

**Titre:** Modèles de génération et de distribution de déplacements en vélopartage  
Title: vélopartage

**Auteur:** Mikael Chibok  
Author:

**Date:** 2018

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Chibok, M. (2018). Modèles de génération et de distribution de déplacements en vélopartage [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie.  
Citation: <https://publications.polymtl.ca/3296/>

## Document en libre accès dans PolyPublie

Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/3296/>  
PolyPublie URL:

**Directeurs de recherche:** Catherine Morency, & Martin Trépanier  
Advisors:

**Programme:** Génie civil  
Program:

UNIVERSITÉ DE MONTRÉAL

MODÈLES DE GÉNÉRATION ET DE DISTRIBUTION  
DE DÉPLACEMENTS EN VÉLOPARTAGE

MIKAËL CHIBOK

DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES  
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION  
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES  
(GÉNIE CIVIL)

AOÛT 2018

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

MODÈLES DE GÉNÉRATION ET DE DISTRIBUTION  
DE DÉPLACEMENTS EN VÉLOPARTAGE

présenté par : CHIBOK Mikaël

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

Mme MILLETTE Louise, Ph. D., présidente

Mme MORENCY Catherine, Ph. D., membre et directrice de recherche

M. TRÉPANIER Martin, Ph. D., membre et codirecteur de recherche

M. BRICE Stéphane, Ph. D., membre

## DÉDICACE

*J'ai faim plus que Alain. J'ai faim plus que Alain,  
parce que lui, on sait, y'a grignoté, toute la  
journée. J'ai faim. Plus que Alain. J'veux des  
craquelins, j'veux des raisins, un morceau de pain.*

*J'croque, croque, croque des concombres, croque,  
croque, croque des concombres, croque, croque,  
croque, croque des concombres.*

*J'croque des concombres !*

*Pourquoi Alain, est-ce que t'as mangé l'intérieur  
de mon frigidaire ? Pourquoi Alain, est-ce que t'as  
mangé l'intérieur de mon frigidaire ?*

*J'ai frette plus que Ginette. J'ai frette plus que  
Ginette. J'ai frette plus que Ginette, parce qu'elle,  
on sait, a s'est tricoté, un gros chandail. Un gros  
chandail.*

*J'ai frette plus que Ginette. J'ai frette plus que  
Ginette. Ouvre ma chaufferette, dans mes  
couvertes. Jlis des magazines de filles.*

*Des magazines de filles !*

*J'adore la mode. (x25)*

*À toutes les femmes que j'ai aimées...*

Les Appendices

## REMERCIEMENTS

Merci à des gens vraiment sérieux : Catherine, Martin, Paula.

Merci à des gens sérieux de la Chaire mobilité, du B-330 et de l'« autre bureau » : Jean-Simon, Pierre-Léo, Gabriel-Lee, madame punaises, prof. Deschaintres, Laurent-David, Boubou, Stephen, Nico, Jérôme, Charles.

Merci aussi à des outils : Google, Stack Overflow et autres, dont les tonnes de sites dénichés au fond du web, de présentations PowerPoint et de fichiers Excel mis en ligne par des profs pour le fun, qui te font vraiment comprendre des choses.

Merci à des gens moins sérieux, mais tout aussi pertinents et surtout passionnés : Emom, Frédé, Anamaïs, Didi.

Merci à d'autres qui m'ont accompagné durant ce périple : Max, Marc (malgré tout), Christine, Clouth, Sass, Do, Black, eXavier, Fred, Manni et les autres.

Merci à des choses vraiment pas sérieuses : *SSB ♥, Le D. Ross, les Steady, Astro-Diane, A.W. !* et j'en passe.

Merci à la famille, à Robert et Astrid, même s'ils n'ont pas vraiment le choix.

Merci à Catherine, la deuxième (dans cette liste), mais non la moindre. Un vrai pilier dans toute cette histoire. On passe peut-être finalement à autre chose ! Jusqu'à ce que je fasse un doc en tout cas.

Merci à moi de toujours me lancer dans des affaires qui se peuvent pas.

## RÉSUMÉ

Le vélopartage est un mode de déplacement qui connaît une popularité soutenue depuis la deuxième moitié de la dernière décennie. Les systèmes de ce genre sont aujourd’hui dotés de technologies de l’information permettant les enregistrements des nombreux emprunts et retours de bicyclettes qui y sont effectués. Ces bases de données grandissantes ont permis une étude de la mobilité urbaine qui a elle-même parallèlement évolué. La Ville de Montréal s’est munie de son propre système de vélos en libre-service nommé BIXI en 2009. Une des particularités de ce dernier sont les stations modulables, retirées à chaque fin d’année et redéployées la saison suivante. Ceci signifie donc que le réseau montréalais peut être optimisé selon l’utilisation et les changements qui se produisent sur le territoire d’année en année. Ceci relève d’une planification du niveau stratégique. En ce sens, l’objectif de ce projet de recherche est d’améliorer les connaissances quant à la grande variation qui existe dans l’usage du système BIXI. Il est également question de mettre à l’épreuve des méthodes de modélisation classiques à ce type d’infrastructure de déplacements contemporain.

D’abord, une mise en contexte permet au lecteur de placer le vélopartage dans le paysage des modes urbains. Pour ce faire, un survol historique des diverses générations qui ont ponctué l’évolution du vélopartage situe les systèmes actuellement utilisés et les possibles pistes de développement des systèmes futurs. Les composantes clés d’un système typique sont également présentées, soit les stations, la capacité de celles-ci, les types d’usagers, la notion de déplacements qui est utilisée dans le mémoire ainsi que les saisons BIXI. Ces éléments permettent de présenter le fonctionnement usuel du réseau. Ensuite, une revue de littérature scientifique révèle de nombreuses variables possédant un effet sur la demande de déplacements en vélopartage, que ce soit à l’Origine ou à la Destination. Effectivement, l’utilisation d’un système de vélos en libre-service semble varier selon le moment dans la journée, dans la semaine et dans l’année, tout comme dans l’espace. Ces diverses variables spatiales sont reprises dans l’étape de modélisation de la génération des déplacements. Cet état de l’art présente aussi quelques études qui se sont penchées sur les déplacements en tant que tels, soit la distance et la durée qui unit chaque Origine et Destination. Cette vue d’ensemble démontre que les études qui sont liées aux modèles de distribution sont peu nombreuses et ce mémoire apporte une contribution en ce sens. La contextualisation se termine avec une présentation de la procédure séquentielle classique puisque les deux premières étapes sont utilisées dans ce document dans le but de modéliser les déplacements en BIXI.

La première partie des analyses se concentre sur les flux aux stations en 2011, soit sur le plan de la production (sortants) ou de l'attraction (entrants) de déplacements. En ce sens, ces flux varient grandement selon la temporalité d'utilisation, ce qui concorde avec la revue de littérature. Un test statistique permet de déterminer que les courbes de distribution des déplacements observés lors des jours ouvrables pendant les mois de juin et de juillet se ressemblent davantage. Dans le but de minimiser la variabilité, les déplacements effectués à l'extérieur de cette période (soit les fins de semaine et les autres mois) sont exclus de la base de données principale. De plus, une brève analyse spatiale réalisée avec une méthode de classement des données démontre une variation selon l'emplacement des stations. Ceci permet d'introduire des variables spatiales dans la modélisation de la génération des déplacements, réalisée par une régression linéaire multiple. Les variables indépendantes utilisées dans ce modèle, retenues selon la revue de littérature, se classent dans les quatre catégories suivantes : sociodémographique, usage du sol et environnement bâti, offre de transport et offre du réseau BIXI. Les variables spatiales sont calculées pour chaque station de vélopartage à l'aide d'une zone tampon de 400 m. Les éléments qui ont un impact significatif sur les flux entrants et sortants sont cohérents avec la littérature, soit certaines variables telles que la densité de population, l'âge des usagers, les emplois, les infrastructures cyclables, l'élévation des stations, ou encore la capacité de ces dernières. Cependant, la validation de ces modèles sur des déplacements observés lors d'une année subséquente s'avère moins certaine puisque l'impact de plusieurs variables est altéré. Des études approfondies davantage sur le sujet seront nécessaires.

En second lieu, la distribution spatiale des déplacements, soit les lignes de désir, est analysée. Pour ce faire, les déplacements de deux années sont utilisés, soit 2011 et 2014, et seulement ceux effectués par des membres lors de la pointe matinale (6 à 9 h) des jours ouvrables sont retenus. Alors que près du tiers des paires Origine-Destination étaient actives en 2011, cette proportion tombe au cinquième en 2014, ce qui s'explique probablement par l'ajout de stations. La spatialisation de la matrice Origine-Destination permet de déceler certaines tendances, notamment en ce qui a trait à la Destination de certains déplacements vers des stations de métro, ce qui suppose une combinaison de modes. Ensuite, les lignes de désir sont modélisées grâce au modèle gravitaire, qui se base sur une fonction de coût afin de déterminer un niveau d'attractivité pour chaque paire Origine-Destination. Quatre métriques sont utilisées dans le cas présent, soit la durée et la distance réseau minimum avec ou sans dénivelé. Des modèles sont donc calibrés et validés sur des déplacements observés lors d'une autre année. Les résultats sont prometteurs, mais certaines lignes

de désir sont plus complexes à modéliser en raison d'une relation spéciale, particulièrement lorsque la station de Destination possède une grande capacité et est située près du centre-ville. Ainsi, une analyse spatiale exploratoire est effectuée afin de déceler des éléments potentiellement pertinents à incorporer dans une mouture bonifiée du modèle gravitaire.

Ce mémoire permet donc d'améliorer la compréhension de certaines dynamiques inhérentes au réseau BIXI et à l'utilisation qui en est faite par les usagers. La méthodologie qui est proposée constitue un jalon vers une modélisation intégrée de la génération et de la distribution des déplacements, ce qui permettrait éventuellement une planification stratégique selon diverses configurations des stations dans le réseau.

## ABSTRACT

Bikesharing is a transport mode that has seen a continuous popularity since the second half of the last decade. It now integrates information technologies which can record every location and return of bikes that are made. Studies on bikesharing grew in parallel, in part with the help of these trip databases. The City of Montreal has launched its own bikesharing system named BIXI in 2009. One of its features is the modular aspect of the stations, which are removed before each winter and reinstalled in the spring. This means that the Montreal network can be optimized based on usage and territorial changes that occur year after year. This is in relation with strategic planning. In this sense, this research project objective is to deepen the understandings of the usage variability that exists in the BIXI network. Furthermore, this document puts forward the application of classical modeling methods on this contemporary transport mode.

Firstly, bikesharing is put in context regarding other urban transportation modes. This is done with a historical presentation of the various generations that characterize the evolution of bikesharing systems. This puts in perspective the systems used nowadays as part of a continuous upgrade and shows certain trends for future developments. Key aspects of a usual bikesharing system are also presented, which are stations and their capacity, the type of users, the type of trips that are referred to in this document, and lastly the BIXI seasons. These components are used to explain how the network functions. Afterwards, a review of the scientific literature reveals that numerous elements have an impact on the bikesharing trips demand, which is either at the Origin or Destination. In fact, usage can vary depending on the time of day, the day of the week, or the period during the year, and on the location of the station. The literature review puts forward that few studies consider trip distribution, based on the distance or the duration between each Origin and Destination. This thesis contributes to the few studies published on bikesharing trip distribution modeling. The contextualization ends by a presentation of the four-step model of which the first two steps are used in this research to model BIXI trips.

Secondly, an analysis is done on flows at stations in 2011, which considers trip production (out) and attraction (in). As discovered in the literature review, these flows vary greatly depending on the time. A statistical test is used and shows that the trip distribution curves observed on business days during the months of June and July are similar. There, to limit the variability of usage, trips done outside this period (weekends and other months) are excluded from the main database. Also,

a simple clustering of out flows shows different usage based on the location of the station. A multiple linear regression to model trip generation is then used based on 15 spatial variables. These, inspired from the literature review, can be grouped in the following four categories: sociodemographic, land use and built environment, transport supply and BIXI network supply. The spatial variables are calculated for each station within a 400 m buffer zone. Factors that have a significant impact on in and out flows are consistent with the literature, such as population density, age of users, jobs, cycling infrastructure, elevation of stations and their capacity. However, the validation of these models on trips of another year seems uncertain because the impact of many variables is altered. More research on this aspect is necessary.

Thirdly, trip distribution, also known as desire lines, is analyzed. Trips observed on two years (2011 and 2014) are used, and only those done by members in the morning peak hours (6 to 9) on business days are kept. Almost a third of the Origin-Destination pairs were active in 2011, while this proportion falls to a fifth in 2014, which is possibly explained by the number of stations added to the network. Spatialization of certain desire lines reveal trends. For example, some trip Destinations are concentrated around subway stations, which hints towards a combination of bikesharing and public transit. Afterwards, desire lines are modeled with the help of the gravity model, which is based on a cost function to determine a certain attractiveness level for every Origin-Destination pair. Four metrics are used in the present case, which is the duration and the minimum network distance with and without accounting for altitude. Models are calibrated and validated on trips observed on another year. Results are promising but certain desire lines are more complex to model because of a special relation between stations, particularly when a Destination has a large capacity and is near the downtown area. Because of this, an exploratory spatial analysis is done to show which variables could be used in an improved version of the gravity model.

This master's thesis improves the understandings of certain inherent dynamics in the BIXI network, and the usage done by customers. The proposed methodology can be considered as a marker towards an integrated trip generation and distribution modeling approach, which could eventually lead to strategic planning based on various stations configurations in the network.

## TABLE DES MATIÈRES

|  |       |
|--|-------|
| DÉDICACE.....  | III   |
| REMERCIEMENTS .....  | IV    |
| RÉSUMÉ.....  | V     |
| ABSTRACT .....   | VIII  |
| TABLE DES MATIÈRES .....   | X     |
| LISTE DES TABLEAUX.....  | XIII  |
| LISTE DES FIGURES.....   | XV    |
| LISTE DES SIGLES ET ABRÉVIATIONS .....                               | XVII  |
| LISTE DES ANNEXES.....   | XVIII |
| CHAPITRE 1 INTRODUCTION.....   | 1     |
| 1.1 Problématique.....   | 2     |
| 1.2 Objectifs .....  | 3     |
| 1.3 Structure du document .....                                      | 4     |
| CHAPITRE 2 MISE EN CONTEXTE.....                                     | 5     |
| 2.1 Le vélopartage dans la mobilité urbaine .....                    | 5     |
| 2.2 Histoire du vélopartage .....                                    | 6     |
| 2.3 Particularités montréalaises.....                                | 8     |
| 2.3.1 Composantes et fonctionnement .....                            | 8     |
| 2.3.2 Tarification usuelle.....                                      | 11    |
| 2.3.3 Répartition spatiale des stations en 2011 .....                | 12    |
| 2.4 Études sur le vélopartage.....                                   | 13    |
| 2.4.1 La demande de déplacements en vélopartage .....                | 13    |
| 2.4.2 La distribution spatiale des déplacements en vélopartage ..... | 16    |

|  |   |           |
|--|---|-----------|
| 2.5  | Méthodologie générale : la procédure séquentielle classique ..... | 16        |
| 2.6  | Conclusion du chapitre.....                                       | 23        |
| <b>CHAPITRE 3 MODÉLISATION DE LA GÉNÉRATION DES DÉPLACEMENTS .....</b>   |   | <b>25</b> |
| 3.1  | Analyse et triage des données .....                               | 25        |
| 3.1.1  | Analyse temporelle.....   | 27        |
| 3.1.2  | Analyse spatiale.....   | 33        |
| 3.2  | Modélisation de la génération des déplacements .....              | 34        |
| 3.2.1  | La régression linéaire .....                                      | 34        |
| 3.2.2  | Développement des modèles .....                                   | 42        |
| 3.2.3  | Résultats de la modélisation .....                                | 46        |
| 3.2.4  | Validation des modèles .....                                      | 49        |
| 3.2.5  | Limites et perspectives .....                                     | 51        |
| 3.3  | Conclusion du chapitre.....                                       | 52        |
| <b>CHAPITRE 4 MODÉLISATION DE LA DISTRIBUTION DES DÉPLACEMENTS .....</b> |   | <b>54</b> |
| 4.1  | Introduction aux concepts .....                                   | 54        |
| 4.2  | Triage de données.....  | 56        |
| 4.3  | Analyse de la distribution des déplacements observés.....         | 57        |
| 4.4  | Modélisation de la distribution des déplacements.....             | 68        |
| 4.4.1  | Le modèle gravitaire.....   | 68        |
| 4.4.2  | Problématique de la diagonale .....                               | 71        |
| 4.4.3  | Développement des modèles .....                                   | 72        |
| 4.4.4  | Résultats de la modélisation .....                                | 72        |
| 4.4.5  | Validation des modèles .....                                      | 85        |
| 4.4.6  | Limites et perspectives .....                                     | 89        |

|   |            |
|---|------------|
| 4.5 Conclusion du chapitre.....                       | 90         |
| <b>CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS .....</b> | <b>92</b>  |
| 5.1 Synthèse de la recherche et contributions .....   | 92         |
| 5.2 Limites.....                                      | 94         |
| 5.3 Perspectives.....                                 | 95         |
| <b>BIBLIOGRAPHIE .....</b>                            | <b>97</b>  |
| <b>ANNEXES .....</b>                                  | <b>101</b> |

## LISTE DES TABLEAUX

|   |    |
|---|----|
| Tableau 2.1 : Tarification du système BIXI en 2018 .....  | 11 |
| Tableau 2.2 : Exemple de matrice Origine-Destination remplie.....   | 19 |
| Tableau 2.3 : Génération de la matrice Origine-Destination .....  | 20 |
| Tableau 2.4 : Distribution de la matrice Origine-Destination .....  | 22 |
| Tableau 3.1 : Statistiques sommaires des flux selon le mois .....   | 27 |
| Tableau 3.2 : Résultats du test de la somme des rangs Wilcoxon des flux sortants selon le mois  | 29 |
| Tableau 3.3 : Statistiques sommaires des flux selon le jour de la semaine (juin et juillet).....  | 30 |
| Tableau 3.4 : Résultats du test de la somme des rangs Wilcoxon des flux sortants selon le jour..  | 32 |
| Tableau 3.5 : Description des variables indépendantes modélisées (juin et juillet 2011).....  | 38 |
| Tableau 3.6 : Coefficients de détermination des diverses transformations de variables.....  | 44 |
| Tableau 3.7 : Résultats de la modélisation 2011 avec transformation de type puissance sans variables corrélées .....  | 46 |
| Tableau 3.8 : Résultats de la modélisation 2011 avec transformation de type puissance .....   | 47 |
| Tableau 3.9 : Résultats de la modélisation 2016 avec transformation de type puissance .....   | 50 |
| Tableau 4.1 : Sommaire des déplacements des membres en période de pointe matinale .....   | 58 |
| Il est également possible de remarquer une hausse de la distance moyenne des déplacements, soit de près de 70 mètres. Ceci pourrait être une résultante de l'augmentation du nombre de stations dans le réseau, plus éparpillées dans le territoire, et donc des déplacements plus longs.Tableau 4.2 Données moyennes des déplacements des membres selon quatre métriques ..... | 59 |
| Tableau 4.3 : Matrice Origine-Destination des membres en 2011 .....   | 63 |
| Tableau 4.4 : Statistiques sommaires des résidus .....  | 77 |
| Tableau 4.5 : Distribution des déplacements en pourcentage selon le type de paires O-D.....   | 81 |
| Tableau 4.6 : Statistiques sommaires des résidus absolus (excluant 0) selon la capacité .....   | 82 |
| Tableau 4.7 : Coefficients de régression linéaire des résidus .....   | 84 |

|   |    |
|---|----|
| Tableau 4.8 : Résultats des modèles selon les indicateurs ..... | 87 |
| Tableau 4.9 : Pourcentage de résidus selon les modèles.....     | 88 |

## LISTE DES FIGURES

|  |    |
|--|----|
| Figure 1.1 : Schéma de la méthodologie générale du projet .....  | 4  |
| Figure 2.1 : La place du vélopartage dans l'offre de transport urbain. ....  | 6  |
| Figure 2.2 : Station BIXI, comprenant borne d'accès et points d'ancrage. ....  | 9  |
| Figure 2.3 : Point d'ancrage avec clavier et fente pour clé. ....  | 10 |
| Figure 2.4 : Spatialisation du réseau 2011 et distance au centre-ville .....   | 12 |
| Figure 2.5 : Interactions prépondérantes entre les trois sous-systèmes.....  | 17 |
| Figure 2.6 : Schéma de la procédure séquentielle classique. ....   | 18 |
| Figure 3.1 : Distribution fréquentielle des flux sortants quotidiens selon le mois .....   | 28 |
| Figure 3.2 : Distribution fréquentielle des flux quotidiens selon le jour de la semaine (juin et juillet)<br>.....                   | 31 |
| Figure 3.3 : Ellipses de dispersion selon les groupes de stations classées par taux sortant moyen                                    | 33 |
| Figure 3.4 : Relation entre le taux sortant quotidien moyen par station et les variables<br>indépendantes .....                      | 39 |
| Figure 3.5 : Relation entre le taux entrant quotidien moyen par station et les variables<br>indépendantes .....                      | 40 |
| Figure 3.6 : Matrice de corrélation entre les variables indépendantes .....  | 41 |
| Figure 3.7 : Spatialisation des résidus du modèle des flux sortants quotidiens moyens .....  | 43 |
| Figure 3.8 : Spatialisation des résidus du modèle des flux sortants quotidiens moyens avec<br>transformation de type puissance ..... | 45 |
| Figure 4.1 : Exemple de trois paires Origine-Destination.....  | 55 |
| Figure 4.2 : Plusieurs itinéraires possibles à gauche, une seule ligne de désir à droite .....                                       | 56 |
| Figure 4.3 : Sommaire des déplacements effectués en 2011 et 2014.....  | 57 |
| Figure 4.4 : Distribution et cumulatif des déplacements selon la distance réseau .....   | 60 |
| Figure 4.5 : Distribution et cumulatif des déplacements selon la durée sur distance réseau .....                                     | 61 |

|   |    |
|---|----|
| Figure 4.6 : Distribution de la somme des déplacements à l'Origine et à la Destination .....    | 62 |
| Figure 4.7 : Somme des déplacements aux stations de l'encadré #1 de la matrice O-D 2011.....    | 64 |
| Figure 4.8 : Somme des déplacements aux stations de l'encadré #2 de la matrice O-D 2011.....    | 65 |
| Figure 4.9 : Somme des déplacements aux stations de l'encadré #3 de la matrice O-D 2011.....    | 66 |
| Figure 4.10 : Lignes de désir plus importantes en 2011 et en 2014.....                          | 67 |
| Figure 4.11 : Différentes fonctions de coût.....  | 70 |
| Figure 4.12 : Fonctions de coût calibrées.....  | 73 |
| Figure 4.13 : Schéma des types de résidus.....  | 74 |
| Figure 4.14 : Cumulatif des déplacements observés et modélisés .....                            | 74 |
| Figure 4.15 : Cumulatif des résidus des modèles 2011 et 2014.....                               | 75 |
| Figure 4.16 : Distribution des résidus (nombre de déplacements prédits) .....                   | 76 |
| Figure 4.17 : Somme des résidus du modèle 2011 selon la distance de la station au centre-ville. | 77 |
| Figure 4.18 : Spatialisation de la somme des résidus du modèle 2011 .....                       | 78 |
| Figure 4.19 : Spatialisation de la somme des résidus du modèle 2014.....                        | 79 |
| Figure 4.20 : Spatialisation des paires O-D les moins bien estimées.....                        | 80 |

## LISTE DES SIGLES ET ABRÉVIATIONS

|      |   |
|------|---|
| CMM  | Communauté métropolitaine de Montréal   |
| GTFS | General Transit Feed Specification      |
| O-D  | Origine-Destination                     |
| OSRM | Open Source Routing Machine             |
| PSC  | Procédure séquentielle classique        |
| RTL  | Réseau de transport de Longueuil        |
| RTM  | Réseau de transport métropolitain (EXO) |
| SIG  | Système d'information géographique      |
| SR   | Secteur de recensement                  |
| STM  | Société de transport de Montréal        |
| VLS  | Vélo en libre-service                   |

## **LISTE DES ANNEXES**

|  |     |
|--|-----|
| Annexe A – Zones tampons de 400 m autour des stations BIXI en 2011.....  | 101 |
| Annexe B – Présentation des SR du recensement de 2011 .....  | 102 |
| Annexe C – Distribution et cumulatif des déplacements selon la distance réseau avec dénivelé                                     | 103 |
| Annexe D – Distribution et cumulatif de la durée des déplacements selon la distance réseau avec dénivelé .....                   | 104 |
| Annexe E – Matrice Origine-Destination des déplacements des membres lors de la pointe matinale des jours ouvrables en 2014 ..... | 105 |
| Annexe F – Statistiques sommaires des résidus absous (excluant 0) selon la distance au centre-ville.....                         | 106 |

## CHAPITRE 1 INTRODUCTION

Depuis le début du nouveau millénaire, la mobilité urbaine, et particulièrement l'offre existante, s'est grandement transformée. Les paysages des villes ne sont plus composés simplement des modes traditionnels tels voiture, autobus et taxi, mais ils sont dotés également de modes dits partagés, comme l'autopartage et le vélopartage. Ce dernier constitue un mode de transport qui existe depuis maintenant 50 ans, mais qui n'a vu son développement se populariser que depuis les années 2000, grâce à l'intégration des technologies de l'information. Effectivement, en 2009, 78 villes opéraient des systèmes de vélos en libre-service (VLS) comprenant un total d'environ 70 000 vélos (Midgley, 2009). Les plus récents chiffres, au début de l'été 2018, estiment l'existence de 1 608 systèmes (avec ou sans stations) avec une flotte totalisant plus de 18 millions de bicyclettes (Meddin & DeMaio, 2018).

