Plusieurs bâtiments contigus ou superposés de même « fonction » sont généralement considérés comme un seul et même objet (seul le contour extérieur est saisi).

Deux objets sont représentés dans le cas où les bâtiments contigus ou superposés présentent les caractéristiques suivantes :

- différence de hauteur entre les deux bâtiments > 10 m environ (ou 3 étages)
- surface de chaque objet résultant > 400 m<sup>2</sup>

Description	Monde réel et modélisation	Modélisation géométrique	Modélisation géométrique
<u>Modélisation d'un bâtiment</u>		Perspective 	Vue de dessus 

**Modélisation géométrique APRES intégration du bâti du cadastre :** Contour extérieur du bâtiment tel qu'il est défini dans le cadastre. La planimétrie est obtenue par translation de la géométrie issue du cadastre de manière à se superposer au mieux au bâtiment auquel elle correspond dans la BD TOPO®.

L'altitude correspondant au contour est une altitude toit médiane, calculée par interpolation, en prenant en compte les altitudes des contours des bâtiments directement contigus si ils existent.

En effet, après intégration des bâtiments issus du cadastre, plusieurs bâtiments contigus sont cette fois-ci considérés comme autant d'objets si la géométrie du cadastre en fait état.

Description	Monde réel et modélisation	Modélisation géométrique	Modélisation géométrique
<u>Modélisation d'un bâtiment</u>		Perspective 	Vue de dessus 

**Contrainte de modélisation :** Deux bâtiments d'attributs distincts adossés l'un à l'autre sont modélisés par deux surfaces ayant un côté commun. L'altitude de ce côté est celle du bâtiment le plus élevé.

- ID

**Définition :** Identifiant du bâtiment.  
Cet identifiant est unique. Il est stable d'une édition à l'autre.

**Type :** Caractères

**Contrainte sur l'attribut :** Valeur obligatoire

- PREC\_PLANI

**Définition :** Précision géométrique planimétrique.  
Attribut précisant la précision géométrique en planimétrie de la donnée.

**Type :** Décimal

**Contrainte sur l'attribut :** Valeur obligatoire

**Valeurs de l'attribut :**

Source des données	Précision	Valeurs de l'attribut
Levé GPS de précision	Inférieure à 0,5 m	0.5
Photogrammétrie, plan ou fichier métrique	0,5 à 1,5 m	1.5
Photogrammétrie des PVA longue focale, levé GPS dynamique, levé terrain	1,5 à 2,5 m	2.5
Orthophotographie, plan ou fichier non métrique, BD TOPO® version antérieure	2,5 à 5 m	5
Carte 1/25000	5 m à 10 m	10
BD CARTO®	Supérieure à 10 m	30

- PREC\_ALTI

**Définition :** Précision géométrique altimétrique.  
Attribut précisant la précision géométrique en altimétrie de la donnée.

**Type :** Décimal

**Contrainte sur l'attribut :** Valeur obligatoire

**Valeurs de l'attribut :**

Source des données	Précision	Valeurs de l'attribut
Photogrammétrie	Inférieure à 1 m	1
GPS	1 m < précision < 2,5 m	2.5
Interpolation MNT	2,5 m < précision	20
Sans Z		9999

- **ORIGIN\_BAT**

**Définition :** Attribut précisant d'où est issu le bâtiment, de quelle base, plan ou levé.

**Type :** Caractères

**Valeurs de l'attribut :**

BDTopo
Cadastre
Terrain
Autre
NR

- **HAUTEUR**

**Définition :** Hauteur du bâtiment correspondant à la différence entre le z le plus élevé du pourtour du bâtiment et un point situé au pied du bâtiment. La hauteur est arrondie au mètre.

Description	Monde réel et modélisation
Signification de l'attribut hauteur	

**Type :** Entier

**Contrainte sur l'attribut :** Valeur obligatoire

**Valeurs de l'attribut :** 0 / 2 / 3 / 4 ...

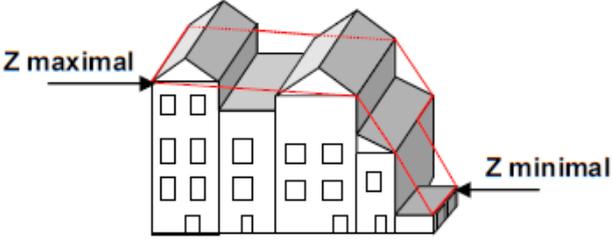
(La valeur 1 est volontairement exclue, ne correspondant pas à la réalité terrain).

- **Z\_MIN**

**Définition :** Altitude minimale d'un objet linéaire ou surfacique. Cette information est issue de la saisie photogrammétrique et concerne des objets dont les points initiaux et finaux ne sont pas caractéristiques de l'altitude de l'objet (objets surfaciques, objets linéaires isolés). Pour des objets surfaciques comme les bâtiments ou les réservoirs, il s'agit de l'altitude minimum du pourtour tel qu'il est défini en planimétrie (généralement l'altitude aux gouttières) et non de l'altitude minimum absolue.

Dans le cas d'un bâtiment obtenu après intégration du bâti issu du cadastre, Z\_MIN correspond à une altitude toit médiane calculée, en prenant en compte les altitudes des contours des bâtiments directement contigus s'ils existent.