Le vélopartage est souvent vanté pour ses nombreuses vertus. En effet, une administration municipale, universitaire ou autre mettant en service une telle infrastructure se dote de plusieurs objectifs, le plus souvent en lien avec le développement et la mobilité durable. Les désirs en ce sens sont multifacettes : le vélopartage devrait améliorer l'intermodalité ; avoir un impact sur la qualité de vie, de l'air et sur la pollution, et donc sur la santé ; diminuer la congestion routière ; et il pourrait même bonifier la vitalité commerciale et la solidarité régionale (Midgley, 2009; Shaheen, Guzman, & Zhang, 2010). Bref, les systèmes de VLS semblent représenter une solution innovante en ce qui concerne de nombreuses problématiques urbaines.

De son côté, la Ville de Montréal s'est lancée dans l'aventure du vélopartage avec BIXI, son propre système mis sur pied en 2009. Le souhait avoué de cette réalisation est de remplacer une portion des déplacements urbains présentement effectués en voiture (Midgley, 2009). La popularité de ce nouveau mode de déplacement à Montréal était telle que la mise en place de la deuxième phase, initialement prévue en 2010, a été devancée à la fin de l'été 2009. À l'époque, la métropole comptait sur 5 000 bicyclettes et 450 stations (Fuller et al., 2011). Le système s'est depuis stabilisé, permettant environ 100 000 déplacements par semaine en 2014 (Morency, Trépanier, Paez, Verreault, & Faucher, 2016). L'offre a récemment été augmentée : on retrouve aujourd'hui 540 stations, 6250 vélos et plus de 4,4 millions de déplacements ont été effectués en 2017, un record pour le réseau (Agence QMI, 2017).

## 1.1 Problématique

L'utilisation d'un système de vélopartage est relativement simple pour un usager. Abonné ou non, il est possible de louer une bicyclette directement aux bornes réparties dans un territoire donné. La location s'entame dès qu'un usager emprunte un vélo et se termine au moment où la bicyclette est rapportée à une station, qui peut être la même qu'à l'origine ou une autre. En revanche, la tâche est bien plus complexe pour l'opérateur, qui doit assurer un équilibre dans le réseau. Effectivement, les déplacements des bicyclettes par les usagers créent un déséquilibre aux stations, c'est-à-dire qu'il risque d'y avoir une distribution inégale des vélos (Vogel, P., Greiser, & Mattfeld, 2011). Or, les utilisateurs ne doivent pas se heurter à deux types de situations : une station vide, ce qui rend impossible une location, ou encore une station pleine, ce qui empêche de rapporter un vélo en location. Afin de minimiser ces risques, il est nécessaire d'effectuer une planification opérationnelle du réseau.

Selon Vogel, P. et al. (2011), il existe trois types de planification typiques dans un système de transport : stratégique, tactique et opérationnelle. Ces niveaux s'appliquent également aux VLS. Le niveau stratégique se déploie à long terme et consiste à optimiser l'emplacement des stations et leur capacité. Le volet tactique, quant à lui, s'effectue plutôt à moyen terme, et est orienté vers les consommateurs, en offrant des mesures leur permettant de mieux balancer le réseau, par exemple, avec des récompenses pécuniaires ou du temps d'utilisation crédité. Enfin, la planification opérationnelle se réalise à court terme et nécessite une redistribution en temps réel des bicyclettes à diverses stations. Celle-ci s'effectue le plus souvent par des camionnettes gérées par l'opérateur et des employés qui effectuent des transferts de vélos d'une station vers une autre, dans le but de rééquilibrer le réseau. Les décisions provenant du niveau stratégique influencent les deux niveaux inférieurs et la planification peut être compensée entre les niveaux : une mauvaise stratégie peut être comblée tactiquement et opérationnellement, et vice versa (Vogel, P. et al., 2011).

Heureusement, le système montréalais possède une grande flexibilité grâce à ses composantes modulables. En effet, les stations, qui comprennent les points d'ancrage des vélos, sont amovibles. Ce design a été développé afin de retirer les stations pendant l'hiver rigoureux que subit la métropole. Elles sont redéployées chaque printemps. Ceci est combiné au fait que le système peut donc s'ajuster en fonction de certains événements d'envergure ou même en fonction de travaux routiers. Ainsi, il est donc possible de changer l'emplacement des stations chaque année, ce qui

relève du niveau stratégique de planification. Cette malléabilité du système permet une optimisation annuelle, contrairement aux systèmes de VLS fixes.

Afin de tendre vers une planification et une utilisation optimale du système, il est pertinent d'évaluer les déplacements effectués en BIXI. Deux manières sont utilisées pour faire ceci, soit en observant les emprunts et retours de bicyclettes aux stations, ainsi qu'en analysant l'Origine et la Destination des déplacements des usagers. Or, ceci est complexe, car la demande de déplacements n'est pas uniforme : elle varie temporellement et spatialement.

## 1.2 Objectifs

Cette recherche s'intéresse principalement à obtenir de meilleurs éléments de compréhension en ce qui concerne la demande de déplacements en vélopartage. En ce sens, elle souhaite mettre en lumière des éléments de réponse concernant la variabilité spatio-temporelle qui existe dans l'utilisation du système et ce, dans le but d'améliorer la planification stratégique. Pour ce faire, les données du système de VLS montréalais BIXI sont mises à profit. La variabilité de l'utilisation de ce système dans le temps et dans l'espace est analysée selon la production et l'attraction des déplacements aux stations, ainsi qu'à une analyse des lignes de désir, c'est-à-dire ce qui lie l'Origine et la Destination des déplacements. De plus, la recherche s'intéresse à l'application de modèles classiques afin de reproduire ces comportements de déplacements. Ainsi, les deux premières étapes de la procédure séquentielle classique (PSC) sont utilisées, soit la production et la distribution de déplacements. Ces modèles sont ensuite validés sur les déplacements observés. La figure 1.1 offre un survol schématique de la méthodologie générale qui est déployée dans ce document.

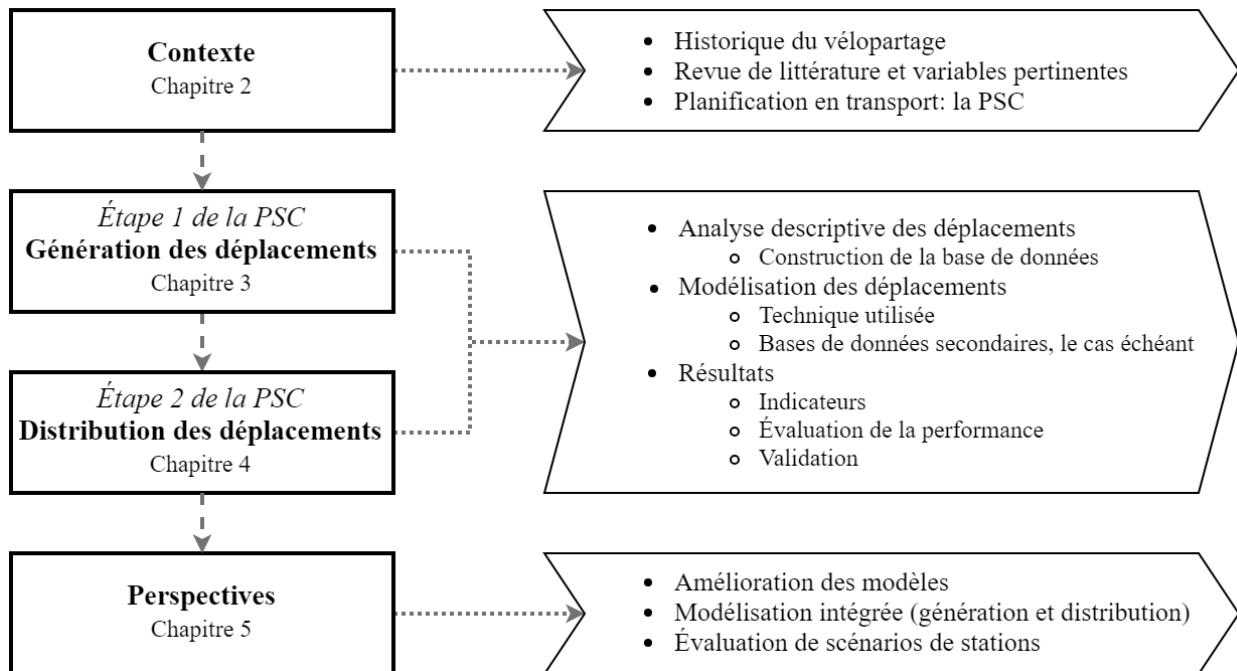


Figure 1.1 : Schéma de la méthodologie générale du projet

### 1.3 Structure du document

Le mémoire comprend 5 chapitres. Le deuxième chapitre est utilisé en guise de contextualisation générale par la présentation de concepts et éléments clés. Il propose un historique et une présentation du fonctionnement du vélopartage, ainsi qu'une revue de littérature des études réalisées sur le sujet. Il se termine sur la planification et la modélisation des déplacements en exposant le fonctionnement de la PSC. Le troisième chapitre est utilisé afin de modéliser la première étape de la PSC, soit la génération des déplacements. Une analyse des flux sortants et entrants aux stations est effectuée et la méthodologie utilisée est précisée. Le quatrième chapitre s'attarde quant à lui à la deuxième étape de la PSC, c'est-à-dire la distribution des déplacements. Dans ce cas, les lignes de désir des usagers sont analysées puis le modèle gravitaire est présenté, qui permet de reproduire les paires Origine-Destination (O-D) des utilisateurs BIXI. Le dernier chapitre propose un retour sur les faits saillants du mémoire en plus de proposer des perspectives de recherche sur les analyses réalisées, les modèles développés et éventuellement de nouvelles pistes à explorer.

## CHAPITRE 2 MISE EN CONTEXTE

La popularité des systèmes de VLS a explosé depuis le début du 21<sup>e</sup> siècle et les recherches sur le sujet se sont également multipliées pendant la même période. Grâce à l'intégration des technologies de l'information, il est possible d'utiliser et d'analyser des données transactionnelles obtenues des infrastructures de vélopartage afin de faire avancer la recherche sur la question des déplacements et le comportement des usagers. Ce chapitre consiste donc à mettre en contexte le vélopartage selon plusieurs aspects. D'abord, il est question de présenter le concept de vélopartage et sa place dans la mobilité urbaine. Puis, un survol de son historique et de son développement est effectué. Le fonctionnement typique de ces systèmes est présenté à travers le système BIXI et son utilisation à Montréal. Ensuite, une revue de littérature permet de faire un état des études existantes sur le vélopartage et sur BIXI spécifiquement, afin de dresser un portrait des connaissances acquises et des tendances qui s'en dégagent. Finalement, il est question de la planification et de la modélisation en transport. Une méthodologie générale est exposée grâce aux modèles à quatre étapes puisque le projet de recherche utilise deux de ces étapes, qui sont présentées plus en détail dans les prochains chapitres.

### 2.1 Le vélopartage dans la mobilité urbaine

Le concept de vélopartage, tel que simplement défini par Shaheen et al. (2010), permet une utilisation partagée d'une flotte de vélos. Il se différencie des services de location de vélo dits « traditionnels » puisque, entre autres, les bicyclettes peuvent être louées et rapportées à des endroits différents (Midgley, 2009) : l'usager est entièrement autonome. De plus, Shaheen et al. (2010) indiquent que ce type de déplacement constitue une stratégie durable face à diverses problématiques d'aménagement du territoire et environnementales. Nés d'un petit plan hollandais des années 1960 qui souhaitait résoudre les problèmes de congestion, les systèmes de VLS se sont, avec le temps, adaptés et modernisés face aux réalités urbaines, pour finalement devenir une véritable solution de mobilité quotidienne.

Le vélopartage possède un type d'utilisation propre à lui, tel qu'illustré à la figure 2.1. Grâce à un coût d'utilisation relativement faible, il est possible d'avoir recours à ce mode pour faire des déplacements de courtes et de moyennes distances. Situé quelque part entre la marche, le transport en commun et la voiture, il est différent d'une bicyclette privée puisque l'utilisateur n'est pas

dépendant de son mode et peut donc en combiner plusieurs. Il n'a pas non plus à se soucier de vols potentiels, ni encore de son entretien et d'un espace pour le garer ou encore le remiser l'hiver.

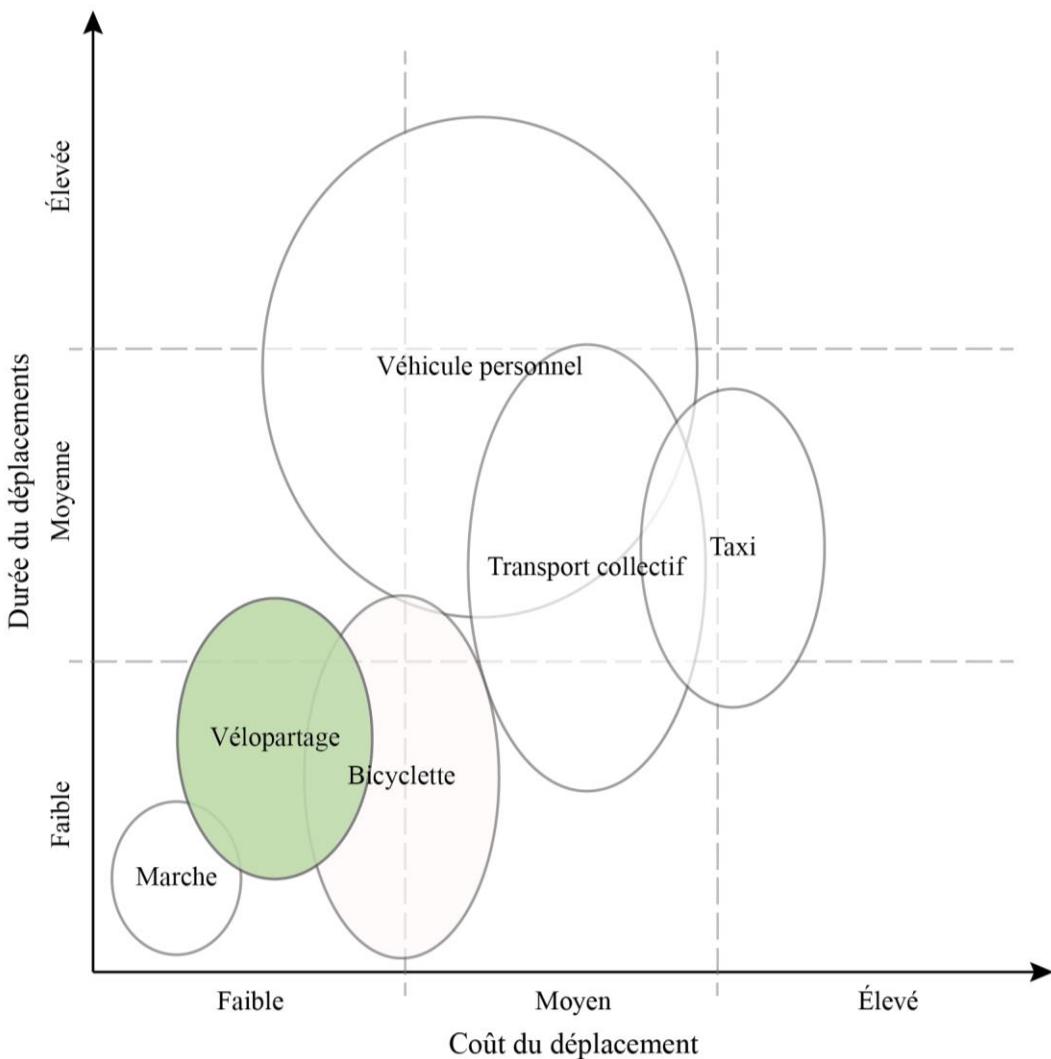


Figure 2.1 : La place du vélopartage dans l'offre de transport urbain.

Dans *TransLink public bike system feasibility study*, 2008, p. 86.

© Quay Communications Inc. (2008). Reproduit et traduit avec permission.

## 2.2 Histoire du vélopartage

La littérature tend à caractériser le développement du vélopartage en « générations », puisque celui-ci s'est effectué en plusieurs phases, relativement distinctes. Il est possible d'établir clairement trois générations, alors que la définition des étapes subséquentes, à partir des années 2000, est

moins certaine, puisque le développement et l'intégration de nouvelles technologies se font très rapidement et ne sont pas uniformes au sein des divers systèmes.

Ainsi, la première génération de vélopartage public apparaît en 1965 à Amsterdam. À ce moment, une flotte de 50 bicyclettes peintes en blanc, d'où le nom *Witte Fietsen* (vélos blancs), est répartie dans la ville, offerte gratuitement pour tous, sans dispositif de sécurité (DeMaio, 2009; Shaheen et al., 2010). Les vols nombreux et le vandalisme mettent fin rapidement à ce projet. Quelques autres moutures similaires voient le jour par la suite, dont celle à La Rochelle, où le projet *Vélos jaunes* est encore en service aujourd'hui avec une infrastructure mise à jour (Shaheen et al., 2010).

La deuxième génération apparaît au Danemark ce qui fait de Copenhague, en 1995, la première ville à planter une infrastructure à grande échelle (1 100 vélos). C'est à cet endroit qu'est introduit le concept de stations, où l'utilisateur peut louer un vélo en déposant l'équivalent de quelques dollars et en récupérant le montant au retour de la bicyclette (DeMaio, 2009; Fishman, 2016; Shaheen et al., 2010). Malgré une nette amélioration par rapport à la première génération et un système qui perdure, ce type de méthode permet tout de même des vols de vélos du à l'anonymat des usagers (DeMaio, 2009; Fishman, 2016).

Les systèmes de VLS de troisième génération présentent des corrections importantes par rapport à leurs prédecesseurs. Ce type, officiellement apparu en tant que système fermé pour les étudiants de l'université Portsmouth en Angleterre en 1996, est surtout popularisé par le Vélo'v lyonnais en 2005 avec ses 1 500 bicyclettes (DeMaio, 2009). En effet, la nouvelle phase inclut différentes avancées technologiques, dont des bornes aux stations avec interface destinées aux utilisateurs, le paiement par carte de crédit ou magnétique, la possibilité de s'abonner et donc une infrastructure permettant de minimiser le vol (DeMaio, 2009; Shaheen et al., 2010). La combinaison de ces différents perfectionnements a permis au vélopartage de gagner une popularité importante dans les milieux urbains.

La littérature évoque également une quatrième génération en devenir. Les caractéristiques de celle-ci ne font pas l'unanimité selon les auteurs, cependant, il semble y avoir consensus sur certains éléments. Les nouveaux systèmes de VLS doivent être munis, entre autres : d'un certain pourcentage de bicyclettes électriques, permettant notamment d'assister les personnes étant incapable de surmonter certains obstacles topographiques, de stations faciles à installer et à énergie solaire ainsi qu'une redistribution des vélos sur le territoire améliorée (DeMaio, 2009; Shaheen et

al., 2010). Or, les caractéristiques de chaque nouveau système sont déterminées en fonction des priorités de l'opérateur, ce qui fait que les infrastructures actuelles sont plutôt hybrides, quelque part entre la troisième et la quatrième génération. D'ailleurs, la nouvelle tendance chez les opérateurs est aujourd'hui d'offrir des vélos sans stations, munis de GPS, ce que Fishman (2016) considère comme une offre de la quatrième génération. Cette nouveauté, aussi connue comme le « Vélopartage 2.0 », est apparue en Chine, mais se propage déjà au Royaume-Uni et aux États-Unis et engendre de nouvelles problématiques liées à la gestion et l'entretien de la flotte (Normandin, 2017).

## **2.3 Particularités montréalaises**

C'est dans ce contexte qu'est né BIXI à Montréal en 2009, qui présente des caractéristiques des troisième et quatrième générations de vélopartage (Shaheen et al., 2010). Comme mentionné précédemment, les stations sont amovibles et elles sont retirées pendant la saison froide, ce qui signifie qu'elles s'installent relativement aisément alors que les bornes fonctionnent à l'énergie solaire. De plus, pour faire suite à un projet pilote permettant aux usagers de louer une bicyclette avec leur carte à puce de transport collectif l'an dernier (Bérubé, 2017), ce concept est progressivement en train de s'étendre à l'ensemble des stations. De plus, un projet pilote de vélos à assistance électrique a été lancé récemment.

### **2.3.1 Composantes et fonctionnement**

Le système montréalais possède un système relativement comparable à de nombreuses autres infrastructures contemporaines de vélopartage. Cette section permet de décrire quelques composantes principales inhérentes au fonctionnement du système qui sont souvent utilisées dans le cadre de ce document, soit les stations, la notion de capacité, le type d'usager, les déplacements et les saisons.

- Stations : les stations agissent à titre d'intermédiaire entre les vélos et les utilisateurs. Elles comprennent des points d'ancrage où sont sécurisées les bicyclettes ainsi qu'une borne d'information, illustrée à gauche sur la figure 2.2. Cette borne est munie d'un écran destiné aux utilisateurs, où il est possible d'acheter un accès de courte durée (la tarification est expliquée en détail dans la section 2.3.2). Chaque point d'ancrage d'une station est muni d'un petit clavier, tel qu'illustré à la figure 2.3, où il est possible d'insérer un code, ou

encore une clé pour les clients abonnés qui n'ont pas à passer par la borne principale, afin de déverrouiller le vélo attaché ;



Figure 2.2 : Station BIXI, comprenant borne d'accès et points d'ancrage.

© BIXI Montréal (2018c). Reproduit avec permission.

- Capacité : la capacité des stations, dont il est question dans ce rapport, constitue la somme des points d'ancrage à une station. Le nombre de points d'ancrage n'est pas identique à chaque station. Normalement, une station située à un endroit susceptible de générer ou d'attirer de nombreux déplacements aura une capacité plus importante, selon l'espace disponible. Les points d'ancrage sont installés sur une plaque au sol avec quatre orifices. Une station comporte minimalement sept points d'ancrage, soit une plaque avec quatre points d'ancrage et une autre avec trois points d'ancrage et une borne. Toute augmentation de capacité se fait donc par quatre points d'ancrage à la fois ;



Figure 2.3 : Point d’ancrage avec clavier et fente pour clé.

© BIXI Montréal (2018b). Reproduit avec permission.

- Type d’usagers : ce document fait référence à deux types d’usagers, soit les membres et les utilisateurs occasionnels. Un membre est un usager qui possède un abonnement à long terme au système, que celui-ci soit hebdomadaire, mensuel ou encore saisonnier. Un usager occasionnel possède plutôt un accès limité au réseau, variant de 24 à 72 h. Les coûts d’accès sont présentés dans la section suivante ;
- Déplacements : un déplacement à BIXI s’entame lors de la location d’une bicyclette à une station et se termine lors du retour de ce vélo à une station, que ce soit la même qu’à l’Origine ou une différente. Ainsi, un déplacement, dans le cadre de ce mémoire, fait référence seulement au trajet effectué d’une station vers une autre et exclut toute combinaison possible avec un autre mode. L’Origine et la Destination sont donc des stations de vélopartage ;
- Saisons : les saisons BIXI se déroulent habituellement du 15 avril au 15 novembre de chaque année depuis 2009. En ce sens, aucune distinction n’est faite entre le terme « annuel » et « saisonnier ».