Dans ce cas Z\_MAX et Z\_MIN prennent la même valeur.

Description	Monde réel et modélisation
Signification des attributs Z minimal et Z maximal pour un bâtiment	

Type : Décimal

Valeurs : L'attribut peut prendre une valeur conventionnelle (-99) lorsque la géométrie de l'objet n'est pas issue d'une saisie photogrammétrique (ex. objets numérisés sur la carte en service).

- **Z\_MAX**

**Définition :** Altitude maximale d'un objet linéaire ou surfacique. Cette information est issue de la saisie photogrammétrique et concerne des objets dont les points initiaux et finaux ne sont pas caractéristiques de l'altitude de l'objet (objets surfaciques, objets linéaires isolés). Pour des objets surfaciques comme les bâtiments ou les réservoirs, il s'agit de l'altitude maximum du pourtour tel qu'il est défini en planimétrie (généralement l'altitude aux gouttières) et non de l'altitude maximum absolue (voir schéma ci-dessus).

Dans le cas d'un bâtiment obtenu après intégration du bâti issu du cadastre, Z\_MAX correspond à une altitude toit médiane calculée, en prenant en compte les altitudes des contours des bâtiments directement contigus s'ils existent. Dans ce cas Z\_MAX et Z\_MIN prennent la même valeur.

Type : Décimal

Valeurs : L'attribut peut prendre une valeur conventionnelle (-99) lorsque la géométrie de l'objet n'est pas issue d'une saisie photogrammétrique (ex. objets numérisés sur la carte en service).

## Annexe.6. Base SIRENE®

Ci-après le descriptif des variables SIRENE®. © Insee.

### Description de la structure de la base de données SIRENE et des fichiers de mise à jour

#### 1 - Dessin du fichier NOTICES 80

NOM	LIBELLE	LONGUEUR	POSITION DEBUT	POSITION FIN
<b>IDENTIFICATION - ADRESSE</b>				
SIRET	Identifiant de l'établissement	14	1	14
L1_NOMEN	Nom ou raison sociale de l'entreprise pour l'adressage	38	15	52
L2_COMP	Complément de nom de l'entreprise pour l'adressage	38	53	90
L3_CADR	Complément d'adresse pour l'adressage	38	91	128
L4_VOIE	Numéro et libellé dans la voie	38	129	166
L5_DISP	Distribution spéciale	38	167	204
L6_POST	Ligne d'acheminement postal pour l'adressage	38	205	242
L7_ETRG	Libellé du pays pour les adresses à l'étranger	38	243	280
ZR1	Zone réservée	11	281	291
<b>LOCALISATION GEOGRAPHIQUE DE L'ETABLISSEMENT</b>				
RPET	Région de localisation de l'établissement	2	292	293
DEPET	Département de localisation de l'établissement	2	294	295
ARRONET	Arrondissement de localisation de l'établissement	2	296	297
CTONET	Canton de localisation de l'établissement	3	298	300
COMET	Commune de localisation de l'établissement	3	301	303
LIBCOM	Libellé de la commune de localisation de l'établissement	32	304	335
DU	Département de l'unité urbaine de la localisation de l'établissement	2	336	337
TU	Taille de l'unité urbaine	1	338	338
UU	Numéro de l'unité urbaine	2	339	340
CODPOS	Code postal	5	341	345
ZR2	Zone réservée	8	346	353
TCD	Tranche de commune détaillée	2	354	355
ZEMET	Zone d'emploi	2	356	357
CODEVOIE	Code voie	5	358	362
NUMVOIE	Numéro dans la voie	4	363	366
INDREP	Indice de répétition	1	367	367
TYPVOIE	Type de la voie de localisation de l'établissement	4	368	371
LIBVOIE	Libellé de la voie de localisation de l'établissement	32	372	403
<b>CARACTERISTIQUES ECONOMIQUES DE L'ETABLISSEMENT</b>				
ENSEIGNE	Enseigne ou nom de l'exploitation	40	404	443
APET700	Activité principale de l'établissement en 700 classes	5	444	448
ZR3 <sup>1</sup>	Zone réservée	2	449	450
SIEGE	Qualité de siège ou non de l'établissement	1	451	451
TEFET	Tranche d'effectif salarié de l'établissement	2	452	453
EFETCENT	Effectif salarié de l'établissement à la centaine près	6	454	459
ORIGINE	Origine de la création de l'établissement	2	460	461
DCRET	Date de création de l'établissement (année, mois)	6	462	467
MMINTRET	Mois d'introduction de l'établissement dans la base de diffusion	2	468	469
ACTIVNAT	Nature de l'activité de l'établissement	2	470	471