### 2.3.2 Tarification usuelle

Les tarifs de BIXI ont évolué au cours des années. Seules les valeurs de 2018 sont présentées dans cette section afin de donner une idée des divers prix et types d'accès au système. Ainsi, une personne souhaitant accéder au réseau BIXI doit d'abord acquérir un certain forfait au système, ce qui passe par un abonnement, par l'entremise du site web, pour diverses périodes (annuelle, 90 jours, 30 jours ou 24 h), ou encore par un accès de courte durée (aller simple, 24 h ou 72 h), ce qui se choisit directement à une borne (BIXI Montréal, 2018d). Les abonnés reçoivent une clé BIXI par la poste afin de faciliter la location d'une bicyclette : ils n'ont qu'à insérer leur clé au point d'ancrage sans passer par la borne principale. Le forfait 24 h est particulier puisqu'il est offert en abonnement et à court terme, comme présenté au tableau 2.1. Ceci s'explique par le fait que l'abonnement permet d'avoir accès à une clé, ce qui facilite la location d'un vélo, évitant de passer par la borne. Certaines stations permettent l'accès avec la carte à puce de transport en commun, ce qui a déjà été mentionné. Plusieurs autres forfaits sont disponibles avec des partenaires, dont la Société de transport de Montréal (STM), l'opérateur du transport collectif de Montréal, ou encore avec des opérateurs de véhicules en libre-service.

Tableau 2.1 : Tarification du système BIXI en 2018

| Forfait      | Méthode  | Période      | Tarif (\$) | Minutes incluses | Premier 1/4 h après minutes incluses (\$) | 1/4 h supp. (\$) |
|--------------|----------|--------------|------------|------------------|---|------------------|
| Abonnement   | En ligne | 1 an         | 91         | 45               | 1,80                                      | 3                |
|              |          | 90 jours     | 57         |                  |   |                  |
|              |          | 30 jours     | 32         |                  |   |                  |
|              |          | 24 h         | 5          |                  |   |                  |
| Courte durée | Stations | 24 h         | 5          | 30               |   |                  |
|              |          | 72 h         | 10         |                  |   |                  |
|              |          | Aller simple | 2,95       |                  |   |                  |

Selon le type d'abonnement ou d'accès, il est possible d'emprunter un vélo pour 30 ou 45 minutes sans frais supplémentaires, et ce, autant de fois que désiré dans la période déterminée selon le forfait. Lorsqu'une location excède la durée incluse avec le forfait, des frais excédentaires sont engendrés pour le premier quart d'heure, et des frais encore plus élevés s'appliquent par tranche de quinze minutes suivantes. Cette méthode vise à restreindre la durée des emprunts et inciter à en

faire plusieurs, au besoin (BIXI Montréal, 2018a), et ce, afin d'assurer une disponibilité de vélos pour les autres utilisateurs.

### 2.3.3 Répartition spatiale des stations en 2011

La figure 2.4 vient conclure ce portrait du réseau BIXI. Il est possible d'y voir la répartition spatiale des stations en l'an 2011 ainsi qu'un point rouge au coin des rues Peel et Sainte-Catherine. Cet endroit est utilisé comme référence au centre-ville dans l'ensemble de ce rapport. De plus, la figure permet également de voir la distance euclidienne (soit à vol d'oiseau) de chaque station par rapport à ce point, élément important dans l'analyse et la modélisation effectuées dans les prochains chapitres. Plusieurs stations sont apparues depuis 2011 et le réseau a été modifié. En revanche, cette année sert de base pour l'ensemble de la recherche et permet ainsi de représenter géographiquement le réseau.

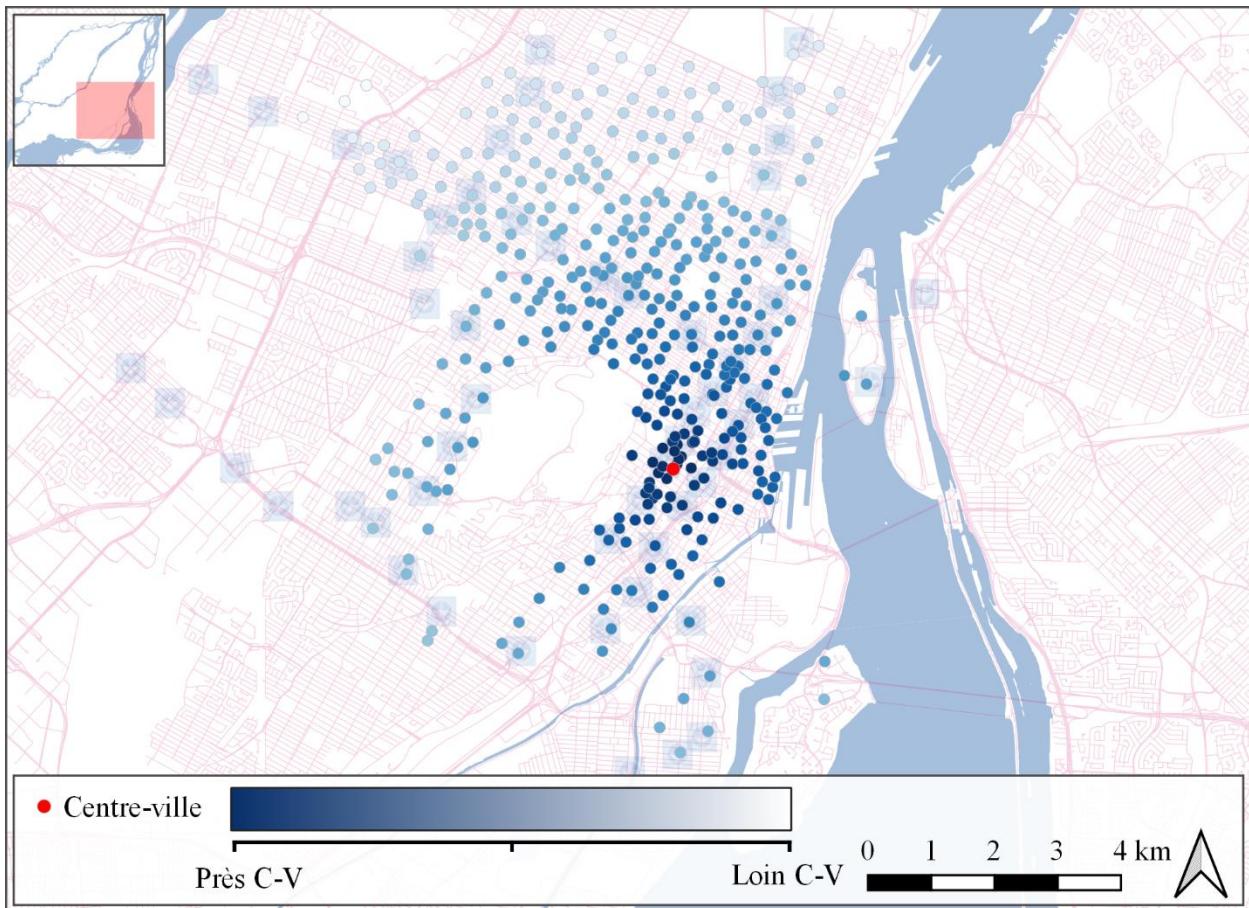


Figure 2.4 : Spatialisation du réseau 2011 et distance au centre-ville

## 2.4 Études sur le vélopartage

Les recherches sur le vélopartage ont augmenté significativement depuis l'arrivée de la troisième génération et ce, grâce à l'intégration des technologies de l'information aux systèmes. Effectivement, celles-ci permettent d'attribuer à chaque déplacement l'heure et la date de la location et du retour de la bicyclette, ainsi que les stations d'Origine et de Destination. Ces données consistent en une excellente source d'analyse des déplacements urbains à bicyclette et même pour comprendre les dynamiques urbaines et les comportements humains (Froehlich, Neumann, & Oliver, 2009).

Or, les sujets de recherche des différentes études sont vastes, variant, entre autres, des manières d'implanter un système dans un nouvel environnement, à l'optimisation de l'emplacement des vélos en temps réel et à court terme. Puisque l'objectif de cette recherche tend vers une planification stratégique, seulement la littérature étant en lien avec ce but est présentée, qui permet de bonifier les connaissances en ce sens. Ces études sont généralement orientées autour de trois composantes principales du vélopartage, soit les usagers qui utilisent le système, les stations et les déplacements entre celles-ci. La revue de littérature est ainsi séparée en deux grandes thématiques, soit la demande aux stations et la distribution spatiale des déplacements.

### 2.4.1 La demande de déplacements en vélopartage

Une grande partie des études réalisées à ce jour consiste à utiliser et analyser les données opérationnelles de différents systèmes de vélopartage. Celles-ci permettent de déterminer la demande de déplacements, en fonction des bicyclettes empruntées et retournées aux stations. Il est plus complexe d'évaluer une demande non comblée, c'est-à-dire lorsqu'une station est vide, ce qui empêche un déplacement potentiel, ou encore lorsqu'une station est pleine, ce qui modifie possiblement le parcours d'un usager, devant retourner la bicyclette à une autre station. Ceci est nécessairement une limite quant à l'évaluation de la demande de déplacements.

Omettant ceci, il semble exister une dépendance spatio-temporelle dans l'utilisation du vélopartage (Vogel, P. et al., 2011), concept qui revient fréquemment dans la littérature. L'étude effectuée par Froehlich et al. (2009) est l'une des premières à démontrer que le comportement des utilisateurs aux stations semble lié à la localisation et à une certaine temporalité. Dans la même veine, Borgnat et al. (2010) analysent les dynamiques du Vélo'v de Lyon et déterminent que certaines stations

sont plus populaires que d'autres selon le moment de la journée (matin, midi, soir), alors qu'il y a une utilisation plus concentrée en fin de semaine pour certaines stations. De la même manière, Vogel, P. et al. (2011) révèlent des dépendances spatio-temporelles selon les tendances d'activités de début et de fin de location de VLS aux stations du Citybike de Vienne.

Les recherches subséquentes précisent les facteurs ayant un impact sur la demande, dont la temporalité et l'environnement urbain, ainsi que des éléments exogènes comme les conditions météorologiques. D'abord, Buck et Buehler (2012) se basent sur une moyenne de déplacements par jour pour conclure que l'installation de stations du système de vélopartage Capital Bikeshare de Washington, D.C., à proximité d'infrastructures cyclables en augmenterait l'utilisation. Rixey (2013), quant à lui, se base sur la moyenne mensuelle de location aux stations de trois villes étatsunies pour déterminer l'importance de certains facteurs, dont la densité de population, le revenu moyen et même la forte présence de minorités visibles dans l'estimation de déplacements. Similairement, Daddio (2012) utilise les déplacements d'un seul mois pour produire une carte de potentialité d'emplacement de stations, à Washington, D.C. Wang, Lindsey, Schoner et Harrison (2016) notent une corrélation entre l'utilisation moyenne par jour aux stations et des lieux d'intérêts, tels campus universitaires et commerces de restauration. Grâce à un sondage en ligne, Raux et Zoubir (2016) déterminent qu'un certain statut social et la proximité résidentielle aux stations sont deux facteurs qui ont un effet positif et indépendant sur la probabilité d'être un membre annuel à Lyon.

Certaines études se penchent également sur l'impact du vélopartage sur d'autres modes de transport. Par exemple, Hampshire et Marla (2012) analysent des données des systèmes de deux villes espagnoles pour déterminer qu'à Barcelone, le vélopartage est un compétiteur aux autres modes de déplacements alternatifs à l'automobile alors qu'il est plutôt complémentaire dans le cas de Séville. Nair, Miller-Hooks, Hampshire et Bušić (2013) révèlent également que la proximité entre une station de métro de Paris et une station Vélib' peut résulter en une utilisation accrue de ces deux modes par un utilisateur, devenant ainsi une valeur ajoutée.

Regardant les VLS d'un autre angle, Mahmoud, El-Assi et Habib (2015) se penchent plutôt sur la question de la capacité des stations du système de Bike Share Toronto, grâce à des données d'utilisation mensuelles. Ils notent qu'une station avec une plus grande capacité possède plus de chances d'attirer et de générer davantage de déplacements et, par le fait même, le nombre de

stations existantes dans un rayon de 200 mètres est corrélé avec le nombre de déplacements attirés. Similairement, Tran, Ovtracht et d'Arcier (2015) notent l'importance de la densité du réseau d'un système de VLS et la capacité des stations quant à l'utilisation.

En plus d'analyser l'environnement dans lequel est implantée une station, des études notent une différence d'utilisation selon le type d'usager. Effectivement, les abonnés utilisent davantage le vélopartage la semaine, aux heures de pointe, alors que les usagers occasionnels sont plus fréquents la fin de semaine, et le temps de déplacement est également généralement plus long. En ce sens, Tran et al. (2015) démontrent qu'à Lyon, les déplacements des abonnés s'expliquent en fonction de la population et des emplois, alors que les déplacements des utilisateurs de courte durée s'expliquent davantage par des variables spatiales liées aux loisirs.

Parallèlement à ces études, plusieurs recherches ont déjà été effectuées spécifiquement sur le cas du BIXI. D'abord, Morency, Trépanier et Godefroy (2011) utilisent les données disponibles de la première saison d'utilisation (2009) pour caractériser le système et les utilisateurs du réseau, en plus de fournir une série d'indicateurs analytiques pertinents. De leur côté, Fuller et al. (2011) effectuent un sondage téléphonique et déterminent un lien potentiel entre l'utilisation du système et la proximité des répondants aux stations, ainsi que l'âge, le genre, l'emploi et le niveau de scolarité. Faghih-Imani, Eluru, El-Geneidy, Rabbat et Haq (2014) quantifient l'influence de plusieurs facteurs sur les départs et les arrivées aux stations et révèlent, entre autres, que les stations près de routes artérielles possèdent moins d'activité d'utilisation que celles près des routes locales et des pistes cyclables. De plus, ces chercheurs font le constat qu'il serait plus pertinent d'installer davantage de petites stations que d'augmenter la capacité de stations existantes. Morency et al. (2016) démontrent que les usagers sont beaucoup plus jeunes que la population de la Ville de Montréal et que l'utilisation est plus importante chez les hommes que chez les femmes. Ceci concorde avec les résultats de la plupart des études sur le sujet (Fishman, 2016; Fishman, Washington, & Haworth, 2013). Ils s'intéressent également à la corrélation des conditions météorologiques et la temporalité de l'utilisation, sur une période de six ans de données transactionnelles. Finalement, Morency, Trépanier, Frappier et Bourdeau (2017) créent une typologie des membres selon leurs tendances d'utilisation du réseau : fréquence et temporalité. Ils dénotent une corrélation entre la proximité des stations et l'usage annuel moyen. De plus, ils étudient le taux de rétention des membres du système à travers les saisons d'opération.

## 2.4.2 La distribution spatiale des déplacements en vélopartage

Ensuite, toujours dans une optique de planification stratégique, il est question d'étudier la distribution de ces déplacements, c'est-à-dire la liaison entre l'Origine et la Destination. Celle-ci occupe un espace beaucoup moins important que la demande de déplacements aux stations dans la littérature.

Faghih-Imani et Eluru (2015) se penchent spécifiquement sur la question des Destinations des déplacements du système Divvy de Chicago. Ils analysent des facteurs ayant une influence selon cinq périodes dans une journée, soit l'âge et le genre de l'utilisateur, la distance entre l'Origine et la Destination, la présence d'infrastructures cyclables, l'usage du sol et l'environnement bâti. D'une manière similaire, González, Melo-Riquelme et de Grange (2016) catégorisent les infrastructures cyclables et incluent d'autres éléments comme la présence d'infrastructure de transport collectif lourd, de bureaux et même d'arbres dans un modèle qui combine la distribution et le choix d'itinéraire.

Dans un autre ordre d'idée, Zhang, He, Dong et Zhou (2016) se penchent plutôt sur la durée des déplacements dans trois villes chinoises qu'ils intègrent dans un modèle gravitaire en tant que fonction de coût dans une étude subséquente (Zhang, Ji, Sheng, & Zhou, 2018).

Ces diverses études constituent de bonnes pistes afin d'étudier certains éléments ayant un impact sur les déplacements dans un système de vélopartage.

## 2.5 Méthodologie générale : la procédure séquentielle classique

La modélisation des déplacements est inhérente à la planification en transport puisque celle-ci souhaite reproduire des comportements observés afin de tenter de dessiner des tendances futures. Ainsi, la méthodologie privilégiée emprunte les deux premières étapes de la PSC, aussi connue comme modèles à quatre étapes. Un survol de ces modèles est effectué dans cette section et les étapes utilisées en tant que méthodologie de modélisation des déplacements sont présentées plus en profondeur. Les explications concernant la PSC sont basées sur celles fournies par Bonnel (2002), McNally (2007) et Ortuzar et Willumsen (2011).

La PSC se veut une structure permettant la modélisation des déplacements effectués sur un territoire en fonction d'étapes successives, dans un but de planification. La modélisation elle-même se base sur des échantillons provenant de données et sur des notions et théories du comportement

humain. Quoiqu'elle provienne des années 1960, elle est toujours pertinente aujourd'hui. Elle se base sur les liens existants entre les infrastructures de transports et le système des activités qui sont dispersées sur un territoire, ainsi que les citoyens qui effectuent des déplacements entre ceux-ci afin d'y effectuer des activités, ce qui est schématisé à la figure 2.5. Les composantes principales des modèles à quatre étapes sont la génération des déplacements, leur distribution dans l'espace, le choix du mode pour se mouvoir dans l'espace, ainsi que l'affectation, c'est-à-dire le trajet emprunté sur le réseau de l'Origine à la Destination.

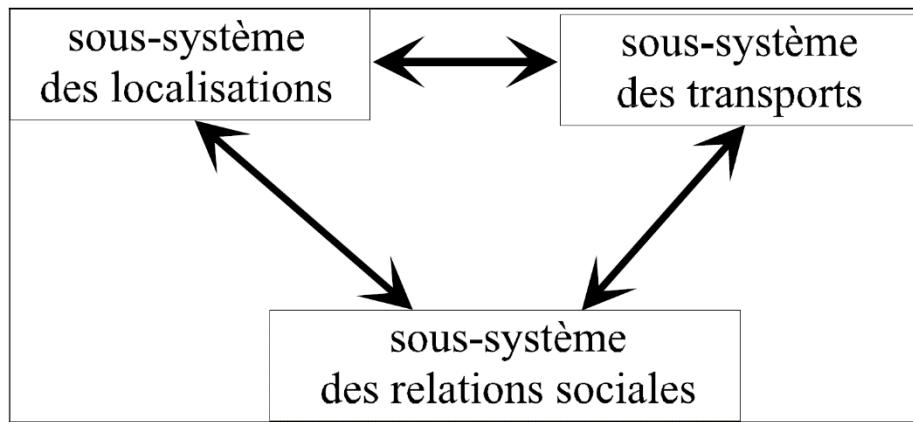


Figure 2.5 : Interactions prépondérantes entre les trois sous-systèmes. Dans *Prévision de la demande de transport*, 2002, p. 32. © Bonnel (2002). Reproduit avec permission.

La figure 2.6 schématise les étapes de la PSC qui sont présentées en bref dans les sections suivantes. D'abord, un territoire est défini, tout comme les réseaux qui y sont compris. Ceci permet d'établir le scénario de référence, soit l'état actuel des choses. À partir de ceci, le développement d'un modèle à chaque étape et les résultats de celui-ci permettent de mener à une étape subséquente de la PSC. Une ou plusieurs phases peuvent être considérées et des itérations sur les étapes peuvent être réalisées selon les besoins. Cette technique presuppose certaines hypothèses ce qui signifie qu'elle comporte son lot de limites, qui sont présentées après les étapes.

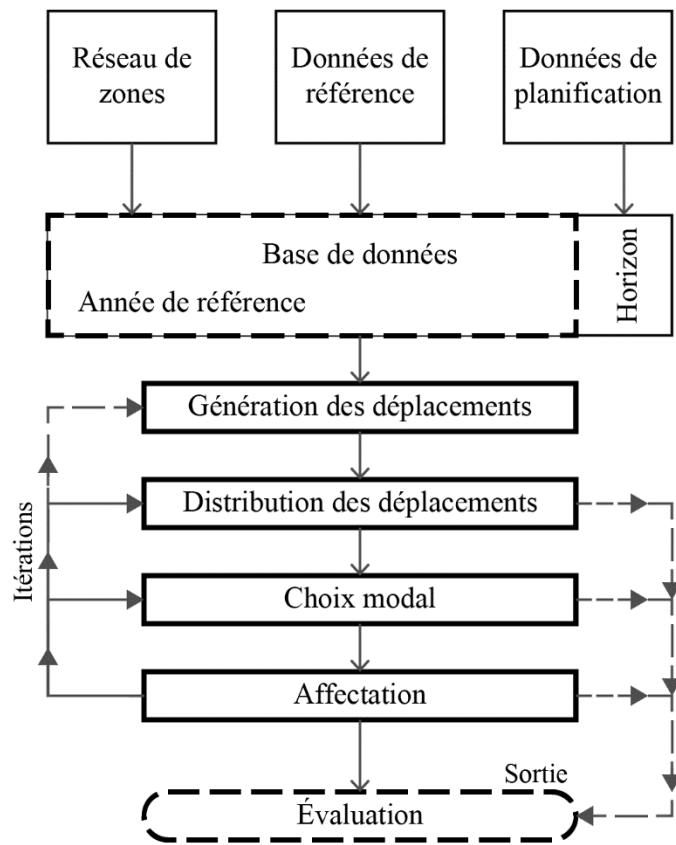


Figure 2.6 : Schéma de la procédure séquentielle classique. Dans *Modelling transport*, 2011, p. 21. © Ortuzar et Willumsen (2011). Reproduit et traduit avec permission.

Avant de passer à ces étapes, il semble pertinent de présenter le concept de matrice O-D, soit le tableau 2.2, qui est crucial pour la bonne compréhension des phases de la PSC. Cette matrice comprend, en rangée, l'Origine des déplacements, et la Destination de ceux-ci est organisée en colonnes. Chaque cellule dans cette matrice représente le nombre de déplacements pour chaque paire O-D. Traditionnellement, l'Origine et la Destination sont des zones géographiques d'un territoire alors que dans le cas présent, il s'agit de stations BIXI. Les marges, quant à elles, constituent les dernières rangée et colonne. Elles sont en fait la somme des déplacements produits et attirés à chaque station. La modélisation permet de prendre une matrice vide et de la remplir d'informations concernant les déplacements.

Tableau 2.2 : Exemple de matrice Origine-Destination remplie

| Stations               |     | Destination  |            |              |     | Somme dépl.<br>produits |
|------------------------|-----|--------------|------------|--------------|-----|-------------------------|
|                        |     | 1            | 2          | 3            | ... |                         |
| Origine                | 1   | 0            | 1          | 0            | ... | <b>615</b>              |
|                        | 2   | 5            | 1          | 10           | ... | <b>569</b>              |
|                        | 3   | 9            | 9          | 2            | ... | <b>835</b>              |
|                        | ... | ...          | ...        | ...          | ... | ...                     |
| Somme dépl.<br>attirés |     | <b>1 169</b> | <b>854</b> | <b>2 134</b> | ... | <b>432 884</b>          |

### 2.5.1.1 Étape 1 : génération des déplacements

Les modèles de génération des déplacements consistent à remplir les marges de la matrice, soit la totalité des déplacements produits et attirés par chaque station BIXI, comme illustré au tableau 2.3. Il est possible d'y constater que les cellules ne contiennent aucune information puisque celle-ci est déterminée à une étape subséquente.

Tableau 2.3 : Génération de la matrice Origine-Destination

| Stations               |     | Destination |     |       |     | Somme dépl.<br>produits |
|------------------------|-----|-------------|-----|-------|-----|-------------------------|
|                        |     | 1           | 2   | 3     | ... |                         |
| Origine                | 1   | -           | -   | -     | -   | 615                     |
|                        | 2   | -           | -   | -     | -   | 569                     |
|                        | 3   | -           | -   | -     | -   | 835                     |
|                        | ... | -           | -   | -     | -   | ...                     |
| Somme dépl.<br>attirés |     | 1 169       | 854 | 2 134 | ... | 432 884                 |

La production d'un déplacement signifie que celui-ci est entamé à partir d'une station alors que l'attraction constitue la station de la fin d'un déplacement. Les termes productions et flux sortants sont utilisés de manière synonyme tout comme les termes attraction et flux entrants. L'étape de génération comprend ces deux éléments, car ils sont estimés de la même manière. Pour ce faire, il est question de déterminer les facteurs ayant un impact sur la probabilité que des déplacements soient effectués à partir d'une certaine station, qui devient ainsi l'Origine, ou vers une certaine station, qui devient la Destination. Les facteurs considérés dans ces modèles sont généralement reliés à des aspects sociodémographiques, à l'usage du sol, à l'environnement bâti ainsi que l'offre existante en transport. Des modèles de génération sont estimés afin de reproduire des déplacements observés aux stations BIXI dans le chapitre 3 et la technique utilisée, soit la régression linéaire multiple, est expliquée en détail à la section 3.2.1.

### 2.5.1.2 Étape 2 : distribution des déplacements

Alors que l'étape de génération permet de remplir les marges d'une matrice Origine-Destination, celle de distribution permet quant à elle de remplir toutes les cellules de la matrice. Effectivement, elle *distribue* tous les déplacements provenant d'une certaine station vers l'ensemble des stations. Pour représenter ceci, le tableau 2.4 est utilisé comme exemple. Cette étape utilise les marges (dernières colonne et rangée), qui constituent la somme des déplacements produits et attirés à chaque station, pour ensuite les distribuer dans chacune des cellules. Sous ses formes les plus simples, cette étape utilise une fonction de coût entre deux stations, afin de déterminer la probabilité que des déplacements y soient effectués. Cette fonction de coût représente généralement la distance ou le temps à parcourir entre une Origine et une Destination et représente un degré d'attractivité. Elle agit de manière négative : plus elle est élevée, moins le nombre de déplacements modélisés est important, donc moins une paire O-D est attractive. En ce sens, elle peut également être appelée fonction de résistance ou de friction. Alors que plusieurs types de modèles existent, la variante gravitaire est celle retenue pour modéliser la distribution des déplacements en BIXI. Les précisions sur ce type de modèle sont présentées à la section 4.4.1, alors qu'il est utilisé afin de reproduire la distribution des déplacements dans le chapitre 4.

Tableau 2.4 : Distribution de la matrice Origine-Destination

| Stations               |     | Destination |     |       |     | Somme dépl.<br>produits |
|------------------------|-----|-------------|-----|-------|-----|-------------------------|
|                        |     | 1           | 2   | 3     | ... |                         |
| Origine                | 1   | 0           | 1   | 0     | ... | 615                     |
|                        | 2   | 5           | 1   | 10    | ... | 569                     |
|                        | 3   | 9           | 9   | 2     | ... | 835                     |
|                        | ... | ...         | ... | ...   | ... | ...                     |
| Somme dépl.<br>attirés |     | 1 169       | 854 | 2 134 | ... | 432 884                 |

### 2.5.1.3 Étape 3 : choix modal des déplacements

Les modèles de choix modal consistent à déterminer le mode de transport utilisé pour effectuer les déplacements. Les facteurs ayant une influence dans ces modèles sont généralement des attributs des personnes, par exemple la possession d'un permis de conduire et d'une voiture, le salaire, et des caractéristiques du ménage, des variables concernant le déplacement, comme le motif, le moment où le déplacement est effectué et les conditions météorologiques, ainsi que des indicateurs décrivant le transport en soi, comme le temps du déplacement et de l'attente pour l'effectuer, le coût de celui-ci, la sécurité, le confort, etc. Bref, cette étape décuple la matrice Origine-Destination en autant de versions que de modes de transports considérés par le modèle.

### 2.5.1.4 Étape 4 : affectation des déplacements

Finalement, l'étape d'affectation tente de déterminer le trajet qui est utilisé pour relier une Origine et une Destination. Ainsi, les matrices O-D sont chargées sur un réseau codifié. Selon le modèle d'affectation préconisé, un calcul permet d'obtenir le coût d'un déplacement sur chaque lien possible entre une Origine et une Destination. Ce coût favorise les distances les plus courtes et dépend majoritairement de la capacité routière des liens, du moins en ce qui concerne le mode automobile. Ces types de modèles sont itératifs afin d'obtenir un état d'équilibre sur l'ensemble

des liens selon la congestion, donc le nombre de véhicules attribués par lien comparativement à la capacité.

### **2.5.1.5 Limites générales de la procédure séquentielle classique**

Chacune des étapes précédemment énumérées possède des limites, c'est-à-dire qu'elles possèdent des hypothèses propres et sont donc munies de certains biais. Les limites pertinentes dans le cas de cette étude sont présentées à la fin des chapitres de modélisation, soit 3 et 4. En plus de celles-ci, la PSC possède également des limites générales qu'il est pertinent de souligner. D'abord, cette manière de procéder divise la planification des déplacements en étapes, ce qui présuppose qu'un choix est fait par l'usager avant chaque phase. Or, cette manière séquentielle de réfléchir ne représente pas le processus de décision d'un humain (Mladenović & Trifunović, 2014). De plus, le degré d'agrégation de la modélisation a nécessairement un impact sur la précision. En ce sens, le niveau utilisé dans cette recherche, soit les stations, est certainement plus précis que lorsque la modélisation est faite de manière zonale, mais tout de même moins précise que lorsque totalement désagrégée, soit la considération de chaque déplacement individualisé. Un dernier élément important est l'aspect statique des modèles à quatre étapes. Quoiqu'un processus itératif permette d'équilibrer les étapes entre elles, il y a nécessairement des données entrantes qui sont exogènes aux modèles qui, en réalité, ont une relation avec les acteurs et la mobilité des usagers. Ceci est le cas, entre autres, lorsque une station est pleine ou vide, ce qui a nécessairement un impact sur la possibilité d'emprunter ou de retourner un vélo. De plus, le rééquilibrage effectué par l'opérateur, soit la redistribution des vélos BIXI par camionnettes, est un élément qui ne pourrait être intégré dans la PSC.

## **2.6 Conclusion du chapitre**

Ce chapitre a permis de présenter le vélopartage non seulement dans le contexte urbain, mais également à travers les diverses étapes qui ont ponctué son développement et les nombreuses variantes techniques. De plus, son fonctionnement général a été présenté à travers le système montréalais BIXI, ainsi que ses composantes clés, soit le concept de station, de capacité, de type d'usager, de déplacement et de saison BIXI, éléments récurrents dans ce mémoire. Le réseau montréalais de stations a également été représenté spatialement, ce qui est pertinent pour les étapes subséquentes d'analyse et de modélisation des déplacements dans les chapitres suivants.

Ensuite, une revue de littérature a dévoilé certaines connaissances acquises sur le vélopartage. Alors que cette recherche possède une vision de planification stratégique, deux sections ont divisé cet état de l'art. La première concerne la demande de déplacements en vélopartage aux stations et a présenté surtout les éléments qui ont un impact sur la production et l'attraction de déplacements, ce qui est en lien avec les modèles de génération. La seconde a fourni quelques études qui se sont penchées sur le choix de Destination ou de la durée des déplacements, thématique qui est reprise dans la modélisation de la distribution des déplacements.

Finalement, le concept de planification en transport a été abordé à travers la PSC. Chaque étape a donc été présentée alors que les deux premières ont été détaillées davantage, car elles sont utilisées par la suite, soit dans les étapes de modélisation. Les limites générales de la PSC ont conclu sa présentation.

## CHAPITRE 3 MODÉLISATION DE LA GÉNÉRATION DES DÉPLACEMENTS

Dans une perspective de planification stratégique, il est question dans ce chapitre d'observer les déplacements effectués aux stations BIXI sur une base annuelle. Une analyse temporelle est d'abord effectuée et permet de voir la fluctuation dans l'usage du système de vélopartage et certaines tendances. Pour ce faire, une base de données des déplacements observés en BIXI est mise à profit et l'analyse permet un certain tri des informations. Ensuite, les flux sont modélisés sous la forme de moyenne quotidienne de déplacements sortants et entrants selon l'étape de génération de la PSC. Pour ce faire, le concept de régression linéaire multiple est présenté en détail. Les bases de données et diverses variables utilisées afin d'expliquer les flux sont également présentées, ainsi que les diverses manipulations effectuées dans le but d'améliorer les résultats de la modélisation. Ainsi, des modèles de génération sont estimés et les résultats sont présentés. Finalement, ils sont validés sur des données d'une année ultérieure.

### 3.1 Analyse et triage des données

Les données utilisées dans ce chapitre proviennent de BIXI de 2011. Cette précision permet d'avoir une synchronicité dans les données, notamment avec une base de données importante utilisée dans le cadre de la modélisation, soit celle du Programme du recensement, réalisé par Statistique Canada de manière quinquennale. Cette base de données est présentée plus en détail à la section sur les variables indépendantes des modèles, soit 3.2.1.2.

La base de données des déplacements BIXI contient la somme quotidienne des flux sortants et entrants à chaque station, ainsi que les coordonnées de ces dernières. Chaque station BIXI permet à un utilisateur d'emprunter et de retourner une bicyclette. Il est donc question de flux sortants lorsqu'il y a production d'un déplacement, soit l'emprunt d'une bicyclette, et à l'inverse, un flux entrant concerne l'attraction d'un déplacement, donc le retour d'un vélo à une station. Afin d'alléger le texte, lorsqu'aucune mention n'est faite, les flux sortants sont considérés comme des déplacements.

La base de données comprend également la capacité de chaque station pour chaque jour, soit le nombre de points d'ancre. Cette capacité est calculée de manière quotidienne en tant que somme des points d'ancre libres et utilisés. Lorsqu'il est question de capacité moyenne dans le texte,

c'est que celle-ci fluctue dans le temps, causé entre autres par des points d'ancrage et des vélos défectueux, ou encore par l'ajout ou le retrait de points d'ancrage.

La première partie du chapitre consiste donc à analyser les données de cette base de données. Ainsi, en 2011, le réseau BIXI comprenait 405 stations et pendant les 215 jours d'opération (15 avril au 15 novembre), 4 167 039 flux entrants et 4 167 265 flux sortants ont été répertoriés. La différence de 226 flux est possiblement due à des erreurs d'enregistrements, à des vols ou encore certains retours de vélo effectués après minuit lors de la dernière journée de la saison. Il y a donc eu, en moyenne, 19 382 déplacements par jour et 49 déplacements quotidiens par station. Or, ces emprunts varient beaucoup selon le moment de l'emprunt et selon la localisation de la station, ce qui est analysé par la suite. Cet élément a également été révélé dans la revue de littérature.

### 3.1.1 Analyse temporelle

D'abord, il est question d'observer les déplacements selon le moment de l'année. Ainsi, le tableau 3.1 permet de constater ces grandes variations dans les déplacements quotidiens selon le mois. Le code de couleur par colonne permet de représenter les valeurs les plus faibles en bleu et les plus élevées en rouge. Il est aisément de constater les grandes disparités dans les déplacements. Les épaules de la saison de BIXI, soit avril et novembre, ont nécessairement des valeurs plus faibles que les autres mois, probablement dû à la basse température typiquement associée à ces périodes. De plus, seulement la moitié de ces mois sont compris dans la saison. Les mois de juin et juillet sont les plus populaires, selon la médiane et la moyenne. De plus, tous les jours du mois de juillet voient au moins 19 703 déplacements, soit plus du double des déplacements minimums quotidiens des autres mois. En outre, juin et juillet sont également les mois généralement les plus chauds, avec le mois d'août.

Tableau 3.1 : Statistiques sommaires des flux selon le mois

|                  | <b>Déplacements quotidiens (flux sortants)</b> |                |                |             |
|------------------|--|----------------|----------------|-------------|
|                  | <b>Min.</b>                                    | <b>Médiane</b> | <b>Moyenne</b> | <b>Max.</b> |
| <b>Avril</b>     | 2 530  | 9 030          | 9 150          | 18 510      |
| <b>Mai</b>       | 4 203  | 18 781         | 17 740         | 29 467      |
| <b>Juin</b>      | 8 994  | 26 092         | 24 207         | 31 205      |
| <b>Juillet</b>   | 19 703   | 24 928         | 25 569         | 29 591      |
| <b>Août</b>      | 5 146  | 23 853         | 22 934         | 28 142      |
| <b>Septembre</b> | 9 138  | 21 698         | 21 130         | 26 755      |
| <b>Octobre</b>   | 4 327  | 14 055         | 14 364         | 20 653      |
| <b>Novembre</b>  | 6 546  | 10 145         | 10 797         | 15 420      |
| <b>Année</b>     | 2 530  | 21 168         | 19 383         | 31 205      |

La figure 3.1 permet d'illustrer cette variabilité quotidienne selon le mois et surtout la concentration qui existe selon le moment de l'année. Il est possible d'y constater les faibles déplacements effectués lors des épaules de la saison, soit avril et novembre, et même en mai et en octobre, alors que les quatre autres mois en génèrent davantage de manière quotidienne. La figure permet également de voir l'ensemble des flux sortants aux stations pour l'année 2011.

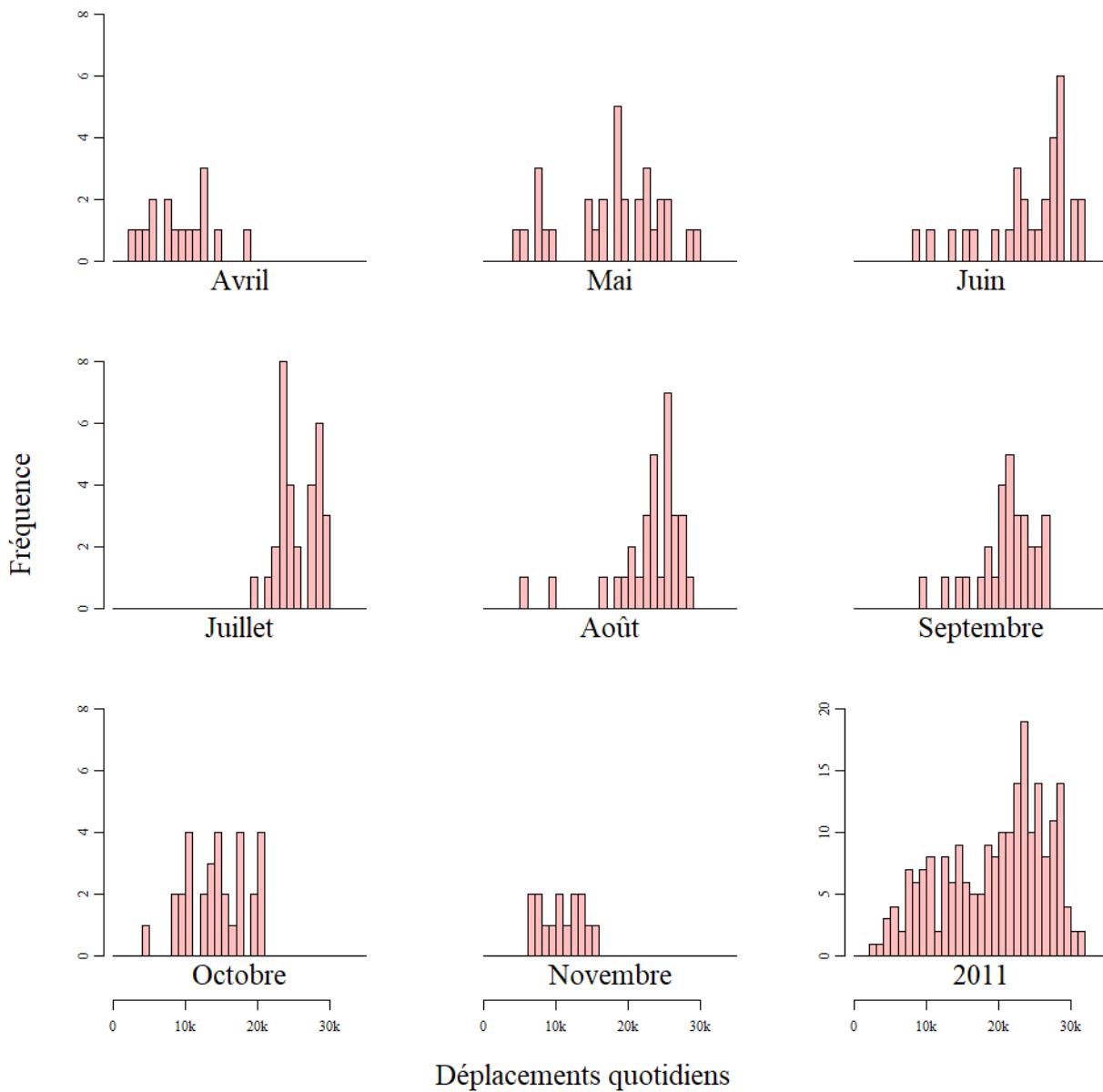


Figure 3.1 : Distribution fréquentielle des flux sortants quotidiens selon le mois

Dans le but de minimiser cette grande variabilité dans l'usage, il est question de trier les données. Pour ce faire, un test de la somme des rangs Wilcoxon est effectué, ce qui permet de déterminer si

les flux sont significativement différents selon la période. Ce test est non paramétrique, où l'hypothèse nulle est que deux populations possèdent les mêmes distributions et médianes, alors que l'alternative signifie qu'il y a un décalage entre deux distributions (Ford, 2017). L'hypothèse nulle est rejetée si la probabilité est plus petite que 0,05.

Les résultats de ce test sont présentés au tableau 3.2 et démontrent les probabilités que les médianes de chaque mois diffèrent des autres. Selon ces résultats, deux mois sont particulièrement similaires, soit juin et juillet, avec une valeur p de 0,77. Deux autres paires de mois sont relativement similaires, soit les mois de juin et d'août (0,09) et les mois d'avril et de novembre (0,22). Les flux des autres mois sont distribués de manière différente.

Tableau 3.2 : Résultats du test de la somme des rangs Wilcoxon des flux sortants selon le mois

|                | <b>Avril</b>  | <b>Mai</b> | <b>Juin</b>   | <b>Juillet</b> | <b>Août</b> | <b>Sept.</b> | <b>Octobre</b> | <b>Nov.</b> |
|----------------|---------------|------------|---------------|----------------|-------------|--------------|----------------|-------------|
| <b>Avril</b>   | NA            | 0,0001     | 0,0000        | 0,0000         | 0,0000      | 0,0000       | 0,0006         | 0,2164      |
| <b>Mai</b>     | 0,0001        | NA         | 0,0001        | 0,0000         | 0,0004      | 0,0452       | 0,0166         | 0,0009      |
| <b>Juin</b>    | 0,0000        | 0,0001     | NA            | 0,7691         | 0,0847      | 0,0021       | 0,0000         | 0,0000      |
| <b>Juillet</b> | 0,0000        | 0,0000     | <b>0,7691</b> | NA             | 0,0235      | 0,0000       | 0,0000         | 0,0000      |
| <b>Août</b>    | 0,0000        | 0,0004     | <b>0,0847</b> | 0,0235         | NA          | 0,0202       | 0,0000         | 0,0000      |
| <b>Sept.</b>   | 0,0000        | 0,0452     | 0,0021        | 0,0000         | 0,0202      | NA           | 0,0000         | 0,0000      |
| <b>Octobre</b> | 0,0006        | 0,0166     | 0,0000        | 0,0000         | 0,0000      | 0,0000       | NA             | 0,0038      |
| <b>Nov.</b>    | <b>0,2164</b> | 0,0009     | 0,0000        | 0,0000         | 0,0000      | 0,0000       | 0,0038         | NA          |

Selon ces résultats, seuls les déplacements des mois de juin et de juillet sont gardés alors que les autres mois sont retirés de la base de données. Ceci permet de limiter l'impact de la variabilité quotidienne selon le mois. Ces deux mois sont également ceux ayant le plus de flux sortants quotidiens selon la moyenne, la médiane, et le nombre maximal absolu, comme présenté précédemment au tableau 3.1, et sont donc critiques dans la modélisation.

Ensuite, les flux selon les jours de la semaine sont étudiés. Le tableau 3.3 affiche quelques statistiques sommaires des flux selon le jour de la semaine, où ceux effectués lors d'un jour férié sont séparés. Le code de couleur utilisé est le même que pour le tableau 3.1. Selon ce tableau, les jours de la semaine semblent générer davantage de déplacements que la fin de semaine et les jours fériés, et le vendredi semble la journée la plus achalandée pour les deux mois à l'étude.

Tableau 3.3 : Statistiques sommaires des flux selon le jour de la semaine (juin et juillet)

|                        | <b>Déplacements quotidiens (flux sortants)</b> |                |                |             |
|------------------------|--|----------------|----------------|-------------|
|                        | <b>Min.</b>                                    | <b>Médiane</b> | <b>Moyenne</b> | <b>Max.</b> |
| <b>Lundi</b>           | 16 718   | 25 147         | 24 754         | 29 501      |
| <b>Mardi</b>           | 23 174   | 28 153         | 25 537         | 28 620      |
| <b>Mercredi</b>        | 23 766   | 27 697         | 26 883         | 30 186      |
| <b>Jeudi</b>           | 15 568   | 25 907         | 25 355         | 30 958      |
| <b>Vendredi</b>        | 25 380   | 28 593         | 28 606         | 31 205      |
| <b>Samedi</b>          | 13 170   | 24 761         | 23 525         | 27 973      |
| <b>Dimanche</b>        | 10 008   | 22 079         | 20 592         | 23 904      |
| <b>Férié</b>           | 8 994  | 16 546         | 16 546         | 24 098      |
| <b>Juin et juillet</b> | 8 994  | 25 860         | 24 899         | 31 205      |

La figure 3.2 permet quant à elle de visualiser cette distribution des flux. Il est possible d'y voir notamment une grande constance dans les déplacements le mardi, c'est-à-dire peu de variabilité dans la fréquence du nombre de déplacements quotidiens. Outre ceci, les distributions sont similaires le lundi, jeudi et samedi, alors que celle des déplacements lors de jours fériés semble distincte. Il est également possible d'apprécier la distribution pour tous les déplacements effectués lors des mois de juin et juillet en 2011.

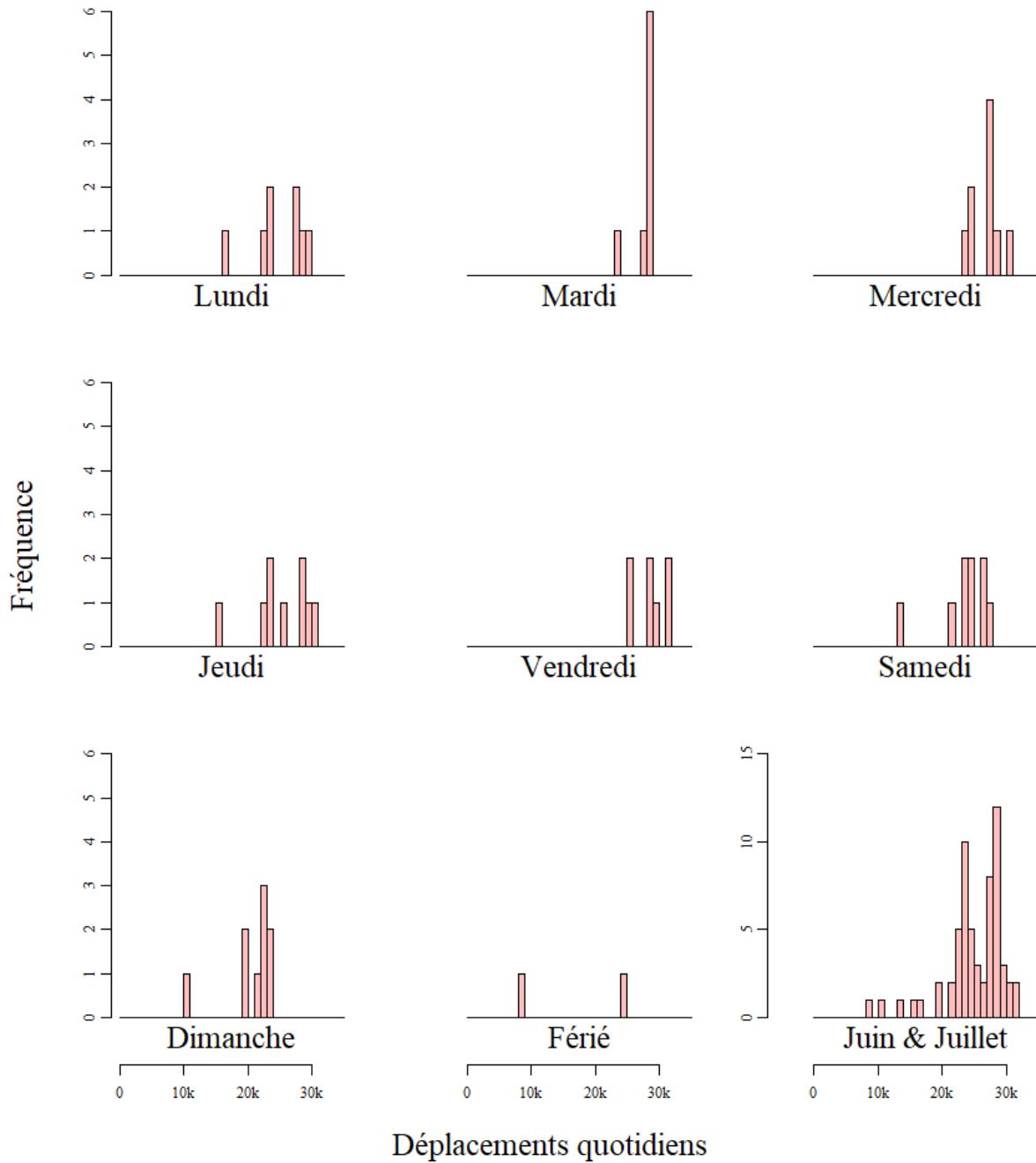


Figure 3.2 : Distribution fréquentielle des flux quotidiens selon le jour de la semaine (juin et juillet)

De la même manière que pour les flux selon le mois, un test de la somme des rangs Wilcoxon est effectué pour déterminer si les distributions des déplacements selon le jour de la semaine sont significativement différentes. Le tableau 3.4 présente les résultats de ce test. Il est possible d'y constater une grande similarité entre tous les jours de la semaine, avec une valeur p variant de 0,07

(lundi et vendredi) à 0,74 (jeudi et lundi ; jeudi et mardi). Le samedi est également semblable à certains jours de la semaine dans une moindre mesure. Fait étonnant, les jours fériés présentent des similitudes avec tous les jours de la semaine. Bref, les valeurs p confirment la grande variabilité qui existe dans la distribution des déplacements et ce, même pour les jours ouvrables. Ceci pourrait être dû en partie aux conditions météorologiques. Or, selon les statistiques et les distributions présentées précédemment, seuls les jours ouvrables sont retenus, soit du lundi au vendredi, ce qui concorde d'ailleurs avec la variabilité retrouvée dans la revue de littérature.