<sup>1</sup> La zone ZR3 remplace la variable APET31 qui ne sera plus fournie

NOM	LIBELLE	LONGUEUR	POSITION DEBUT	POSITION FIN
LIEUACT	Lieu de l'activité de l'établissement	2	472	473
ACTISURF	Type de magasin	2	474	475
SAISONAT	Caractère saisonnier ou non de l'activité de l'établissement	2	476	477
MODET	Modalité de l'activité principale de l'établissement	1	478	478
DAPET	Date de validité de l'activité principale de l'établissement	4	479	482
DEFET	Date de mise à jour de l'effectif salarié de l'établissement (année)	4	483	486
EXPLET	Etablissement exploitant tout ou partie des moyens de production	1	487	487
PRODPART	Participation particulière à la production de l'établissement	1	488	488
AUXILT	Caractère auxiliaire de l'activité de l'établissement	1	489	489
EAEANT	Année de validité des rubriques de niveau établissement provenant des EAE*	4	490	493
EAEAPET	Activité principale de l'établissement issue des EAE*	5	494	498
EAESECT1	Première activité secondaire de l'établissement déclarée dans les EAE*	5	499	503
EAESECT2	Deuxième activité secondaire de l'établissement déclarée dans les EAE*	5	504	508
<b>CARACTERISTIQUES ECONOMIQUES DE L'ENTREPRISE</b>				
NOMEN	Nom ou raison sociale de l'entreprise	38	509	546
SIGLE	Sigle de l'entreprise	20	547	566
CIVILITE	Civilité des entrepreneurs individuels	1	567	567
CJ	Catégorie juridique ou professionnelle pour les entrepreneurs individuels	4	568	571
TEFEN	Tranche d'effectif salarié de l'entreprise	2	572	573
EFENCENT	Effectif salarié de l'entreprise à la centaine près	6	574	579
APEN700	Activité principale de l'entreprise en 700 classes	5	580	584
ZR4 <sup>2</sup>	Zone réservée	1	585	585
APRM <sup>3</sup>	Activité principale au registre des métiers	6	586	591
TCA	Tranche de chiffres d'affaires pour les entreprises enquêtées par les EAE*	1	592	592
RECME	Appartenance au Répertoire des Entreprises Contrôlées Majoritairement par l'Etat	1	593	593
DAPEN	Date de validité de l'activité principale de l'entreprise	4	594	597
DEFEN	Date de mise à jour de l'effectif salarié de l'entreprise (année)	4	598	601
DCREN	Date de création de l'entreprise (année, mois)	6	602	607
MMINTREN	Mois d'introduction de l'entreprise dans la base de diffusion	2	608	609
MONOACT	Indice de monoactivité de l'entreprise	1	610	610
MODEN	Modalité de l'activité principale de l'entreprise	1	611	611
EXPLEN	Entreprise exploitant tout ou partie des moyens de production	1	612	612
EAEANN	Année de validité des rubriques de niveau entreprise en provenance des EAE*	4	613	616
EAEAPEN	Activité principale de l'entreprise issue des EAE*	5	617	621
EAESECT1N	Première activité secondaire de l'entreprise déclarée dans les EAE*	5	622	626
EAESECT2N	Deuxième " " " " "	5	627	631
EAESECT3N	Troisième " " " " "	5	632	636
EAESECT4N	Quatrième " " " " "	5	637	641

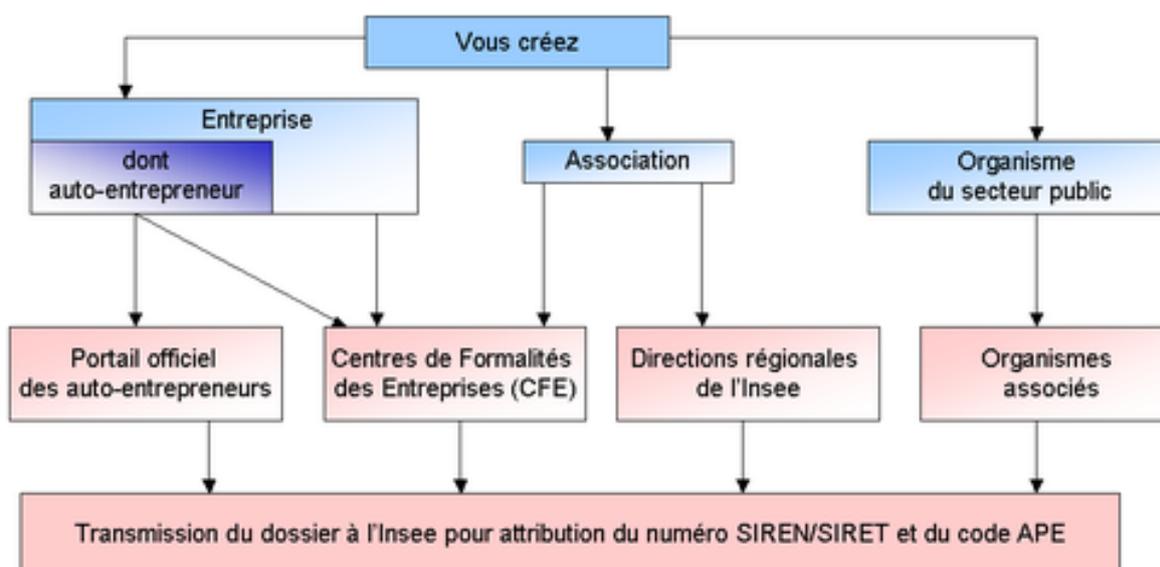
<sup>2</sup> La zone ZR4 remplace en partie la variable APEN31 (1<sup>er</sup> caractère) qui ne sera plus fournie.

<sup>3</sup> La variable APRM passe à 6 caractères.

NOM	LIBELLE	LONGUEUR	POSITION DEBUT	POSITION FIN
NBETEXPL	Nombre d'établissements exploitants actifs de l'entreprise	4	642	645
TCAEXP	Tranche de chiffres d'affaires à l'exportation pour les entreprises enquêtées par les EAE*	1	646	646
REGIMP	Principale région de localisation de l'entreprise	2	647	648
MONOREG	Indice de mono-régionalité de l'entreprise	1	649	649
<b>LOCALISATION DU SIEGE</b>				
RPEN	Région de localisation du siège de l'entreprise	2	650	651
DEPCOMEN	Département et commune de localisation du siège de l'entreprise	5	652	656
<b>DONNEES POUR LES MISES A JOUR</b>				
VMAJ	Nature de la mise à jour (création, suppression, modification)	1	657	657
VMAJ1	Indicateur de mise à jour n°1	1	658	658
VMAJ2	Indicateur de mise à jour n°2	1	659	659
VMAJ3	Indicateur de mise à jour n°3	1	660	660

(\*) EAE = Enquêtes Annuelles Entreprises

Schématisme de la construction de la base SIRENE<sup>®</sup> selon l'Insee. © Insee.



Annexe.7. Exemple de calcul de la longueur d'un connecteur

Données de la population et des emplois						Fonction Tout ou rien					
Population	Cumul_pop	Distance_pop	Emploi	Cumul_emp	Distance_emp	Poids1_pop	Cumul_poids1_pop	Poids1_emp	Cumul_poids1_emp		
0	0	15.3993449	1	1	23.46197227	0	0	23.461972	23.46197227		
6	6	26.9224642	5	6	47.54743449	161.53479	161.5347854	237.73717	261.1991447		
1	7	43.2201625	3	9	52.127618	43.220163	204.7549479	156.38285	417.5819987		
0	7	51.1415375	1	10	72.02604517	0	204.7549479	72.026045	489.6080439		
26	33	54.7226132	1	11	72.02604517	1422.7879	1627.542892	72.026045	561.634089		
2	35	60.6512033	1	12	72.02604517	121.30241	1748.845299	72.026045	633.6601342		
18	53	65.5874683	3	15	72.02604517	1180.5744	2929.419728	216.07814	849.7382697		
0	53	70.5157605	1	16	72.02604517	0	2929.419728	72.026045	921.7643148		
0	53	74.388486	15	31	86.89699121	0	2929.419728	1303.4549	2225.219183		
3	56	74.8579416	2	33	89.48274345	224.57382	3153.993553	178.96549	2404.18467		
2	58	80.7147242	2	35	89.48274345	161.42945	3315.423002	178.96549	2583.150157		
13	71	88.5802981	71	106	89.48274345	1151.5439	4466.966877	6353.2748	8936.424942		

Tampon	Connecteur
800m	489
1000m	546
1200m	577
1400m	617

# BIBLIOGRAPHIE

---

- Agence d'Urbanisme de Lyon, Communauté urbaine de Lyon, SYTRAL, Département du Rhône; Ville de Lyon; Ville de Villeurbanne; Région Rhône-Alpes, 2013. La marche.
- Agence Métropolitaine de Transport, 2012. Autorités organisatrices de transport.
- Baass, 1981. Travel demand forecasting and data considerations. National Academy of Sciences, Washington, D.C.
- Biba, S., Curtin, K.M., Manca, G., 2010. A new method for determining the population with walking access to transit. *Int. J. Geogr. Inf. Sci.* 24, 347–364. doi:10.1080/13658810802646679
- Binetti, M., Ciani, E., 2002. Effects of traffic analysis zones design on transportation models, in: *Proceedings of the 9th Meeting of the Euro Working Group on Transportation*, Bari, Italy.
- Bonnell, P., 2004. *Prévoir la demande de transport*. Presses de l'École nationale des ponts et chaussées, [Paris].
- Bovy, P.H.L., Jansen, G.R.M., 1983. Network Aggregation Effects upon Equilibrium Assignment Outcomes: An Empirical Investigation. *Transp. Sci.* 17, 240.
- Cambridge Systematics, Inc, AECOM Consult, 2007. *A Recommended Approach to Delineating Traffic Analysis Zones in Florida*.
- Carr Smith Corradino, Gannett Fleming, Inc, Florida State University Survey Research Laboratory, Yvonne Ziel Traffic Consultant, Inc., PMG Associates, Inc., Dickey Consulting Services, David Fierro & Associates, 2000. *Southeast Florida Regional Travel Characteristics Study (Executive Summary Report)*.
- Cervero, R., Kockelman, K., 1997. Travel demand and the 3Ds: density, diversity, and design. *Transp. Res. Part Transp. Environ.* 2, 199–219.
- Chang, K., Khatib, Z., Ou, Y., 2002. Effects of zoning structure and network detail on traffic demand modeling. *Environ. Plan. B Plan. Des.* 29, 37–52. doi:10.1068/b2742
- Chapleau, R., 1992. La modélisation de la demande de transport urbain avec une approche totalement désagrégée, in: *World Conference on Transportation Research Proceedings*. pp. 937–948.
- Chu, X., 2004. *Ridership Models at Stop Level*.
- Commissariat général au développement durable, 2009. *CORINE Land Cover France: Clés d'interprétation de la nomenclature*.
- Daniels, R., Mulley, C., 2011. Explaining walking distance to public transport: the dominance of public transport supply. *World* 28, 30.
- Ding, C., 1994. Impact analysis of spatial data aggregation on transportation forecasted demand: a GIS approach. *URISA Proc. URISA 1460*, 362–375.
- Ding, C., 1998. The GIS-based human-interactive TAZ design algorithm: examining the impacts of data aggregation on transportation-planning analysis. *Environ. Plan. B Plan. Des.* 25, 601–616. doi:10.1068/b250601
- El-Geneidy, A., Grimsrud, M., Wasfi, R., Tétreault, P., Surprenant-Legault, J., 2014. New evidence on walking distances to transit stops: identifying redundancies and gaps using variable service areas. *Transportation* 41, 193–210. doi:10.1007/s11116-013-9508-z
- Ewing, R., Cervero, R., 2001. Travel and the built environment: a synthesis. *Transp. Res. Rec. J. Transp. Res. Board* 1780, 87–114.