Tableau 3.4 : Résultats du test de la somme des rangs Wilcoxon des flux sortants selon le jour

|               | <b>Lundi</b>  | <b>Mardi</b>  | <b>Merc.</b>  | <b>Jeudi</b>  | <b>Vend.</b>  | <b>Samedi</b> | <b>Dim.</b>   | <b>Férié</b> |
|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|--------------|
| <b>Lundi</b>  | NA            | 0,1949        | 0,2359        | 0,7430        | 0,0721        | 0,5414        | 0,0592        | 0,4000       |
| <b>Mardi</b>  | <b>0,1949</b> | NA            | 0,3704        | 0,7430        | 0,2319        | 0,0037        | 0,0003        | 0,0889       |
| <b>Merc.</b>  | <b>0,2359</b> | <b>0,3704</b> | NA            | 0,7304        | 0,1416        | 0,0770        | 0,0001        | 0,1455       |
| <b>Jeudi</b>  | <b>0,7430</b> | <b>0,7430</b> | <b>0,7304</b> | NA            | 0,1416        | 0,3401        | 0,0188        | 0,3273       |
| <b>Vend.</b>  | <b>0,0721</b> | <b>0,2319</b> | <b>0,1416</b> | <b>0,1416</b> | NA            | 0,0052        | 0,0002        | 0,0556       |
| <b>Samedi</b> | <b>0,5414</b> | 0,0037        | <b>0,0770</b> | <b>0,3401</b> | 0,0052        | NA            | 0,0244        | 0,3273       |
| <b>Dim.</b>   | <b>0,0592</b> | 0,0003        | 0,0001        | 0,0188        | 0,0002        | 0,0244        | NA            | 1,0000       |
| <b>Férié</b>  | <b>0,4000</b> | <b>0,0889</b> | <b>0,1455</b> | <b>0,3273</b> | <b>0,0556</b> | <b>0,3273</b> | <b>1,0000</b> | NA           |

La base de données qui résulte de ce tri ne comprend donc que les flux aux stations effectués du lundi au vendredi pendant les mois de juin et de juillet pendant l'année 2011, soit 41 jours. Finalement, une dernière manipulation consiste à observer la présence des stations selon le nombre de jours pendant cette période, puisque certaines stations peuvent être installées de manière temporaire ou déplacées pendant la saison. Ainsi, toute station n'ayant pas d'enregistrement pendant au moins 80 % de ces 41 jours est retirée de la base de données. Ce seuil est arbitraire : un seuil de 100 % aurait éliminé 27 stations, soit 6,7 % de l'ensemble des stations, ce qui paraissait trop élevé. De plus, il n'est pas nécessairement souhaité de n'avoir que des stations actives à tous les jours, mais bien d'éliminer celles qui ne sont pas régulières. De ce fait, cette étape élimine 11 stations, soit 2,7 % du réseau. Après ce tri, la base de données restante comporte 393 stations avec un total de 1 072 874 déplacements effectués pendant la période étudiée, ou 26 168 déplacements quotidiens en moyenne. Ces déplacements quotidiens moyens constituent la variable à expliquer dans le modèle de génération qui est présenté dans la section 3.2. Avant de passer à la modélisation, une brève analyse spatiale est présentée à la section suivante.

### 3.1.2 Analyse spatiale

Les déplacements possèdent également une variabilité selon la localisation de la station. Afin de représenter les déplacements aux stations spatialement, la moyenne des flux sortants à chaque station, ainsi que la capacité, est calculée pour la période analysée. Ces informations permettent de calculer un taux sortant, soit le nombre moyen de déplacements sortants par station sur sa capacité. Un algorithme de k-moyennes est ensuite utilisé avec ce taux. Cet algorithme est une technique de classement (*clustering*) de données non supervisée et nécessite en tant qu'entrant le nombre de groupes désiré. L'algorithme tente de minimiser la distance entre chaque valeur et le centre de chacun des groupes grâce à de multiples itérations (Vogel, M. et al., 2014). Ainsi, ce regroupement en cinq groupes est effectué grâce au logiciel RStudio selon l'algorithme développé par Hartigan et Wong (1979), où la distance entre les valeurs est euclidienne avec 10 itérations.

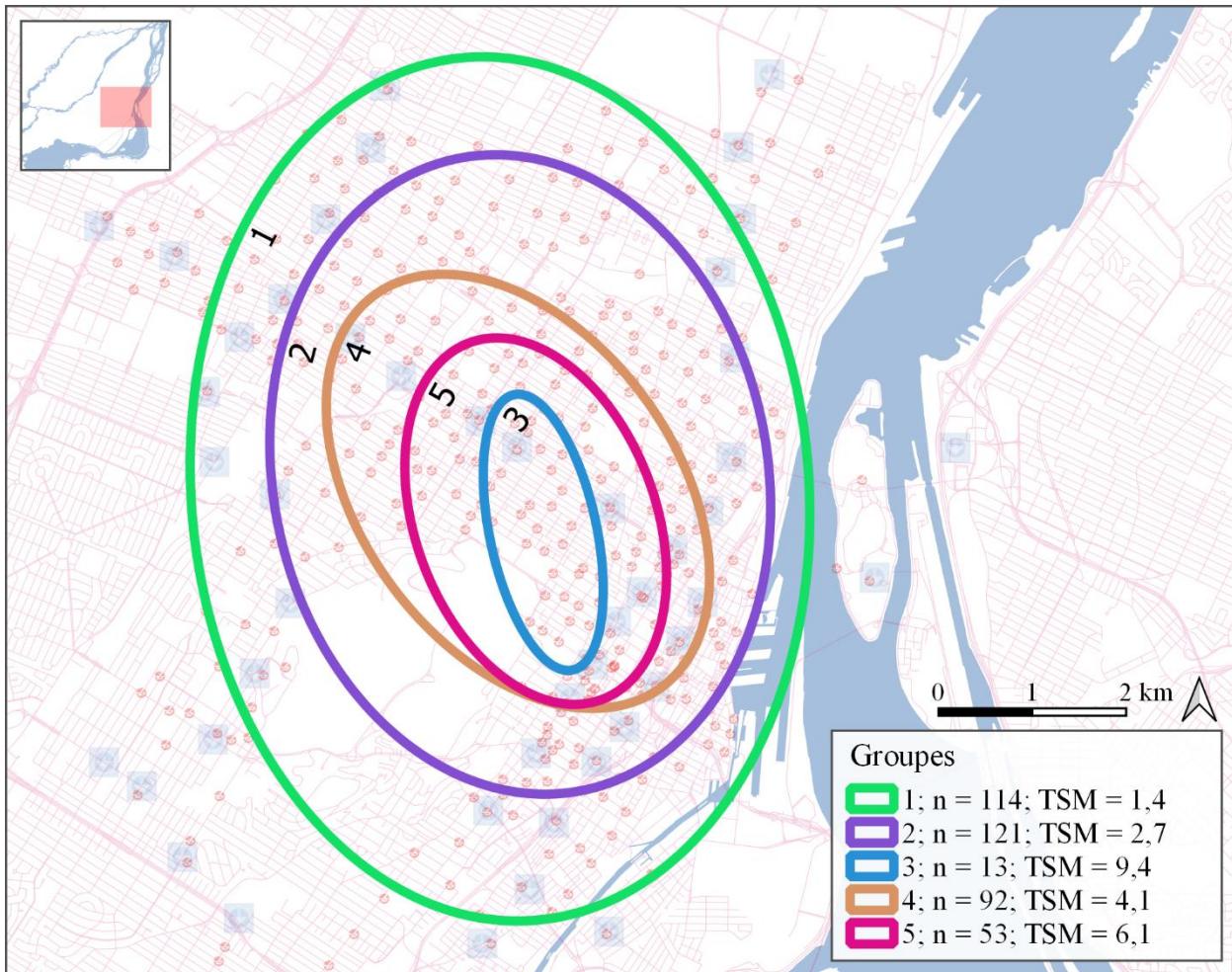


Figure 3.3 : Ellipses de dispersion selon les groupes de stations classées par taux sortant moyen

Les groupes de stations sont représentés grâce à des ellipses de dispersion à la figure 3.3. La figure permet d'observer les tendances de ce groupement basé seulement sur le taux sortant moyen (TSM) des flux aux stations, où les taux les plus importants sont très concentrés au centre, soit les groupes 3 et 5 avec des taux moyens de 9,4 et 6,1. À l'inverse, les groupes avec des taux plus faibles contiennent davantage de stations. Effectivement, les groupes 1 et 2, avec des taux faibles de 1,4 et 2,7, comprennent ensemble 235 stations, soit 60 % de la totalité.

Cette visualisation simple permet de voir que l'utilisation du système BIXI varie dans l'espace. En ce sens, certains éléments spatiaux ont nécessairement un effet sur cette utilisation, ce que la modélisation doit faire ressortir dans la section suivante.

## 3.2 Modélisation de la génération des déplacements

Les modèles utilisés souhaitent expliquer le nombre de déplacements moyens quotidiens produits et attirés par chaque station en 2011. La modélisation emprunte donc la première étape de la PSC, soit la génération des déplacements. Comme vu à la section 2.5, il est question de remplir les marges de la matrice O-D, soit la totalité des déplacements moyens quotidiens produits et attirés par chacune des stations. La technique utilisée afin de reproduire ces flux aux stations est la régression linéaire multiple et est présentée en détail ci-après. Les informations détaillées se basent sur celles de Bonnel (2002), McNally (2007), et Ortuzar et Willumsen (2011) ainsi que des apprentissages de Tran (2016).

### 3.2.1 La régression linéaire

La régression linéaire est la technique privilégiée dans le cas présent afin de modéliser la génération et l'attraction de déplacements aux stations du réseau BIXI, car celle-ci est réputée pour sa robustesse. La régression linéaire tente d'expliquer l'influence d'une variable indépendante  $X_i$ , ou plusieurs dans le cas d'une régression multiple, sur une autre variable  $Y_i$ , donc dépendante. Or, il est impossible de représenter la distribution de chaque variable dépendante pour chaque variable indépendante, ce qui signifie que le modèle de régression se base sur des hypothèses.

- Les distributions des variables expliquées, pour chaque variable indépendante, ont la même variance, soit le carré de l'écart-type ( $\sigma^2$ ) et sont donc homogènes (notion d'homoscédasticité) ;

- Les différentes variables dépendantes sont statistiquement indépendantes les unes des autres, c'est-à-dire qu'elles n'ont pas d'influence entre elles ;
- De plus, dans le cas d'une régression multiple, les variables dépendantes doivent présenter peu de corrélation, afin de minimiser la problématique de la multicolinéarité ;
- Chaque variable indépendante possède une influence linéaire avec celle qui est dépendante, de sorte que la moyenne de chaque distribution est la vraie droite de régression, donc  $E\left(\frac{Y_i}{X_i}\right) = \mu_i = \alpha + \beta X_i$ ;
- Le paramètre  $\beta$  estimé  $\hat{\beta}$ , suit une distribution normale, ce qui est généralement le cas lorsque le nombre d'observations de l'échantillon est supérieur à 30.

Dans le cas des modèles de production et d'attraction de déplacements en BIXI, plusieurs variables explicatives (indépendantes) sont considérées, ce qui mène à une régression linéaire multiple. La formulation de cette dernière est présentée à l'équation 3.1, où  $\hat{Y}$  est la valeur expliquée prédite par le modèle,  $\alpha$  est une constante,  $\beta_i$  le coefficient de régression pour chaque variable explicative  $X_i$ , alors que  $e$  constitue les résidus permettant d'obtenir la vraie droite de régression.

$$\hat{Y} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + e \quad 3.1$$

Les équations 3.2 et 3.3 permettent d'estimer les paramètres  $\hat{\alpha}$  et  $\hat{\beta}$ , où  $\bar{X}$  est la moyenne des  $X_i$  et  $\bar{Y}$  la moyenne des  $Y_i$ . L'ensemble de ces manipulations sont effectués avec le logiciel RStudio.

$$\hat{\beta} = \frac{\sum_i ((X_i - \bar{X}) * (Y_i - \bar{Y}))}{\sum_i (X_i - \bar{X})^2} \quad 3.2$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta} \bar{X} \quad 3.3$$

### 3.2.1.1 Variable dépendante

Ainsi, une régression linéaire est effectuée pour chaque station à l'Origine et à la Destination. La variable dépendante,  $Y$ , est le nombre de déplacements moyens quotidiens produits et attirés par chacune des stations en 2011, provenant de la base de données triée présentée à la section 3.1.1. Il est donc question d'expliquer ces déplacements en fonction de plusieurs variables indépendantes qui permettent de caractériser les stations selon leur emplacement sur le territoire.

### 3.2.1.2 Variables indépendantes

Les modèles de régression permettent d'estimer la production et l'attraction des déplacements aux stations en fonction de plusieurs variables indépendantes,  $X_i$ . Celles-ci sont le plus souvent utilisées pour traduire certains éléments existants à l'Origine ou à la Destination. Elles agissent à titre d'indicateurs permettant de traduire la temporalité, l'espace ou encore des facteurs qui risquent d'avoir un impact sur la probabilité d'effectuer un déplacement. Les variables utilisées dans ce cas sont regroupées en quatre catégories : sociodémographique, usage du sol et environnement bâti, offre de mobilité et puis offre BIXI. Les indicateurs retenus ont été déterminés en fonction de leur popularité dans la littérature. Alors que certains éléments sont ponctuels, comme l'altitude ou encore la capacité de la station, d'autres sont plutôt spatiaux, tels les emplois ou le nombre de stations de métro à proximité. Les études sur le sujet suggèrent ainsi de considérer les variables spatiales grâce au calcul d'une zone tampon autour de chaque station de vélopartage. La distance de ces zones varie de 200 m à 800 m dans la revue de littérature. Des tests avec des zones tampons de différentes tailles ont été effectués et il a été décidé de présenter les informations et résultats relatifs aux zones tampons de 400 m. Cette distance agit à titre d'un minimum quant à l'accessibilité piétonne et est équivalente à cinq minutes de marche (Active healthy communities, 2017). Ces zones tampons sont calculées autour de chaque station BIXI de manière euclidienne, qui, en guise de référence, sont présentées à l'annexe A. Pour ce faire, la base de données des flux entrants et sortants aux stations BIXI est importée dans un système d'information géographique (SIG), ce qui permet par la suite de la combiner avec d'autres jeux de données. Un des enjeux importants est d'avoir des données dont la date est similaire, afin d'assurer une certaine synchronicité dans celles-ci pour améliorer la qualité des modèles. Ce faisant, l'année de production des différentes bases de données est également mentionnée.

Tel que mentionné plus tôt, certaines données proviennent du recensement canadien de 2011. Les données dans ce cas sont calculées au niveau du secteur de recensement (SR), qui contiennent de 2 500 à 8 000 résidents (Statistique Canada, 2015). Cette méthode lie les informations d'un SR donné à chaque station contenue dans celui-ci. La représentation cartographique des SR est disponible à l'annexe B. Cette base de données permet également de calculer le nombre d'emplois par SR, qui est calculé grâce à l'endroit déclaré du lieu d'emploi habituel, excluant le télétravail. Quatre variables sont donc calculées grâce à la base de données du recensement : la population, la densité de population par kilomètre carré, l'âge médian et le nombre d'emplois.

Les données sur l'usage du sol et l'environnement bâti proviennent de diverses sources et permettent de calculer trois variables. Les informations concernant les lieux d'éducation postsecondaire proviennent du site web des données ouvertes du gouvernement du Québec (<https://www.donneesquebec.ca/fr>) Alors que cette information est relativement récente, soit 2017, les 26 lieux d'établissements supérieurs recensés dans la région montréalaise ont été observés afin de s'assurer qu'ils étaient tous présents en 2011. Les aires de parc proviennent de l'usage du sol obtenu des produits cartographiques de la Communauté métropolitaine de Montréal (CMM) et celui-ci date de 2012 (Communauté métropolitaine de Montréal, 2018). L'aire totale du parc est considérée lorsqu'une zone tampon d'une station entre en intersection avec elle. L'élévation des stations provient de Ressources naturelles Canada (2014).

Quant aux données sur les infrastructures de transport collectif, elles proviennent des données *General Transit Feed Specification* (GTFS) de la Société de transport de Montréal (STM) et du réseau de transport métropolitain (RTM), l'agence de transport métropolitaine, et datent chacune de 2011. Les infrastructures cyclables proviennent des données ouvertes de la Ville de Montréal (<http://donnees.ville.montreal.qc.ca>) et ont été produites en 2013. Les informations retirées de ces bases de données sont la présence d'une station de métro, l'occurrence d'une gare de train de banlieue, la fréquence quotidienne de passage des autobus aux arrêts ainsi que la somme de la longueur des infrastructures cyclables.

Le dernier groupe comprend quatre variables permettant de décrire le réseau BIXI qui sont recueillies à même la base de données principale : la capacité moyenne de la station, le nombre de stations avoisinantes, la somme de la capacité moyenne de ces stations avoisinantes et la distance de la station par rapport au centre moyen du réseau. En guise de rappel, la capacité des stations utilisée est une moyenne due à sa fluctuation temporelle.

Ainsi, les statistiques sommaires des 15 variables explicatives et deux variables à expliquer sont présentées dans le tableau 3.5. De plus, il agit à titre de rappel afin de savoir s'il s'agit d'une variable ponctuelle ou spatiale, calculée avec la zone tampon, la source et la date des données ainsi que le type de variable. Les variables binaires sont utilisées lorsque peu de variation est présente dans les données, et celles-ci prennent donc la forme de 1 lorsqu'il y a au moins une observation et 0 quand il n'y en a pas.

Tableau 3.5 : Description des variables indépendantes modélisées (juin et juillet 2011)

|                             | Variable                                  | Définition  | Unité                | Niveau      | Données                 | An   |
|-----------------------------|---|---|----------------------|-------------|-------------------------|------|
| <b>Dépendante</b>           | <b>Sortants</b>                           | Moyenne quotidienne de flux sortants                                    | —                    | Station     | BIXI                    | 2011 |
|                             | <b>Entrants</b>                           | Moyenne quotidienne de flux entrants                                    | —                    | Station     | BIXI                    | 2011 |
| <b>Sociodémo.</b>           | <b>Population</b>                         | Résidents par aire  | —                    | SR          | Statistique Canada      | 2011 |
|                             | <b>Densité de population</b>              | Résidents par aire  | Pers/km <sup>2</sup> | SR          | Statistique Canada      | 2011 |
|                             | <b>Âge médian</b>                         | 50 <sup>e</sup> percentile de l'âge des résidents par aire              | —                    | SR          | Statistique Canada      | 2011 |
|                             | <b>Emplois</b>                            | Emplois par aire  | —                    | SR          | Statistique Canada      | 2011 |
| <b>Usage du sol et env.</b> | <b>Postsecond.</b>                        | Présence d'établissements postsecondaires                               | Binaire              | Zone tampon | Données ouvertes Québec | 2017 |
|                             | <b>Parc</b>                               | Total de l'usage de parc par aire                                       | m <sup>2</sup>       | Zone tampon | CMM                     | 2012 |
|                             | <b>Élévation</b>                          | Altitude de la station  | m                    | Station     | Ress. nat. Canada       | —    |
| <b>Offre de transport</b>   | <b>Station de métro</b>                   | Présence de station de métro par aire                                   | Binaire              | Zone tampon | STM (GTFS)              | 2011 |
|                             | <b>Gare</b>                               | Présence de gare par aire   | Binaire              | Zone tampon | RTM (GTFS)              | 2011 |
|                             | <b>Fréquence de bus</b>                   | Total quotidien de passage de bus aux arrêts par aire                   | —                    | Zone tampon | STM (GTFS)              | 2011 |
|                             | <b>Infra. cyclable</b>                    | Somme de la distance par aire   | m                    | Zone tampon | Ville de Montréal       | 2013 |
| <b>Offre BIXI</b>           | <b>Capacité</b>                           | Moyenne quotidienne de points d'ancre par station                       | —                    | Station     | BIXI                    | 2011 |
|                             | <b>Stations avoisinantes</b>              | Nombre de stations par aire   | —                    | Zone tampon | BIXI                    | 2011 |
|                             | <b>Capacité des stations avoisinantes</b> | Somme de la moyenne quotidienne de points d'ancre des stations par aire | —                    | Zone tampon | BIXI                    | 2011 |
|                             | <b>Distance au centre</b>                 | Distance au centre moyen du réseau BIXI                                 | m                    | Station     | BIXI                    | 2011 |

Les deux figures ci-dessous permettent d'apprécier l'influence de chacune des variables explicatives en fonction des taux des flux sortants et entrants moyens. Il est possible de voir des tendances similaires dans les deux cas, notamment celle de la densité de population, les infrastructures cyclables, la fréquence de passages des autobus aux arrêts, ainsi que l'ensemble des variables ayant trait à l'offre BIXI. Il est également intéressant de noter l'influence inverse que possède l'élévation de la station sur les flux sortants par rapport aux flux entrants. La modélisation est utilisée afin d'évaluer lesquelles de ces variables possèdent une influence significative sur les déplacements aux stations.