- Frank, L.D., Pivo, G., 1994. Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transp. Res. Rec.* 44–44.
- Friedrich, M., Galster, M., 2009. Methods for Generating Connectors in Transport Planning Models. *Transp. Res. Rec. J. Transp. Res. Board* 2132, 133–142. doi:10.3141/2132-15
- Godinot, C., 2008. Mise en forme du réseau de transports collectifs sous DaVISUM dans le cadre du projet SIMBAD.
- Guerra, E., Cervero, R., Tischler, D., 2011. The Half-Mile Circle: Does It Represent Transit Station Catchments? *Univ. Calif. Transp. Cent.*
- Gutiérrez, J., Cardozo, O.D., García-Palomares, J.C., 2011. Transit ridership forecasting at station level: an approach based on distance-decay weighted regression. *J. Transp. Geogr.* 19, 1081–1092. doi:10.1016/j.jtrangeo.2011.05.004
- Horowitz, A.J., 2001. Computational issues in increasing spatial precision of traffic assignments. *Transp. Res. Rec. J. Transp. Res. Board* 1777, 68–74.
- Hsiao, S., Lu, J., Sterling, J., Weatherford, M., 1997. Use of geographic information system for analysis of transit pedestrian access. *Transp. Res. Rec. J. Transp. Res. Board* 1604, 50–59.
- IGN, 2008. Descriptif du contenu de la BD TOPO Version 2.
- Insee, 2013. *documentation-complete-donnees-a-200m.pdf*.
- INSEE, 2014. Découpage infracommunal en IRIS. URL <http://www.insee.fr/fr/methodes/default.asp?page=zonages/iris.htm> (accessed 8.25.14).
- Kuby, M., Barranda, A., Upchurch, C., 2004. Factors influencing light-rail station boardings in the United States. *Transp. Res. Part Policy Pract.* 38, 223–247. doi:10.1016/j.tra.2003.10.006
- le secrétariat à l'enquête Origine-Destination, 2003. *Enquête Origine-Destination 2003. La mobilité des personnes dans la région de Montréal.*
- Martinez, L.M., Dupont-Kieffer, A., Viegas, J.M., 2010. An integrated application of zoning for mobility analysis and planning: the case of Paris Region, in: *Proceedings of the 12th WCTR Conference.*
- Martínez, L.M., Viegas, J.M., Silva, E.A., 2009. A traffic analysis zone definition: a new methodology and algorithm. *Transportation* 36, 581–599. doi:10.1007/s11116-009-9214-z
- Martínez, L., Viegas, J., Silva, E., 2007. Zoning Decisions in Transport Planning and Their Impact on the Precision of Results. *Transp. Res. Rec. J. Transp. Res. Board* 1994, 58–65. doi:10.3141/1994-08
- 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, Rapport de recherche du Laboratoire d'Économie des Transports Pour le compte de la DRI (MEEDDAT) et de l'ADEME dans le cadre du groupe.*
- Openshaw, S., 1978. An empirical study of some zone-design criteria. *Environ. Plan. A* 10, 781–794.
- Ortúzar S, Juan D, Willumsen, L.G., 2011. *Modelling transport.* Wiley-Blackwell, Oxford.
- O'Sullivan, S., Morrall, J., 1996. Walking distances to and from light-rail transit stations. *Transp. Res. Rec. J. Transp. Res. Board* 1538, 19–26.

- Parsons Brinckerhoff Quade & Douglas, Inc., 1993. Building orientation a supplement to the pedestrian environment.
- PTV VISUM, 2012. Manuel de l'utilisateur VISUM 12.5.
- Sean Qian, Z., Zhang, H.M., 2012. On centroid connectors in static traffic assignment: Their effects on flow patterns and how to optimize their selections. *Transp. Res. Part B Methodol.* 46, 1489–1503. doi:10.1016/j.trb.2012.07.006
- SYTRAL, Agence d'Urbanisme de Lyon, 2011. Fiche technique conférence des maires Lyon-Villeurbanne.
- Szu-Han Chen, 2013. Ridership\_and\_built\_environment.pdf.
- WATS, 2008. Zone centroid and centroid connector placement.
- Zhao, F., Chow, L.-F., Li, M.-T., Ubaka, I., Gan, A., 2003. Forecasting transit walk accessibility: regression model alternative to buffer method. *Transp. Res. Rec. J. Transp. Res. Board* 1835, 34–41.

# LISTE DES FIGURES

---

FIGURE 1 : EXEMPLE DE DÉFINITION DE CONNECTEURS VP ET TC SUR LA BASE DE LA POSITION GÉOGRAPHIQUE DU CENTROÏDE DE ZONE (SOURCE SIMBAD) .....	8
FIGURE 2 : PÉRIMÈTRE SIMBAD (FOND DE CARTE OPENSTREETMAP).....	11
FIGURE 3 : AGENCEMENT DES MODULES INTÉGRÉS À SIMBAD (NICOLAS ET AL., 2009).....	12
FIGURE 4 : DÉCOUPAGE EN IRIS DE L'AIRES SIMBAD, FOCUS SUR LA COMMUNE DE LYON .....	13
FIGURE 5 : CAS DE ZONES ISSUES DE SIMBAD NON CONFORMES AUX RÉGLES DE DÉCOUPAGE	14
FIGURE 6 : EXEMPLE DE DEUX CONFIGURATIONS DE CODIFICATION DES CENTROÏDES .....	16
FIGURE 7 : EXEMPLE DE ZONES À CALUIRE-ET-CUIRE AVEC UNE DENSITÉ NON UNIFORME (ORHTOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2009).....	17
FIGURE 8 : EXEMPLE D'UNE ZONE AVEC UN TISSU URBAIN DISCONTINU (ORHTOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012).....	18
FIGURE 9 : NATURE DU TISSU URBAIN SUR L'AIRES SIMBAD (DONNÉES CORINE LAND COVER 2006) .....	19
FIGURE 10 : EXEMPLE DE CONNECTEURS DE L'IRIS 693860101 ENJAMBANT LE RHÔNE (ORHTOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012).....	20
FIGURE 11 : EXEMPLE D'UNE SITUATION DE RÉFÉRENCE D'UNE CONFIGURATION DE CONNECTEURS.....	21
FIGURE 12 : LES QUATRE CONFIGURATIONS TESTÉES.....	22
FIGURE 13 : EXEMPLE D'IRIS ET DE DISPOSITIONS GÉOGRAPHIQUES DES ARRÊTS DES LIGNES TC (ORHTOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012).....	26
FIGURE 14 : EXEMPLE D'ACCESSIBILITÉ DES IRIS .....	27
FIGURE 15 : MUTUALISATION D'UN POINT DE DESSERTE ENTRE LIGNES (BUS 1ET-34) ET DIRECTIONS (ALLER-RETOUR) DE BUS (FOND GOOGLE MAPS).....	29
FIGURE 16 : EXEMPLE DE PARTITION EN CARROYAGE RÉSUMANT L'INFORMATION DÉSAGRÉGÉE.....	32
FIGURE 17 : DIFFÉRENCE ENTRE AIRES DE CHALANDISE POUR UNE MÊME STATION SELON LA MÉTHODOLOGIE DE CALCUL DE LA DISTANCE (SZU-HAN CHEN, 2013) .....	38
FIGURE 18 : DONNÉES D'ENTRÉE ET DE SORTIE ET LEUR INTERACTION .....	45
FIGURE 19 : EXTRAIT DE LA BASE CARROYÉE CORRESPONDANT À L'AIRES SIMBAD.....	46
FIGURE 20 : EXTRAIT DE LA BASE CARROYÉE AU NIVEAU DE LA COMMUNE DE LYON .....	47
FIGURE 21 : EXTRAIT DE LA BD TOPO <sup>®</sup> CORRESPONDANT À L'AIRES SIMBAD.....	48
FIGURE 22 : EXTRAIT DE LA BD TOPO <sup>®</sup> AU NIVEAU DE LYON .....	48
FIGURE 23 : EXEMPLE DE BÂTIMENTS ISSUS DE LA BD TOPO <sup>®</sup> (FOND GOOGLEMAPS) .....	49
FIGURE 24 : EXEMPLES DE BÂTIMENTS NON RÉSIDENTIELS INCLUS DANS LA CLASSE BÂTI_INDIFÉRENCIÉ DE LA BD TOPO <sup>®</sup> (FOND GOOGLEMAPS).....	50
FIGURE 25 : RÉSULTAT D'IDENTIFICATION DES BÂTIMENTS NON RÉSIDENTIELS (FOND GOOGLEMAPS).....	51
FIGURE 26 : EXEMPLE DE CONSTRUCTIONS CODÉES AVEC UNE HAUTEUR NULLE DANS LA BD TOPO <sup>®</sup> (FOND BINGMAP) .....	52
FIGURE 27 : LOCALISATION DES EMPLOIS SUR L'AIRES SIMBAD SELON LA BASE SIRENE TRAITÉE .....	53