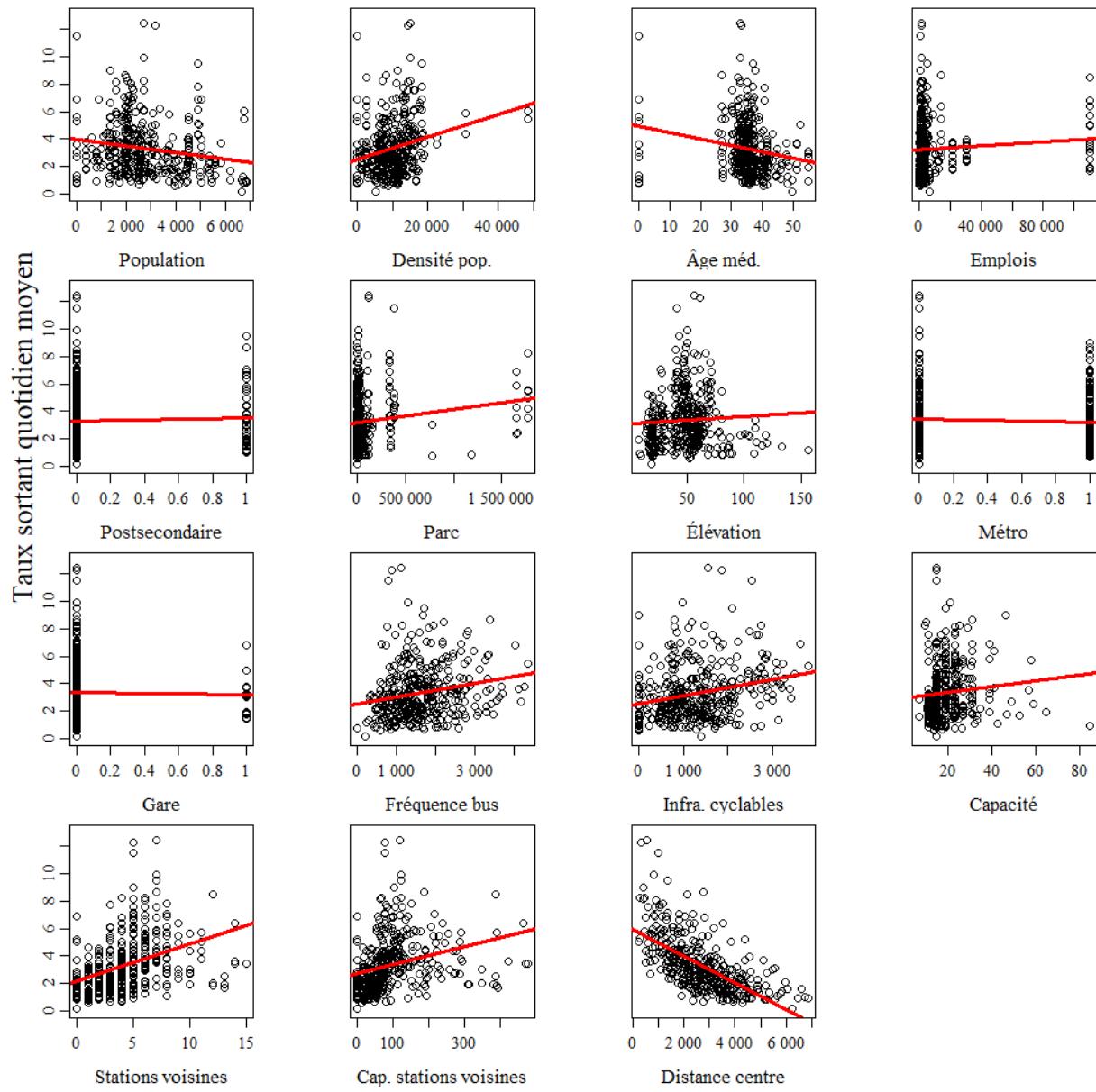


Figure 3.4 : Relation entre le taux sortant quotidien moyen par station et les variables indépendantes

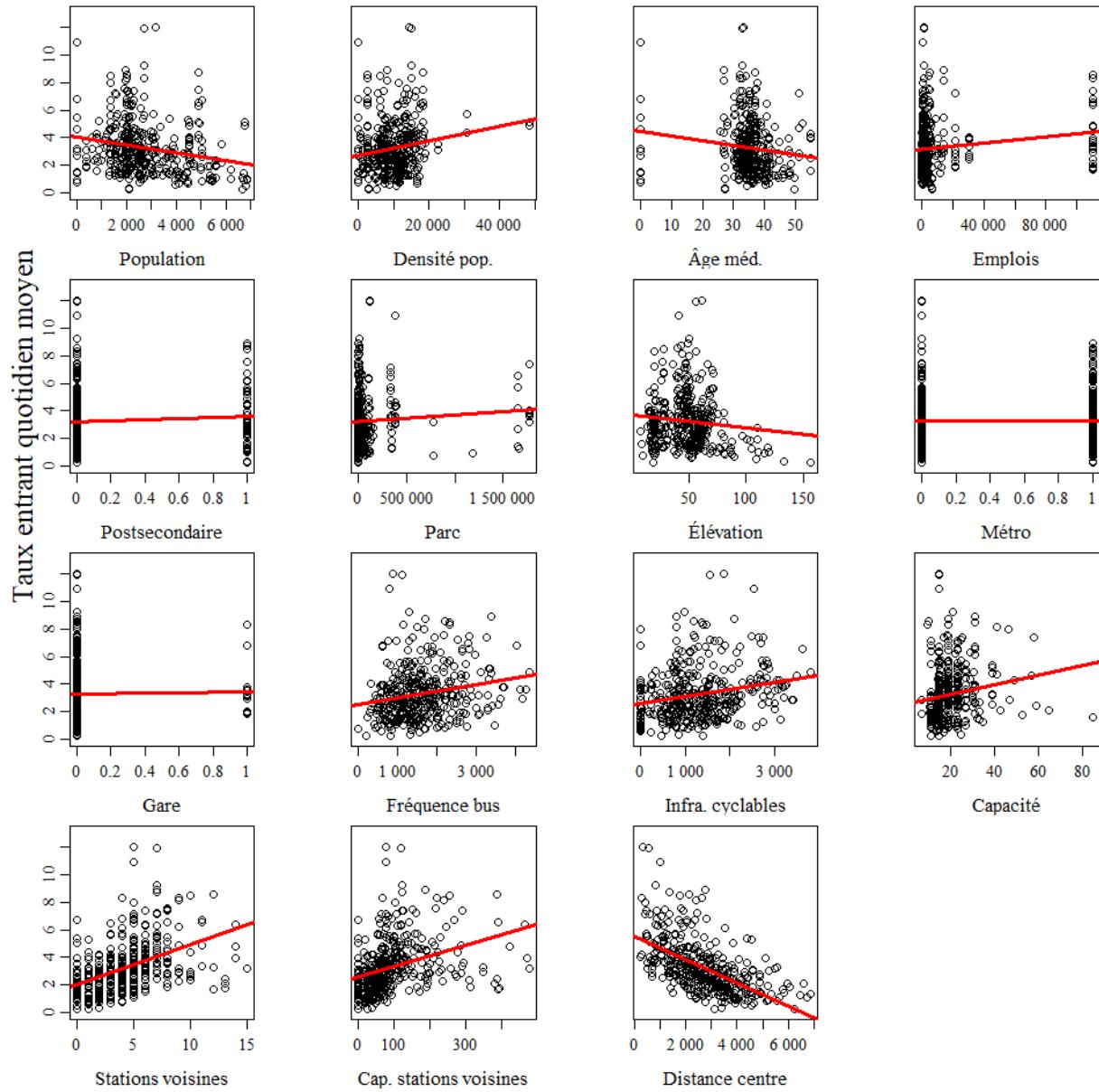


Figure 3.5 : Relation entre le taux entrant quotidien moyen par station et les variables indépendantes

Finalement, il semble pertinent de mentionner que plusieurs autres variables ont été considérées pour l'étape de modélisation. Effectivement, la revue de littérature fait état d'éléments intéressants comme le revenu de la population ou encore la présence de commerces, pour ne mentionner que ceux-ci. Des variables permettant de considérer ces éléments ont été retirées en amont, car trop corrélées avec les variables exposées. Certaines de ces dernières présentent également une corrélation importante, tel qu'illustiqué à la figure 3.6. Cette matrice présente la corrélation Pearson

qui varie de -1 à 1, signifiant une corrélation très forte entre les variables, alors qu'un 0 traduit une corrélation inexistante entre les données. Généralement, une corrélation supérieure à 0,4 (ou inférieure à -0,4) est considérée comme importante.

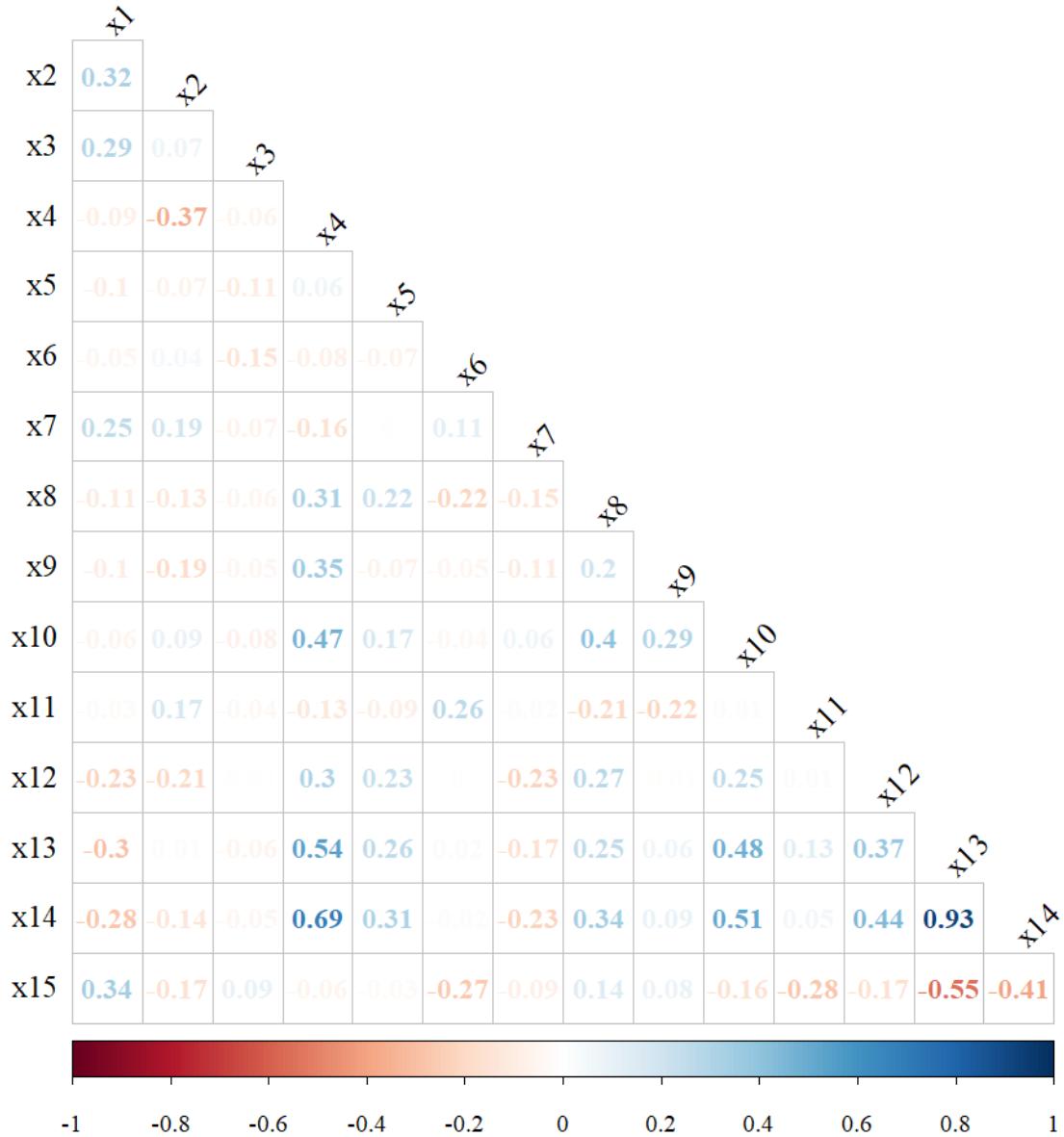


Figure 3.6 : Matrice de corrélation entre les variables indépendantes

*x1 : population ; x2: densité de population ; x3 : âge médian ; x4 : emplois ; x5 : postsecondaire ; x6 : parc ; x7 : élévation ; x8 : station de métro ; x9 : gare ; x10 : fréquence de bus ; x11 : infrastructures cyclables ; x12 : capacité ; x13 : stations avoisinantes ; x14 : capacité des stations avoisinantes ; x15 : distance au centre*

Il est possible de constater une corrélation importante entre la variable emplois (x4) et la fréquence de passage des autobus (x10), le nombre de stations avoisinantes (x13) et la capacité de celles-ci (x14). La fréquence de passage de bus possède également une corrélation élevée avec les deux dernières variables, soit de 0,48 et 0,51. L'ensemble des variables concernant l'offre BIXI est également corrélé à divers niveaux, variant de -0,55 à 0,93. Généralement, ces variables ne devraient pas être incluses simultanément dans une régression linéaire, car elles ne sont pas indépendantes. Cependant, il a été décidé de les garder dans le cas présent afin de déterminer l'ampleur de leur impact sur les variables à expliquer. Une comparaison de modèle avec et sans les variables corrélées est toute de même réalisée dans la section des résultats afin de voir si les coefficients présentent des différences importantes.

### 3.2.2 Développement des modèles

Les 15 variables explicatives sont utilisées dans des modèles tentant de décrire les flux entrants et sortants quotidiens moyens aux stations. La formulation utilisée pour la modélisation est une régression linéaire multiple, démontrée à l'équation 3.1.

Une fois la régression estimée, le coefficient de détermination  $R^2$  est un bon indicateur pour évaluer le modèle. Il représente le ratio de la variance expliquée sur la variance totale, tel que présenté aux équations 3.4, 3.5 et 3.6, où  $Y_i$  représente les variables observées,  $\bar{Y}$  la moyenne de ces dernières et  $\hat{Y}_i$  les valeurs prédictes par le modèle. Étant une proportion, le  $R^2$  se situe entre 0 et 1 : plus le coefficient est élevé, plus la variance est expliquée par le modèle, donc meilleur est ce dernier.

$$Y_i - \bar{Y} = (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \quad \text{écart total} \quad \text{écart expliqué} \quad \text{écart non expliqué}$$

3.4

$$\sum_i (Y_i - \bar{Y})^2 = \sum_i (\hat{Y}_i - \bar{Y})^2 + \sum_i (Y_i - \hat{Y}_i)^2 \quad 3.5$$

variance totale      variance expliquée      variance non expliquée

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \quad 3.6$$

Or, dans le cas d'une régression multiple, plus il y a de variables explicatives dans le modèle, plus le coefficient de détermination augmente. Conséquemment, il est pertinent de calculer également

le  $R^2$  ajusté,  $\overline{R^2}$ , qui nécessite la taille de l'échantillon  $n$  et le nombre de variables explicatives  $k$ , tel que présenté à l'équation 3.7.

$$\overline{R^2} = \left( R^2 - \frac{k}{n-1} \right) * \left( \frac{n-1}{n-k} \right) \quad 3.7$$

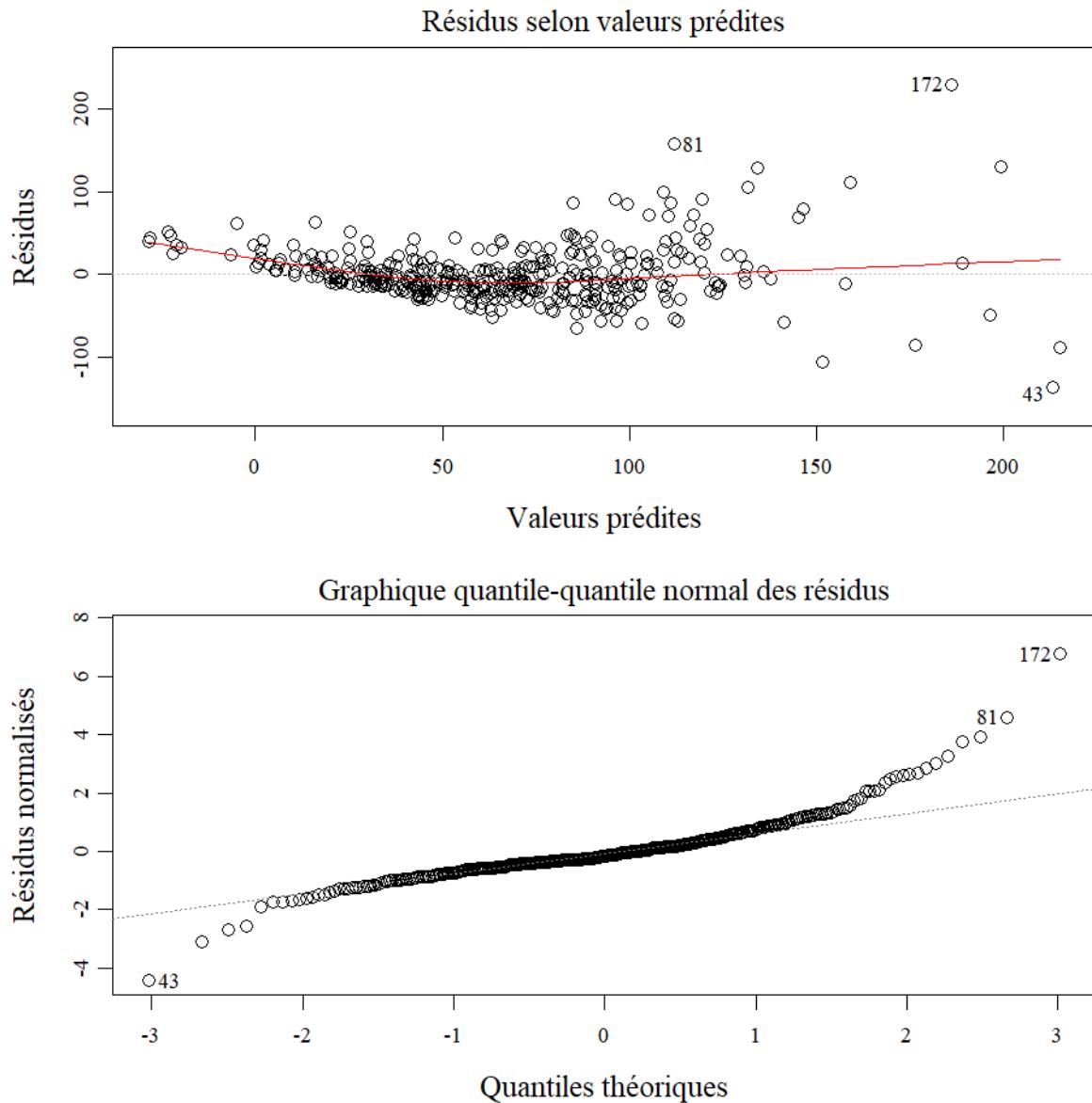


Figure 3.7 : Spatialisation des résidus du modèle des flux sortants quotidiens moyens

Les résultats des premiers tests donnent des résultats intéressants en se fiant aux coefficients de détermination. Effectivement, la régression linéaire multiple des flux sortants moyens quotidiens produit un  $R^2$  de 0,57 et cette valeur est à 0,58 dans le cas des flux entrants. Cependant, la

spatialisation des résidus des modèles démontre des tendances particulières. La figure 3.7 illustre le résultat du modèle selon les flux sortants à titre d'exemple. Les résidus, qui correspondent aux différences entre les valeurs estimées et observées, présentent une certaine concentration lorsque observés selon les valeurs prédictives. Ils ne semblent pas normalement distribués. De manière similaire, le graphique quantile-quantile permet de constater une droite qui s'éloigne de la diagonale parfaite, qui exprimerait une distribution aléatoire des résidus. En ce sens, la modélisation peut être améliorée.

Tableau 3.6 : Coefficients de détermination des diverses transformations de variables

| <b>Transformation</b>                 | <b>Flux moyens quotidiens</b> |                             |                      |                             |
|---------------------------------------|-------------------------------|-----------------------------|----------------------|-----------------------------|
|                                       | <b>Sortants</b>               |                             | <b>Entrants</b>      |                             |
|                                       | <b>R<sup>2</sup></b>          | <b>R<sup>2</sup> ajusté</b> | <b>R<sup>2</sup></b> | <b>R<sup>2</sup> ajusté</b> |
| <b>Linéaire (sans transformation)</b> | 0,57                          | 0,56                        | 0,58                 | 0,56                        |
| <b>Exponentielle</b>                  | 0,66                          | 0,65                        | 0,66                 | 0,64                        |
| <b>Quadratique</b>                    | 0,65                          | 0,63                        | 0,65                 | 0,63                        |
| <b>Réciproque</b>                     | 0,39                          | 0,36                        | 0,37                 | 0,34                        |
| <b>Logarithmique</b>                  | 0,63                          | 0,61                        | 0,59                 | 0,58                        |
| <b>Puissance</b>                      | 0,71                          | 0,70                        | 0,69                 | 0,68                        |

Ceci s'effectue grâce à une transformation des variables, qui permet de changer le type de relation qui existe entre elles et améliorer la linéarité entre la variable dépendante et celles qui sont explicatives. Plusieurs transformations ont été testées et les coefficients de détermination de chacune d'elles sont présentés au tableau 3.6. Tout type de transformation semble être bénéfique pour la modélisation sauf celle de type réciproque. De plus, les deux modèles fonctionnent de manière similaire concernant les flux sortants et entrants. En observant les coefficients de détermination, la transformation de type puissance des variables est celle qui fonctionne le mieux.

De plus, c'est également celle qui présente la spatialisation des résidus la plus intéressante, illustrée à la figure 3.8.

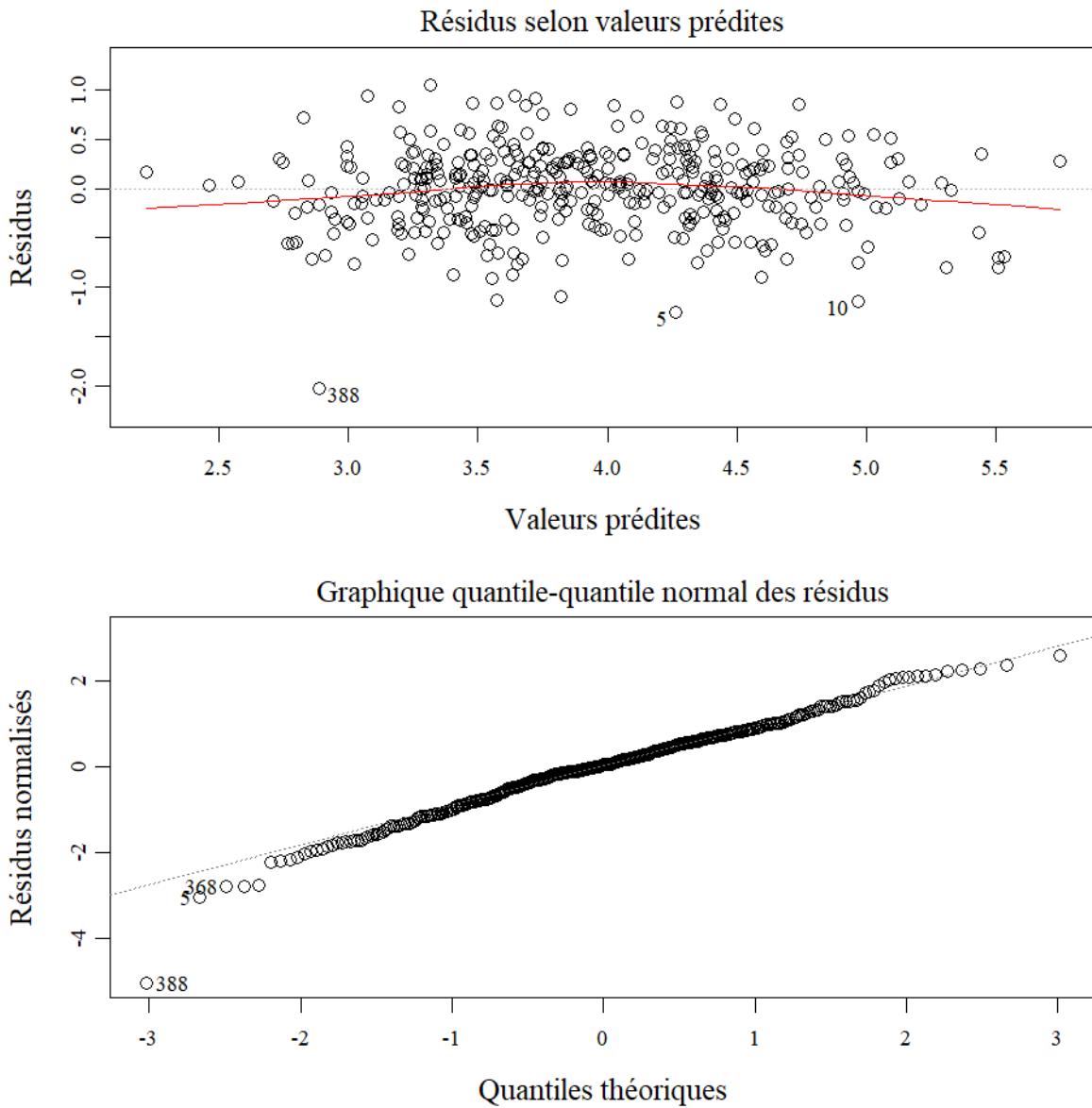


Figure 3.8 : Spatialisation des résidus du modèle des flux sortants quotidiens moyens avec transformation de type puissance

C'est donc la version puissance qui est retenue aux fins de modélisation, qui consiste en une transformation logarithmique de toutes les variables, formulée à l'équation 3.8. Lorsqu'une variable continue possède un minimum de 0, la transformation présentée à l'équation 3.9 est effectuée.

$$\log Y_i = \beta_0 + \beta_1 \log X_1 + \beta_2 \log X_2 + \dots + \beta_n \log X_n + e \quad 3.8$$

$$\log(X_i + 1) \quad 3.9$$

### 3.2.3 Résultats de la modélisation

Des modèles de génération et d'attraction sont donc estimés selon les variables indépendantes présentées précédemment autour de chaque station BIXI. Le tableau 3.7 présente les résultats de la modélisation avec seulement les variables qui ne sont pas corrélées, tel qu'illustré à la figure 3.6. Seules 11 variables sont présentes puisque quatre variables sont exclues : fréquence de bus (x10); le nombre de stations avoisinantes (x13); la capacité de ces stations avoisinantes (x14) et la distance au centre (x15). Les variables en gras sont significatives au moins à 95 % alors que l'intervalle de confiance est indiqué entre parenthèses. L'intercept est la valeur de la variable indépendante lorsque toutes les valeurs dépendantes sont à 0.