FIGURE 28 : EXEMPLE D'IMPRÉCISION DE LA GÉOLOCALISATION SIRENE AU NIVEAU DE LA PART-DIEU (ORHTOPHOTOGRAPHIE DEPUIS LA BASE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012).....	54
FIGURE 29 : CENTROÏDES DES CARREAUX SUR L'AIRESIMBAD.....	56
FIGURE 30 : CARROYAGE ET CENTROÏDES AU NIVEAU DE LA COMMUNE DE LYON.....	56
FIGURE 31 : PROCESSUS D'ÉLABORATION DE LA MÉTHODE 1.....	57
FIGURE 32 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA MÉTHODE 1.....	57
FIGURE 33 : ERREUR RELATIVE AU NIVEAU DE LA COMMUNE DE LYON PAR LA MÉTHODE 1 ..	59
FIGURE 34 : PROCESSUS D'ÉLABORATION DE LA MÉTHODE 2.....	61
FIGURE 35 : EXEMPLE D'ENTITÉS COLORIÉES EN FONCTION DE L'IRIS D'APPARTENANCE.....	61
FIGURE 36 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA MÉTHODE 2.....	62
FIGURE 37 : ERREUR RELATIVE AU NIVEAU DE LA COMMUNE DE LYON PAR LA MÉTHODE 2 ..	63
FIGURE 38 : EXEMPLE DE RÉPARTITION DE LA POPULATION AU NIVEAU DE L'IRIS DU PARC (FOND BINGMAP).....	65
FIGURE 39: PROCESSUS D'ÉLABORATION DE LA MÉTHODE 3.....	67
FIGURE 40 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA METHODE 3.1.....	68
FIGURE 41 : ERREUR RELATIVE AU NIVEAU DE LYON PAR LA METHODE 3.1.....	69
FIGURE 42 : EXEMPLE DE RÉPARTITION DU BÂTI EN COULEUR SUR LES DIFFÉRENTES ENTITÉS.....	71
FIGURE 43 : ERREUR RELATIVE ENTRE POPULATIONS ESTIMÉE ET RECENSÉE PAR LA METHODE 3.2.....	72
FIGURE 44 : ERREUR RELATIVE AU NIVEAU DE LYON PAR LA METHODE 3.2.....	73
FIGURE 45 : EXEMPLE D'ESTIMATION DE LA POPULATION DES BÂTIMENTS (FOND GOOGLMAPS).....	78
FIGURE 46 : CAS DE L'IRIS NUMÈRE 13330000 DE SAINT-ANDRÉ-DE-CORCY (FOND DE CARTE OPENSTREETMAP).....	80
FIGURE 47 : APPLICATION DE LA FONCTION 'TOUT OU RIEN' AVEC UNE DISTANEC DE 800M ..	81
FIGURE 48 : CAS DE L'IRIS 10430000 DE BEYNOST (FOND DE CARTE OPENSTREET MAP).....	84
FIGURE 49 : RÉPARTITION DES ÉLÉMENTS DE L'IRIS 5 SUR 2 STATIONS (ORHTOPHOTOGRAPHIE ISSUE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012).....	85
FIGURE 50 : CAS DE L'IRIS 693860303 DE VITTON (FOND DE CARTE OPENSTREETMAP) .....	88
FIGURE 51 : CARTE DE SITUATION DE L'IRIS 693860303 ET DES LIGNES ET ARRÊTS À PROXIMITÉ (FOND DE CARTE OPENSTREETMAP).....	89
FIGURE 52 : DÉTERMINATION DES ARRÊTS ACCESSIBLES (ORHTOPHOTOGRAPHIE ISSUE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012) ..	90
FIGURE 53 : CONNECTEURS GÉNÉRÉS PAR SIMBAD DANS LE CAS DE L'IRIS VITTON (ORHTOPHOTOGRAPHIE ISSUE DE LA PLATEFORME D'INTÉGRATION DE DONNÉES GÉOGRAPHIQUES DU GRAND LYON, 2012).....	91

# LISTE DES GRAPHIQUES

---

GRAPHIQUE 1 : PRMSE DE TOUTES LES LIGNES DE TRANSPORT EN COMMUN .....	23
GRAPHIQUE 2 : RMSE DES QUATRE EXPÉRIENCES SELON LA CATÉGORIE DE LA LIGNE.....	23
GRAPHIQUE 2 : PRMSE DES QUATRE EXPERIENCES SELON LA CATEGORIE DE LA LIGNE.....	24
GRAPHIQUE 3 : DISTANCE D'ACCÈS AUX ARRÊTS EN FONCTION DU MODE (SOURCE AGENCE D'URBANISME DE LYON ET AL., 2013) .....	27
GRAPHIQUE 4 : RÉPARTITION DES DÉPACEMENTS DES RÉSIDENTS DE LYON-VILLEURBANNE SELON LE MOTIF (SYTRAL AND AGENCE D'URBANISME DE LYON, 2011) .....	34
GRAPHIQUE 5 : PART MODALE SELON LE MOTIF DE DÉPLACEMENT SUR LYON-VILLEURBANNE .....	35
GRAPHIQUE 6 : EFFECTIFS DES USAGERS ACCEDANT AUX STATIONS EN FONCTION DE LA DISTANCE DE MARCHÉ À PIED (ZHAO ET AL., 2003) .....	36
GRAPHIQUE 7 : FONCTION DÉGRESSIVE AGRÉGÉE SELON ZHAO (ZHAO ET AL., 2003) .....	39
GRAPHIQUE 8 : FONCTION DÉGRESSIVE DÉSAGRÉGÉE SELON GUTIÉRREZ (GUTIÉRREZ ET AL., 2011) .....	39
GRAPHIQUE 9 : FRÉQUENCE CUMULÉE DU NOMBRE DE PASSAGERS EN FONCTION DE LA DISTANCE DE MARCHÉ À PIED (EL-GENEIDY ET AL., 2014) .....	42
GRAPHIQUE 10 : DISTRIBUTION DES ENTREPRISES SELON LA TAILLE.....	54
GRAPHIQUE 11 : COURBE DE FRÉQUENCE DE L'ERREUR RELATIVE SUR L'AIRESIMBAD.....	58
GRAPHIQUE 12 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD .....	58
GRAPHIQUE 13 : COURBE DE FRÉQUENCE DE L'ERREUR RELATIVE SUR LYON.....	59
GRAPHIQUE 14 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON .....	60
GRAPHIQUE 15 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR L'AIRESIMBAD.....	62
GRAPHIQUE 16 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD .....	63
GRAPHIQUE 17 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR LYON .....	64
GRAPHIQUE 18 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON .....	64
GRAPHIQUE 19 : COURBE DE FRÉQUENCE DE L'ERREUR RELATIVE SUR L'AIRESIMBAD.....	68
GRAPHIQUE 20 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD .....	69
GRAPHIQUE 21 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR LYON.....	70
GRAPHIQUE 22 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON .....	70
GRAPHIQUE 23 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR L'AIRESIMBAD.....	72
GRAPHIQUE 24 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD .....	73
GRAPHIQUE 25 : COURBE DE FREQUENCE DE L'ERREUR RELATIVE SUR LYON.....	74
GRAPHIQUE 26 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON .....	74
GRAPHIQUE 27 : COMPARAISON DES ERREURS REATIVES GÉNÉRÉES PAR LES 4 MÉTHODES SUR L'AIRESIMBAD.....	75
GRAPHIQUE 28 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR L'AIRESIMBAD .....	75

GRAPHIQUE 29 : COMPARAISON DES ERREURS REATIVES GENERÉES PAR LES 4 MÉTHODES SUR LYON .....	76
GRAPHIQUE 30 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR L'ERREUR RELATIVE SUR LYON .....	76
GRAPHIQUE 31 : SOMME DES ÉCARTS ENTRE POPULATIONS ESTIMÉE ET RECENSÉE SELON L'AIRE D'ÉTUDE.....	77
GRAPHIQUE 32 : POURCENTAGE DE LA POPULATION RECENSÉE CONCERNÉE PAR UNE ERREUR RELATIVE DE -20% À 20%.....	77

# LISTE DES TABLES

---

TABLE 1 : MÉTHODOLOGIE DE PONDÉRATION DU CALCUL DE LA DISTANCE MOYENNE À L'ARRÊT.....	37
TABLE 2 : EXEMPLE DE DISTANCES MOYENNES DE RABATTEMENT À PIED DANS LE GRAND LYON (AGENCE D'URBANISME DE LYON ET AL., 2013).....	38
TABLE 3 : QUELQUES FONCTIONS DÉGRESSIVES (ZHAO ET AL., 2003).....	39
TABLE 4 : QUELQUES FORMES DE LA FONCTION COMBINÉE (GUTIÉRREZ ET AL., 2011; ZHAO ET AL., 2003).....	40
TABLE 5 : ÉTENDUE DE L'AIRES DE SERVICE DE CERTAINS MODES DE TRANSPORT EN COMMUN AU CANADA ET AUX ÉTATS-UNIS (O'SULLIVAN AND MORRALL, 1996) .....	41
TABLE 6 : STATISTIQUES DE LA DISTANCE DE MARCHÉ À PIED VERS ET DEPUIS LES ARRÊTS (EL-GENEIDY ET AL., 2014) .....	42
TABLE 7 : RAYON DES AIRES DE SERVICE EN FONCTION DU MODE DE TRANSPORT.....	43
TABLE 8 : ATTRIBUTS CONTENUS DANS LA COUCHE BÂTI INDIFFÉRENCIÉ DE LA BD TOPO <sup>®</sup> (VOIR ANNEXE.5).....	49
TABLE 9 : LONGUEUR DU CONNECTEUR DE L'IRIS 46 EN FONCTION DE L'AIRES DE SERVICE CHOISIE (VOIR ANNEXE.7 POUR LE CALCUL).....	81
TABLE 10 : LONGUEUR DU CONNECTEUR EN FONCTION DE LA DISTANCE TAMPON CHOISIE POUR LA FONCTION COMBINÉE .....	83
TABLE 11 : LONGUEUR DU CONNECTEUR SELON LA DISTANCE TAMPON CHOISIE.....	86
TABLE 12 : LONGUEUR DU CONNECTEUR PAR LA FONCTION COMBINÉE AGRÉGÉE SELON LA DISTANCE TAMPON CHOISIE.....	87
TABLE 13 : LONGUEUR DU CONNECTEUR PAR LA FONCTION COMBINÉE DÉAGRÉGÉE SELON LA DISTANCE TAMPON CHOISIE.....	87
TABLE 14 : EXEMPLE D'AIRES DE SERVICE APPLIQUÉES À L'IRIS VITTON .....	88
TABLE 15 : LONGUEUR DE CONNECTEURS POUR DIFFÉRENTES LIGNES.....	90