Tableau 3.7 : Résultats de la modélisation 2011 avec transformation de type puissance sans variables corrélées

|                                  | <b>Production</b>                             | <b>Attraction</b>              |
|----------------------------------|---|--------------------------------|
| <b>Intercept</b>                 | -2,30*** (-3,07, -1,53)                       | -1,23*** (-2,03, -0,44)        |
| <b>Population</b>                | <b>-0,55***</b> (-0,68, -0,43)                | <b>-0,56**</b> (-0,69, -0,43)  |
| <b>Densité de population</b>     | <b>0,53***</b> (0,44, 0,63)                   | <b>0,54***</b> (0,44, 0,63)    |
| <b>Âge médian</b>                | -0,17 (-0,39, 0,06)                           | -0,14 (-0,36, 0,09)            |
| <b>Emplois</b>                   | <b>0,21***</b> (0,16, 0,26)                   | <b>0,20***</b> (0,15, 0,25)    |
| <b>Postsecondaire</b>            | <b>-0,21**</b> (-0,38, -0,04)                 | <b>-0,24***</b> (-0,42, -0,07) |
| <b>Parc</b>                      | 0,001 (-0,01, 0,01)                           | 0,002 (-0,01, 0,01)            |
| <b>Élévation</b>                 | <b>0,27***</b> (0,16, 0,38)                   | -0,04 (-0,15, 0,08)            |
| <b>Métro</b>                     | <b>-0,21***</b> (-0,33, -0,10)                | <b>-0,21***</b> (-0,33, -0,09) |
| <b>Gare</b>                      | 0,06 (-0,23, 0,36)                            | 0,09 (-0,21, 0,39)             |
| <b>Infrastructures cyclables</b> | <b>0,05***</b> (0,02, 0,08)                   | <b>0,05***</b> (0,02, 0,08)    |
| <b>Capacité</b>                  | <b>1,21***</b> (1,05, 1,37)                   | <b>1,23***</b> (1,06, 1,40)    |
| Observations                     | 393   | 393                            |
| R <sup>2</sup>                   | 0,58  | 0,58                           |
| R <sup>2</sup> ajusté            | 0,57  | 0,56                           |
| Écart-type résiduel (df = 381)   | 0,50  | 0,51                           |
| Statistique F (df = 11 ; 381)    | 47,61***                                      | 46,99***                       |
| Significativité                  | <i>*p&lt;0,1 ; **p&lt;0,05 ; ***p&lt;0,01</i> |                                |

La plupart des variables qui ne sont pas corrélées entre elles ont une influence significative sur les flux entrants et sortants. Le tableau 3.8 présente quant à lui les coefficients des 15 variables malgré la corrélation entre certaines d'entre elles. Outre le changement de valeur des coefficients qui est généralement plus faible, ils ne présentent pas de changement quant à leur influence positive ou négative. Entre d'autres mots, le signe ne change pas. Le coefficient de détermination est plus grand puisqu'il y a plus de variables explicatives, ce qui est normal pour cet indicateur. Ce gain de 0,13 est intéressant pour l'ajout de seulement quatre variables. D'ailleurs l'ensemble des variables ajoutées dans cette version du modèle sont significatives, outre le nombre de stations avoisinantes et leur capacité sur les flux sortants.

Tableau 3.8 : Résultats de la modélisation 2011 avec transformation de type puissance

|   | <b>Production</b>                             | <b>Attraction</b>               |
|---|---|---------------------------------|
| <b>Intercept</b>                          | 3,68 *** (2,46, 4,90)                         | 4,09 *** (2,80, 5,39)           |
| <b>Population</b>                         | <b>-0,15 **</b> (-0,27, -0,02)                | <b>-0,17 **</b> (-0,30, -0,03)  |
| <b>Densité de population</b>              | <b>0,22 ***</b> (0,12, 0,32)                  | <b>0,23 ***</b> (0,13, 0,34)    |
| <b>Âge médian</b>                         | <b>-0,28 ***</b> (-0,47, -0,09)               | <b>-0,25 **</b> (-0,45, -0,06)  |
| <b>Emplois</b>                            | <b>0,10 ***</b> (0,05, 0,14)                  | <b>0,09 ***</b> (0,04, 0,13)    |
| <b>Postsecondaire</b>                     | -0,06 (-0,21, 0,08)                           | -0,10 (-0,26, 0,05)             |
| <b>Parc</b>                               | -0,01 * (-0,02, 0,001)                        | -0,01 (-0,02, 0,003)            |
| <b>Élévation</b>                          | 0,04 (-0,06, 0,15)                            | <b>-0,26 ***</b> (-0,37, -0,16) |
| <b>Métro</b>                              | -0,10 * (-0,20, 0,004)                        | <b>-0,12 **</b> (-0,23, -0,01)  |
| <b>Gare</b>                               | 0,20 (-0,05, 0,45)                            | 0,20 (-0,06, 0,47)              |
| <b>Fréquence bus</b>                      | <b>0,09 **</b> (0,003, 0,17)                  | <b>0,12 ***</b> (0,03, 0,21)    |
| <b>Infrastructures cyclables</b>          | <b>0,03 ***</b> (0,01, 0,06)                  | <b>0,03 **</b> (0,002, 0,05)    |
| <b>Capacité</b>                           | <b>1,02 ***</b> (0,88, 1,16)                  | <b>1,06 ***</b> (0,91, 1,21)    |
| <b>Nombre de stations avoisinantes</b>    | 0,17 (-0,04, 0,39)                            | <b>0,35 ***</b> (0,12, 0,58)    |
| <b>Capacité des stations avoisinantes</b> | -0,06 (-0,15, 0,03)                           | <b>-0,14 ***</b> (-0,23, -0,04) |
| <b>Distance au centre du réseau</b>       | <b>-0,54 ***</b> (-0,63, -0,44)               | <b>-0,48 ***</b> (-0,58, -0,38) |
| Observations                              | 393   | 393                             |
| R <sup>2</sup>                            | 0,71  | 0,69                            |
| R <sup>2</sup> ajusté                     | 0,70  | 0,68                            |
| Écart-type résiduel (df = 377)            | 0,42  | 0,44                            |
| Statistique F (df = 15 ; 377)             | 60,85 ***                                     | 55,65 ***                       |
| <i>Significativité</i>                    | <i>*p&lt;0,1 ; **p&lt;0,05 ; ***p&lt;0,01</i> |                                 |

Dans la catégorie sociodémographique, la population autour des stations BIXI semble avoir un impact négatif sur les déplacements entrants et sortants, ce qui est contre-intuitif. Une explication possible de ce phénomène est que les SR contenant davantage de population se situent plus loin des aires centrales (ce qui se confirme en partie par la corrélation positive de 0,34 entre la population et la distance du centre moyen du réseau) et qu'ils ont des superficies variables. La densité résidentielle possède un des impacts les plus importants et positifs, ce qui semble naturel. L'âge médian semble suggérer un effet négatif sur la génération de déplacements, à l'Origine comme à la Destination, ce qui est cohérent avec la revue de littérature. La présence de lieux d'emplois a également un impact significatif sur la production et l'attraction des flux, concordant avec plusieurs recherches évoquées dans la revue de littérature (Rixey, 2013; Tran, 2016; Tran et al., 2015).

En ce qui concerne l'usage du sol et l'environnement bâti, les lieux postsecondaires que sont les universités et cégeps ont un impact négatif sur les flux et non significatif, mais significatif dans la version des modèles sans variables corrélées. Il est fort probable que ceci soit une résultante de la temporalité des modèles, basés sur deux mois d'été, là où l'activité dans ce genre d'établissement est nécessairement à son plus faible, contrairement aux mois situés dans les épaules des saisons BIXI. Sans surprise, l'altitude des stations possède un effet négatif sur l'attraction des flux, avec un coefficient important de -0,26.

Les variables relatant l'offre de transport sont intéressantes. Effectivement, la présence de stations de métro a un effet négatif sur les flux et significativement sur l'attraction de déplacements. Il est possible dans ce cas que la zone tampon de 400 m soit trop grande pour tenir compte de déplacements multimodaux qui risquent de commencer et de se terminer près des stations. La fréquence de passage des autobus aux arrêts possède quant à elle un effet positif qu'il faut prendre avec un grain de sel, car il y a de fortes chances qu'une zone ayant une fréquence plus élevée soit plus centrale et donc les flux s'expliquent par une multitude d'éléments. La somme des infrastructures cyclables possède un impact, quoique faible, significatif pour les deux modèles ce qui est intéressant, et confirme également d'autres études sur le sujet.

Finalement, dans le cas de l'offre BIXI, il faut être prudent avec l'ensemble des variables. La capacité de la station est la variable la plus importante de toutes, mais les stations avec le plus de points d'ancrage sont nécessairement à des endroits où la demande est plus forte, ce qui est

également vrai pour le nombre de stations à proximité sur l'importance d'attraction des flux. Ceci est similaire à ce qu'évoquent Wang et al. (2016), Mateo-Babiano, Bean, Corcoran et Pojani (2016) et Zhao, Deng et Song (2014). De la même manière, une station étant plus loin du centre moyen du réseau de vélopartage risque de produire et d'attirer moins de déplacements.

### **3.2.4 Validation des modèles**

Les résultats de la précédente modélisation sont intéressants, mais il importe également de voir si elle est aussi efficace avec des données qui sont basées sur une autre temporalité. Ainsi, la validation de ces modèles est effectuée avec les déplacements BIXI de 2016. Cette année a été choisie puisqu'il s'agit d'une autre année de recensement, ce qui rend possible la mise à jour d'une des bases de données importantes. La validation des modèles sur une année ultérieure permet aussi de déterminer quelles variables, issues des différents modèles, ont un effet sur la demande des déplacements de manière continue.

Ainsi, le processus de sélection de données est effectué de nouveau avec les déplacements de 2016. La base de données comprend 464 stations et une saison de 215 jours. Un test de la somme des rangs Wilcoxon révèle que les mois de juin et août sont les plus similaires pour cette année, avec une valeur de 0,52. Cependant, afin de garder une certaine constance, les mois de juin et juillet sont conservés dans la base de données (probabilité de similarité de 0,13), comme pour l'an 2011. En ce qui concerne les jours de la semaine, les jours ouvrables possèdent des distributions similaires, tout comme en 2011, alors seuls les déplacements effectués du lundi au vendredi sont retenus. De plus, 11 stations sont également retirées des données, car leur période d'activité est inférieure à 80 % sur l'ensemble des 41 jours. Ce faisant, la base de données résultante contient 452 stations et 0,98 million de déplacements, ou 23 859 quotidiennement, en moyenne.

De plus, les autres jeux de données utilisés sont également mis à jour et datent tous de 2016, sauf les informations concernant les lieux d'éducation postsecondaire, qui datent de 2017. Il faut aussi tenir compte de l'évolution du réseau BIXI, puisque six stations sont présentes sur le territoire de la Ville de Longueuil en 2016. Les variables spatiales considérant ce territoire sont incluses dans les modèles de 2016. Les données sur les infrastructures cyclables pour ce lieu proviennent de leur portail de données ouvertes (<https://www.longueuil.quebec/fr/donnees-ouvertes>) et celles sur les passages d'autobus proviennent des GTFS du Réseau de transport de Longueuil (RTL).

Les résultats des modèles sont présentés au tableau 3.9. Un des premiers constats est le nombre de variables avec un effet significatif qui est plus faible sur les flux entrants et sortants en 2016 qu'en 2011. L'effet de la population et de la densité est encore plus important en 2016. L'âge médian possède un effet beaucoup moins important, ce qui pourrait signifier que des usagers plus vieux s'approprient ce mode de transport, mais il faut mentionner que cette variable n'est pas significative. La variable emplois, quant à elle, est toujours aussi importante en 2016. Il serait intéressant d'observer l'effet de cette variable sur les déplacements lors d'une période de pointe.

Tableau 3.9 : Résultats de la modélisation 2016 avec transformation de type puissance

|   | <b>Production</b>                             | <b>Attraction</b>              |
|---|---|--------------------------------|
| <b>Intercept</b>                          | 3,02*** (1,64, 4,41)                          | 3,37*** (1,95, 4,80)           |
| <b>Population</b>                         | <b>-0,32***</b> (-0,46, -0,18)                | <b>-0,40***</b> (-0,55, -0,25) |
| <b>Densité de population</b>              | <b>0,28***</b> (0,17, 0,38)                   | <b>0,33***</b> (0,22, 0,44)    |
| <b>Âge médian</b>                         | -0,06 (-0,29, 0,16)                           | 0,01 (-0,22, 0,24)             |
| <b>Emplois</b>                            | <b>0,10***</b> (0,05, 0,15)                   | <b>0,12***</b> (0,07, 0,17)    |
| <b>Postsecondaire</b>                     | 0,05 (-0,11, 0,21)                            | 0,02 (-0,14, 0,19)             |
| <b>Parc</b>                               | -0,01 (-0,02, 0,01)                           | -0,01 (-0,03, 0,01)            |
| <b>Élévation</b>                          | 0,08 (-0,02, 0,19)                            | <b>-0,21***</b> (-0,32, -0,10) |
| <b>Métro</b>                              | 0,01 (-0,10, 0,12)                            | -0,02 (-0,13, -0,10)           |
| <b>Gare</b>                               | 0,07 (-0,23, 0,37)                            | 0,06 (-0,25, 0,37)             |
| <b>Fréquence bus</b>                      | 0,01 (-0,02, 0,04)                            | 0,01 (-0,08, 0,10)             |
| <b>Infrastructures cyclables</b>          | 0,01 (-0,02, 0,04)                            | 0,01 (-0,02, 0,05)             |
| <b>Capacité</b>                           | <b>0,93***</b> (0,77, 1,10)                   | <b>0,89***</b> (0,72, 1,06)    |
| <b>Nombre de stations avoisinantes</b>    | -0,18 (-0,43, 0,07)                           | -0,14 (-0,39, 0,12)            |
| <b>Capacité des stations avoisinantes</b> | <b>0,12**</b> (0,02, 0,21)                    | <b>0,10**</b> (-0,01, -0,20)   |
| <b>Distance au centre du réseau</b>       | <b>-0,42***</b> (-0,52, -0,31)                | <b>-0,34***</b> (-0,44, -0,23) |
| Observations                              | 452   | 452                            |
| R <sup>2</sup>                            | 0,64  | 0,61                           |
| R <sup>2</sup> ajusté                     | 0,63  | 0,60                           |
| Écart-type résiduel (df = 436)            | 0,49  | 0,51                           |
| Statistique F (df = 15 ; 436)             | 52,12***                                      | 45,69***                       |
| <i>Significativité</i>                    | <i>*p&lt;0,1 ; **p&lt;0,05 ; ***p&lt;0,01</i> |                                |

Les coefficients des variables qui traitent de l'usage du sol et de l'environnement bâti changent peu. La fréquence des bus possède également un impact amoindri et non significatif. Ceci pourrait

être conséquence de plusieurs facteurs, que ce soit un changement important dans le service BIXI ou encore une stabilisation dans l'offre de transport en commun. Une analyse approfondie permettrait d'éclaircir cette particularité. Les variables qui traitent de l'offre BIXI sont quant à elles toujours aussi importantes, soit la capacité des stations et la distance de celles-ci par rapport au centre moyen du réseau. Il est intéressant de noter également le coefficient de détermination ajusté de 0,61, plus faible en 2016 qu'en 2011, alors qu'il est de 0,69. Des recherches supplémentaires sont nécessaires afin de mieux comprendre la fluctuation de l'impact des variables sur les flux tout en considérant que certaines d'entre elles montrent des signes importants de corrélation.

### **3.2.5 Limites et perspectives**

Il convient de souligner que cette modélisation et les résultats se basent sur plusieurs hypothèses, telles que présentées à la section 3.2.1. La régression linéaire suppose qu'il existe une relation linéaire entre chaque variable indépendante et celle à expliquer. De plus, ce type de modèle possède une hypothèse d'homoscédasticité, c'est-à-dire que les variances des variables explicatives doivent être identiques. Puis, le risque de multicolinéarité des variables indépendantes entraîne des effets problématiques dans la modélisation, car il est plus ardu de séparer les effets de ces variables. Donc, malgré une analyse fine des données, de leur distribution, de leur corrélation, de leur transformation et des résidus du modèle, il se peut que certains de ces éléments posent problème.

Outre ces limites qui sont intrinsèques à la régression linéaire, d'autres limites décrites par Bonnel (2002) existent sur les modèles de génération de manière générale. Premièrement, les flux sortants et entrants sont modélisés de manière indépendante. Ceci signifie qu'il y a forcément une différence entre l'ensemble des déplacements produits et attirés par le modèle contrairement à ce qui se produit dans la réalité. Il faudrait donc y avoir une rectification éventuelle pour pallier ce problème.

Deuxièmement, alors que certaines variables qui sont incluses dans les modèles tentent de décrire l'offre de transport, celles-ci ne tiennent pas compte de changements dans cette offre qui risque d'affecter les usagers, qui pourraient utiliser un autre mode de transport, par exemple.

Troisièmement, les modèles de génération se basent sur des paramètres fixés dans le temps. Les coefficients de la régression vont nécessairement changer en fonction de divers éléments. Effectivement, les flux des déplacements sont perturbés par la disponibilité ou non des vélos aux stations en temps réel. C'est le genre de constat que la moyenne quotidienne basée sur une année

ne permet pas de prendre en compte. C'est une des raisons qui explique la validation des modèles de 2011 avec des données de 2016. Dans une perspective idéale, il faudrait nécessairement avoir une série de ces modèles permettant de voir l'évolution des coefficients dans le temps.

Quatrièmement, les variables indépendantes utilisées constituent un choix. La littérature sur les modèles de génération de manière générale et celle en particulier sur le vélopartage propose de nombreuses variables ayant un impact sur ce mode de déplacement. Or, ce choix a nécessairement un impact sur la qualité de prédiction des modèles. En plus de pouvoir ajouter, retirer, ou modifier certaines des variables, il est également possible d'utiliser d'autres paramètres. Dans le cas présent, l'accent est mis sur la planification stratégique, ce qui explique pourquoi la modélisation des flux est effectuée sur une journée moyenne pour toute une année. Plusieurs modèles peuvent être estimés pour différentes temporalités avec des résultats qui varient. De plus, la zone tampon de 400 m pour les variables spatiales est un choix arbitraire éclairé par la littérature qui a nécessairement une influence sur les résultats. Il se peut que, selon les motifs de déplacements, des usagers acceptent de se déplacer sur une plus grande distance (ou plus courte) pour accéder à une station BIXI. Bref, le choix des variables explicatives est en soi une limite à l'étape de génération.

Cinquièmement, la qualité des données récoltées et utilisées constitue la dernière limite énoncée. Plusieurs bases de données ont été utilisées dans la modélisation. Un grand effort a été déployé afin d'assurer une certaine synchronicité dans les données, afin que les analyses et les résultats soient les plus cohérents possible. Malgré ceci, la jointure de bases de données indépendantes implique nécessairement des découpages territoriaux différents ou encore des méthodes de récolte de données différentes ce qui a un impact sur leur fiabilité.

### **3.3 Conclusion du chapitre**

Ce chapitre a permis une analyse exploratoire des données sur les déplacements effectués aux stations par les usagers BIXI. Il a mis en lumière la grande variabilité qui existe dans l'utilisation du vélopartage, en relation avec le temps, comme le moment dans l'année ou dans une semaine, ou dans l'espace, soit la localisation de l'emprunt et du retour d'un vélo. Ceci est cohérent avec la revue de littérature. Il serait également possible de se pencher spécifiquement sur des périodes horaires, comme la pointe matinale ou du soir afin de montrer cette fluctuation. Cette plage horaire est d'ailleurs la temporalité de prédilection du prochain chapitre qui se penche sur la deuxième étape de la PSC.

Bref, après analyse des déplacements, la régression linéaire multiple a été présentée, la technique utilisée afin de modéliser la génération des déplacements en vélopartage. Après avoir présenté les hypothèses sur lesquelles se base ce type de modèle, les variables dépendantes et indépendantes qui composent les modèles sont détaillées, en plus de présenter le processus de développement afin d'améliorer les résultats.

Ces derniers ont été présentés et révèlent un impact important des variables sociodémographiques et celles ayant trait à l'offre de mobilité et à l'offre BIXI sur les flux sortants et entrants. Les modèles ont par la suite été validés sur une autre année où une disparité dans les résultats fait surface. Ceci implique donc qu'une analyse plus approfondie doit être effectuée. Il pourrait être question, par exemple, de développer une typologie de jours basée sur des facteurs externes ou même de créer des groupes de stations selon un taux de flux produits et attirés, tel que cela a été fait brièvement dans l'analyse spatiale.

## **CHAPITRE 4 MODÉLISATION DE LA DISTRIBUTION DES DÉPLACEMENTS**

Ce chapitre s'attarde à la distribution spatiale des déplacements effectués par les usagers en vélopartage, soit les paires O-D du réseau BIXI. Puisque ce document a une portée stratégique, il est question d'observer ces paires sur une base annuelle, mais seulement pour la pointe matinale. Deux années sont à l'étude, soit 2011 et 2014. Les bases de données des déplacements de ces périodes sont mises à profit afin de faire ressortir des tendances intéressantes. Par la suite, les paires O-D sont modélisées grâce au modèle gravitaire, qui se base sur une métrique de distance ou de durée afin de reproduire les observations. Modèle et ses composantes sont présentés en détail avant de passer aux résultats. Ces derniers sont analysés à l'aide de statistiques sommaires et de représentations cartographiques. Finalement, les modèles calibrés sont appliqués sur une année ultérieure et divers indicateurs sont employés en guise de validation.

### **4.1 Introduction aux concepts**

Les déplacements dans un système de vélos en libre-service sont effectués à partir d'une station, l'Origine, vers une autre, la Destination. Les possibilités de déplacements dépendent du nombre de stations  $N$ , soit  $N * N$ . Chaque possibilité de déplacement comprend une Origine et une Destination unique, soit une paire Origine-Destination (O-D), schématisée à la figure 4.1. Par exemple, un réseau comprenant 100 stations signifie qu'il compte 10 000 paires O-D. Dans ce réseau fictif, 100 de ces paires sont constituées d'Origine et de Destination identiques puisqu'un déplacement qui commence et se termine à la même station constitue également une paire O-D. Ceci permet d'introduire la problématique des déplacements aux mêmes Origines et Destinations.

Un utilisateur peut effectivement emprunter et retourner un vélo à la même station. Cependant, un usager qui emprunte une bicyclette et la retourne à la même station pour quelque raison (bris d'équipement, changement d'idée ou autre) est également considéré comme un déplacement par le système. Ainsi, afin de minimiser ces faux déplacements, seuls les enregistrements supérieurs à cinq minutes où l'Origine et la Destination sont les mêmes sont gardés dans la base de données. Les déplacements avec une Origine et une Destination différente ne sont pas filtrés.

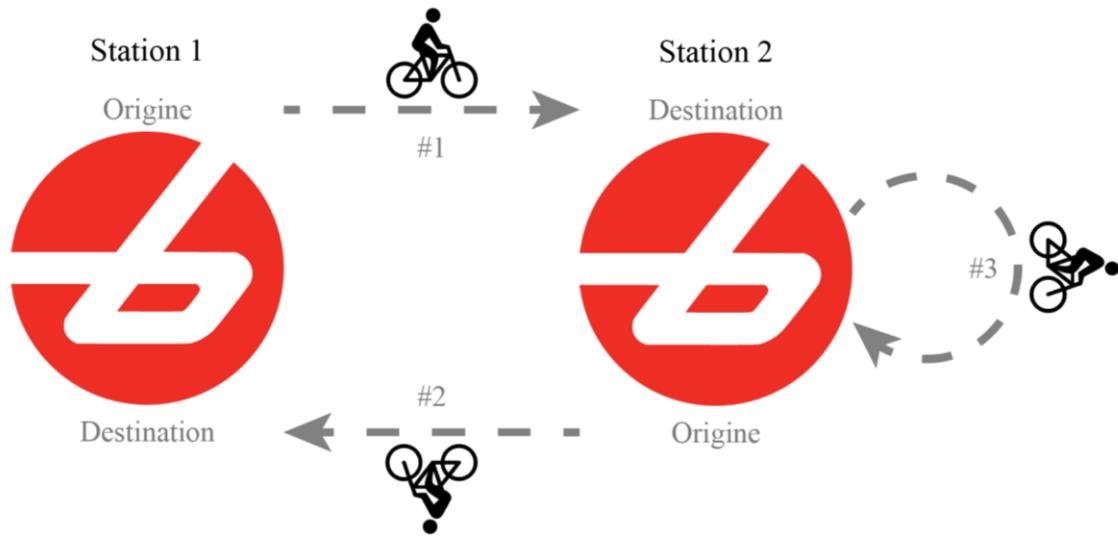


Figure 4.1 : Exemple de trois paires Origine-Destination

Dans un même ordre d'idée, un utilisateur qui effectue un déplacement entre deux stations possède un très grand nombre d'itinéraires possibles. Alors que cette information pourrait nécessairement être pertinente, elle est indisponible, puisque seules les stations d'Origine et de Destination sont connues. En ce sens, le concept de ligne de désir est utilisé, qui est en fait une représentation directe entre l'Origine et la Destination d'un déplacement à vol d'oiseau, comme illustré à la figure 4.2. Une ligne de désir est donc la représentation d'un déplacement pour une paire O-D donnée, qui ne prend pas en considération le réseau routier : il est plutôt question de visualiser les observations.



Figure 4.2 : Plusieurs itinéraires possibles à gauche, une seule ligne de désir à droite

## 4.2 Triage de données

Les données utilisées dans ce chapitre proviennent en partie de la base de données BIXI sur les déplacements aux stations présentée à la section 3.1. Celle-ci fournit les coordonnées des stations, leur capacité et les flux aux stations. Or, ceux-ci ne sont pas reliés. Puisqu'il est question de la distribution des déplacements, il est nécessaire de connaître l'Origine et la Destination des déplacements, ce qui constitue une autre base de données. De plus, les informations sur la distribution des déplacements des usagers BIXI de deux années distinctes sont analysées afin de vérifier si une certaine stabilisation du système de vélopartage s'est effectuée depuis son lancement en 2009. Ainsi, les deux années de référence sont 2011 et 2014 et, pour chacune de ces périodes, il y avait respectivement 215 et 216 jours dans les saisons, ainsi que 405 et 460 stations. Après un traitement de données, il s'est avéré qu'une des stations en 2014 était en fait un doublon, soit une station située au même endroit avec un identifiant qui a changé au cours de l'année, ce qui fait

passer le compte à 459 stations. Ceci signifie qu'il y avait 164 025 paires O-D en 2011 et 210 681 en 2014.

Afin de réduire la variabilité dans les déplacements, la somme de ces derniers a été calculée pour chaque paire O-D pour les deux années à l'étude et ce, seulement pour la pointe matinale lors de jours ouvrables, soit de 6 à 9 h. La somme de ces déplacements correspond à environ 15 % de l'ensemble des déplacements effectués pendant les deux années, tel que représenté à la figure 4.3. De plus, l'analyse et la modélisation de la distribution des déplacements ne sont effectuées que pour les usagers membres du système, car ils effectuent la grande majorité de ceux-ci (96 et 98 % selon l'année).

|                        | 2011                  | 2014                  |
|------------------------|-----------------------|-----------------------|
| <i>Stations</i>        | 405                   | 459                   |
| <i>Paires O-D</i>      | 164 025               | 210 681               |
| Déplacements totaux    | 4,2 M                 | 3,2 M                 |
| Jours ouvrables        | 3,1 M (73 %)          | 2,4 M (74 %)          |
| <b>Pointe matinale</b> | <b>451 715 (15 %)</b> | <b>389 088 (16 %)</b> |
| Membres                | 432 884 (96 %)        | 380 093 (98 %)        |
| Occasionnels           | 18 831 (4 %)          | 8 992 (2 %)           |

Figure 4.3 : Sommaire des déplacements effectués en 2011 et 2014

### 4.3 Analyse de la distribution des déplacements observés

Cette nouvelle base de données permet de constater que l'ensemble des stations produisent et attirent au moins un déplacement pendant la période analysée en 2011. En 2014, toujours pendant la pointe matinale, les membres ont utilisé 456 stations à l'Origine d'un déplacement et 457 stations à la Destination, ce qui constitue une activité à plus de 99 % des 459 stations. Cependant, dans les valeurs présentées au tableau 4.1, 29,3 % des paires O-D ont été utilisés par les membres pendant

la période observée en 2011, soit près du tiers, et cette proportion tombe à un cinquième (20,6 %) en 2014. Ceci s'explique possiblement par une hausse du nombre de stations (et donc de paires O-D possibles) pour cette année. Les autres statistiques du tableau sont plutôt similaires entre les années et permettent de constater un grand écart-type du nombre de déplacements effectués dans les paires O-D qui sont actives. Une paire O-D active signifie qu'il y a eu moins un déplacement effectué entre ces stations.

Tableau 4.1 : Sommaire des déplacements des membres en période de pointe matinale

| <b>Statistiques sommaires</b>                           | <b>2011</b>     | <b>2014</b>     |
|---|-----------------|-----------------|
| <b>Stations actives à l'Origine</b>                     | 405 (100 %)     | 456 (99,4 %)    |
| <b>Stations actives à la Destination</b>                | 405 (100 %)     | 456 (99,4 %)    |
| <b>Paires O-D actives</b>                               | 48 129 (29,3 %) | 43 432 (20,6 %) |
| <b>Déplacements moyens par paire O-D active</b>         | 9,0             | 8,8             |
| <b>Écart-type des déplacements par paire O-D active</b> | 19,4            | 19,1            |
| <b>Nombre de déplacements maximal par paire O-D</b>     | 665             | 656             |

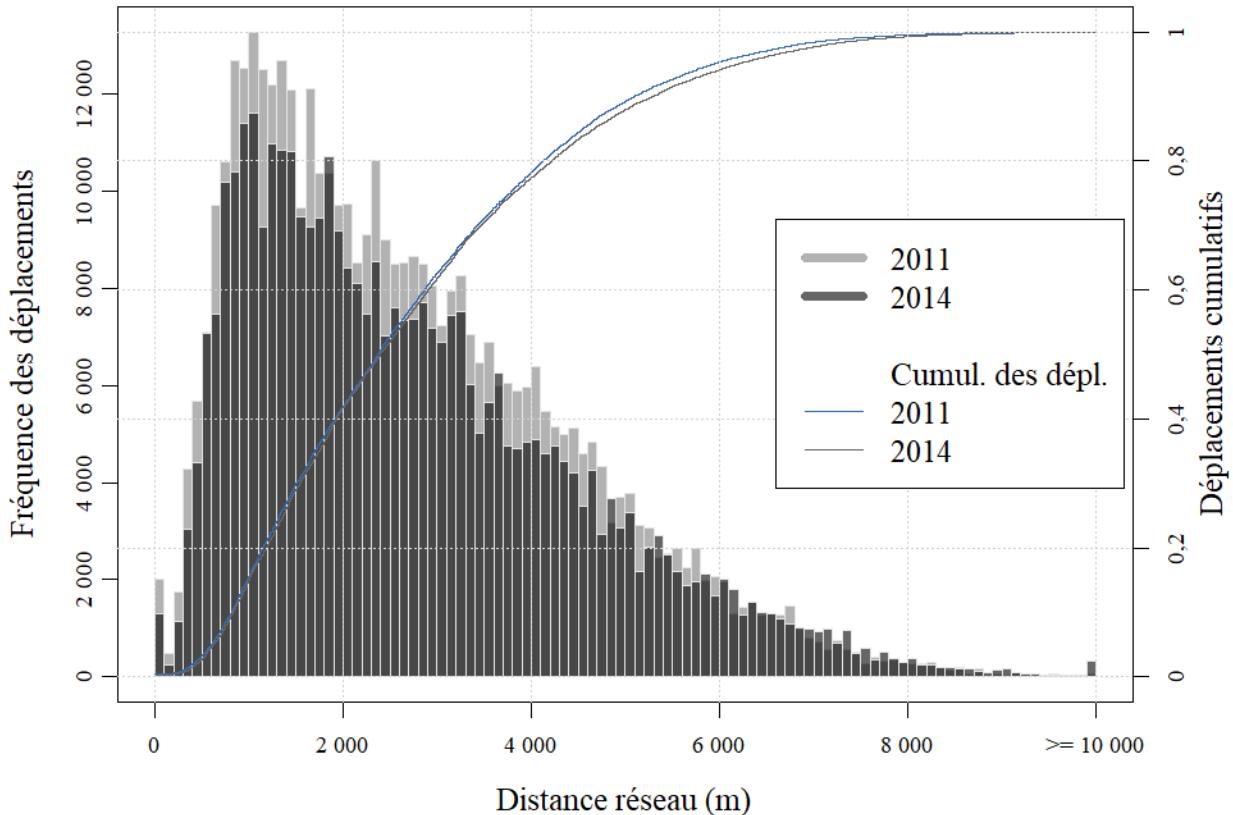
Ensuite, quatre métriques qui sont utilisées pour déduire une fonction de coût dans des variations du modèle gravitaire, détaillé à la section 4.4.1.2, sont présentées. Celles-ci sont utilisées en tant que distance et durée de déplacement moyenne entre chaque paire O-D. Ces valeurs sont estimées avec le calculateur de trajets *Open Source Routing Machine* (Luxen & Vetter, 2011), soit OSRM, qui permet de calculer la distance à parcourir sur le réseau routier et cycliste entre toutes les stations du réseau. De plus, alors que les durées des déplacements sont disponibles, celles-ci varient beaucoup selon les usagers. Aussi, une vitesse moyenne théorique de 15 km/h est imposée et la durée est calculée selon la distance réseau estimée pour chaque paire O-D. Puis, une seconde manière de calculer la distance et la durée des déplacements se base quant à elle sur une étude de Scarf et Grehan (2005). Elle permet de considérer les pentes sur le réseau routier afin d'ajouter une pénibilité pour les montées. Le calcul applique un parcours supplémentaire de huit mètres horizontaux pour chaque différence d'élévation d'un mètre. Dans tous les cas, la distance et la

durée calculées se basent sur le chemin le plus court entre deux stations. Les valeurs moyennes de ces métriques sont présentées au tableau 4.2.

Il est également possible de remarquer une hausse de la distance moyenne des déplacements, soit de près de 70 mètres. Ceci pourrait être une résultante de l'augmentation du nombre de stations dans le réseau, plus éparpillées dans le territoire, et donc des déplacements plus longs. Tableau 4.2 Données moyennes des déplacements des membres selon quatre métriques

| Métriques  | 2011  | 2014  |
|--|-------|-------|
| <b>Distance réseau moyenne des déplacements (m)</b>  | 2 697 | 2 763 |
| <b>Durée moyenne théorique des déplacements selon distance réseau (min.)</b>               | 11,8  | 12,0  |
| <b>Distance réseau estimée moyenne des déplacements avec dénivelé (m)</b>                  | 2 703 | 2 768 |
| <b>Durée moyenne théorique des déplacements selon distance réseau avec dénivelé (min.)</b> | 12,0  | 12,2  |

Ainsi, la figure 4.4 permet de constater la distribution des déplacements selon l'une de ces métriques, soit la distance réseau. Il est possible d'y voir que les déplacements sous les 6 km étaient plus nombreux en 2011 qu'en 2014. Ceci n'est pas étonnant puisqu'il y avait simplement davantage de déplacements effectués pendant la première année. Les déplacements les plus nombreux sont sous 2 km pour les deux années alors que 80 % des déplacements sont d'un peu plus de 4 km.



De manière similaire, la figure 4.5 illustre la distribution des déplacements selon la durée moyenne théorique de ceux-ci, calculée à partir de la distance réseau, qui constitue une autre métrique. La grande majorité des déplacements (plus de 80 %) seraient de 20 minutes ou moins alors que plus de 95 % de ceux-ci sont inférieurs à 30 minutes, période incluse gratuitement avec toute location de BIXI. La durée qui revient le plus souvent est relativement courte, soit environ cinq minutes. Les graphiques présentant les courbes de distribution des déplacements selon les deux autres métriques, celles prenant en compte le dénivelé sont présentées aux annexes C et D.

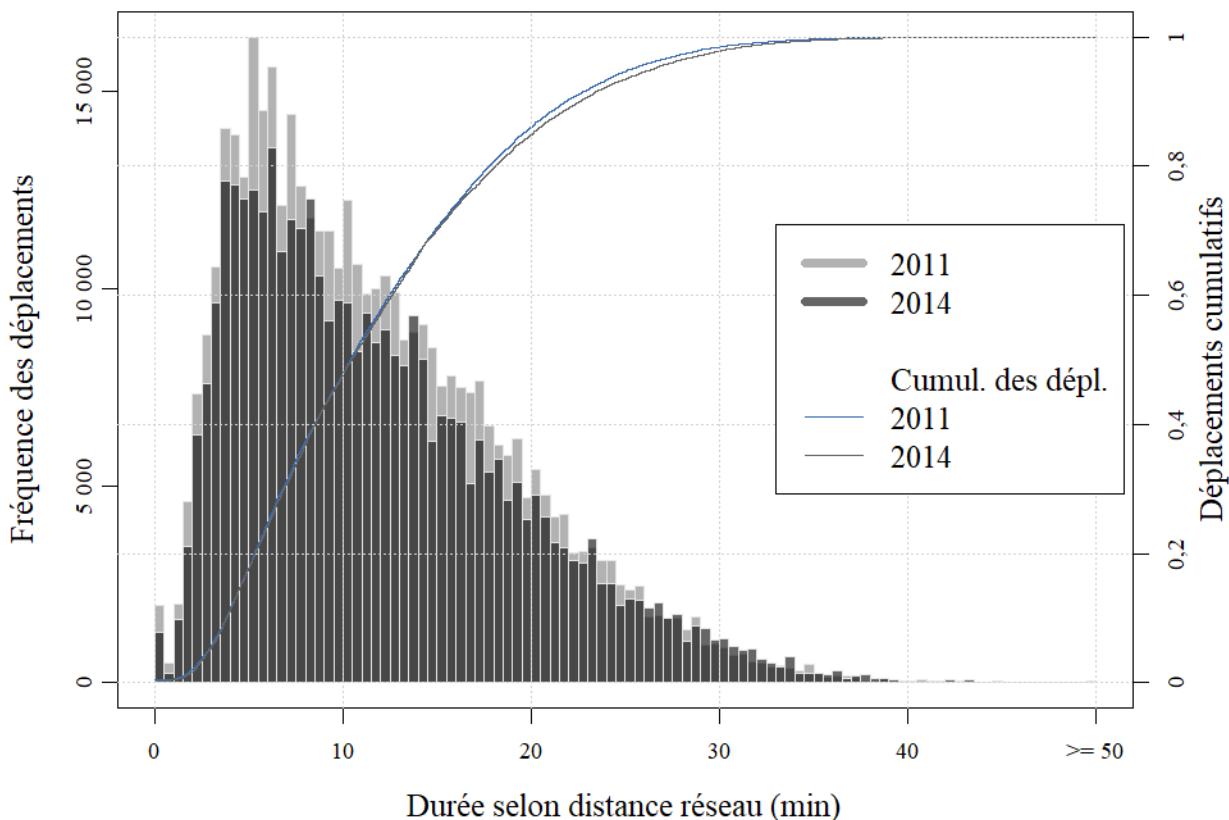


Figure 4.5 : Distribution et cumulatif des déplacements selon la durée sur distance réseau

Par la suite, les déplacements effectués sur chaque paire O-D ont été additionnés selon chaque Origine et Destination, ce qui est présenté à la figure 4.6. Premièrement, elle permet de constater qu'il y a peu de changement en 2014 par rapport à 2011, si ce n'est qu'encore plus de stations sont peu utilisées en tant que Destination. Deuxièmement, elle permet de voir que de nombreuses stations sont peu utilisées en tant que Destination, alors qu'une quinzaine sont très achalandées pour la période à l'étude, c'est-à-dire qu'elles reçoivent plus de 5 000 déplacements (jusqu'à 12 791) chacune. En d'autres termes, l'Origine des déplacements est plutôt dispersée alors que la

Destination est concentrée. La somme maximale des déplacements à l'Origine est de 4 526 en 2011 et 3 660 en 2014 et la médiane est de 903 et 713 respectivement. La médiane des déplacements aux stations de Destination est plus faible, soit, pour les mêmes années, de 482 et 363.

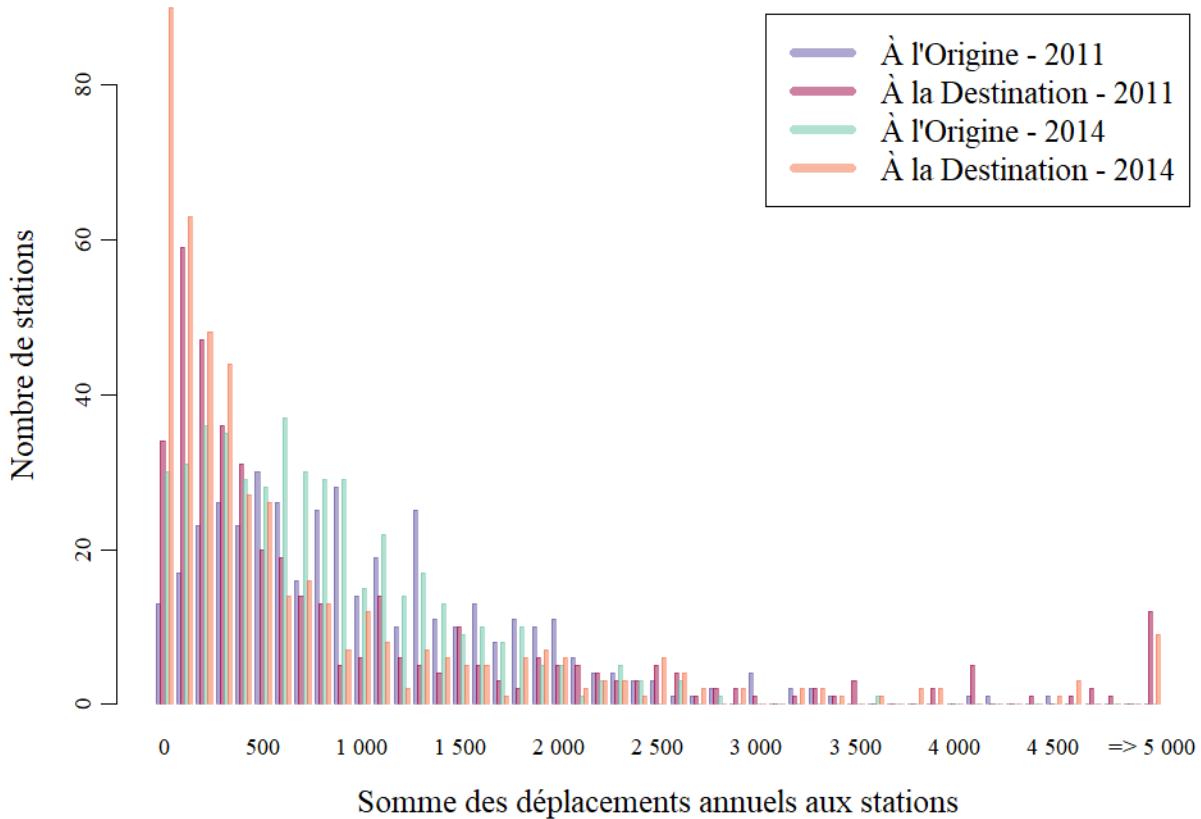


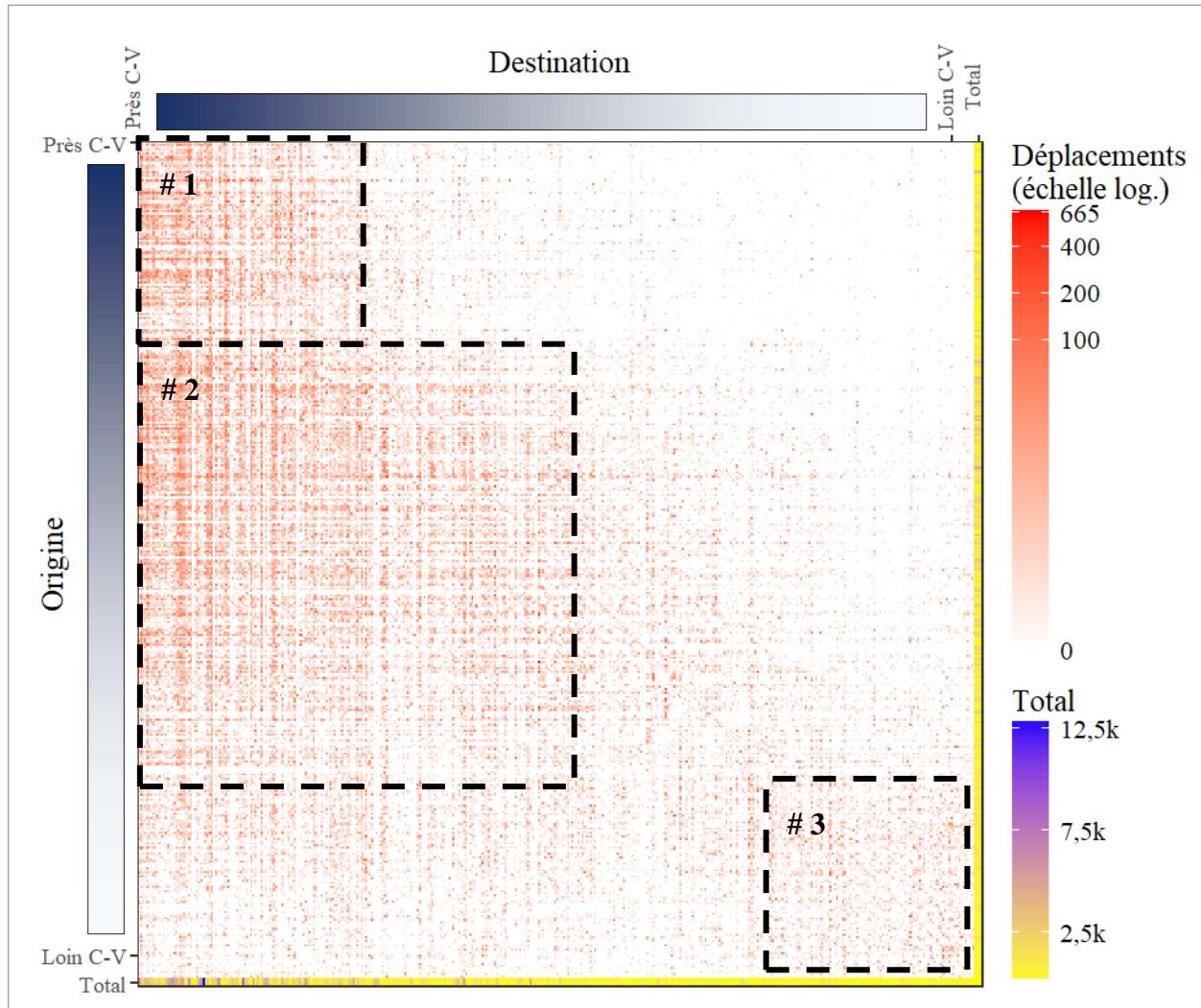
Figure 4.6 : Distribution de la somme des déplacements à l'Origine et à la Destination

Ensuite, afin de visualiser les déplacements sur l'ensemble des paires O-D, une matrice de ces dernières est produite. Ainsi, pour reprendre l'explication de la section 2.5, ce genre de matrice présente tous les déplacements selon l'Origine (les rangées) et la Destination (les colonnes). Afin d'organiser la matrice et de faciliter la visualisation, les stations ont été classées selon leur distance au centre-ville, tel que présenté précédemment à la figure 2.4. Ceci implique donc que deux stations, qui ne sont pas nécessairement près l'une de l'autre, mais qui sont équidistantes du centre-ville, peuvent se retrouver voisines dans la matrice. Ceci permet tout de même de créer un certain ordre.

Ainsi, la matrice Origine-Destination des déplacements effectués par les membres pour l'année 2011 est présentée au tableau 4.3 alors que celle de 2014 est à l'annexe E. Une seule des matrices permet amplement l'analyse des déplacements par paire O-D puisque les tendances sont

similaires. Le classement des stations selon la distance au centre-ville implique que les stations dans le coin supérieur gauche sont plus près de celui-ci, alors que plus elles sont éloignées sur l'axe des abscisses et des ordonnées, plus elles sont en périphérie du réseau. Ce classement est représenté par les barres violettes situées à gauche et en haut de la matrice, à titre indicatif seulement, permettant de se retrouver avec la figure 2.4. La diagonale de la matrice est constituée des déplacements dont l'Origine et la Destination sont les mêmes stations et exclut les déplacements inférieurs à cinq minutes.

Tableau 4.3 : Matrice Origine-Destination des membres en 2011



La somme des déplacements aux stations d'Origine et de Destination, aussi appelée les marges, est représentée aux extrémités droite et inférieure grâce à une autre échelle de couleur. Elle permet de constater la répartition des déplacements aux stations d'Origine et par le fait même la concentration des déplacements aux Destinations à proximité du centre-ville, alors que la plupart totalisent bien moins de 2 500 attractions. Il est possible de voir des régions de la matrice qui comprennent très peu de déplacements, comme dans le coin supérieur droit, ce qui signifie que très peu de déplacements quittent le centre-ville pour se diriger vers des stations plus éloignées. La grande présence de blanc dans la matrice s'explique puisque seulement 29,3 % des paires O-D sont actives en 2011, c'est-à-dire ayant au moins un déplacement, et de manière un peu plus faible, cette proportion est à 20,6 % en 2014. Ceci a été présenté au tableau 4.1. Certaines concentrations importantes de déplacements entre des paires O-D sont encadrées et spatialisées dans les figures suivantes.

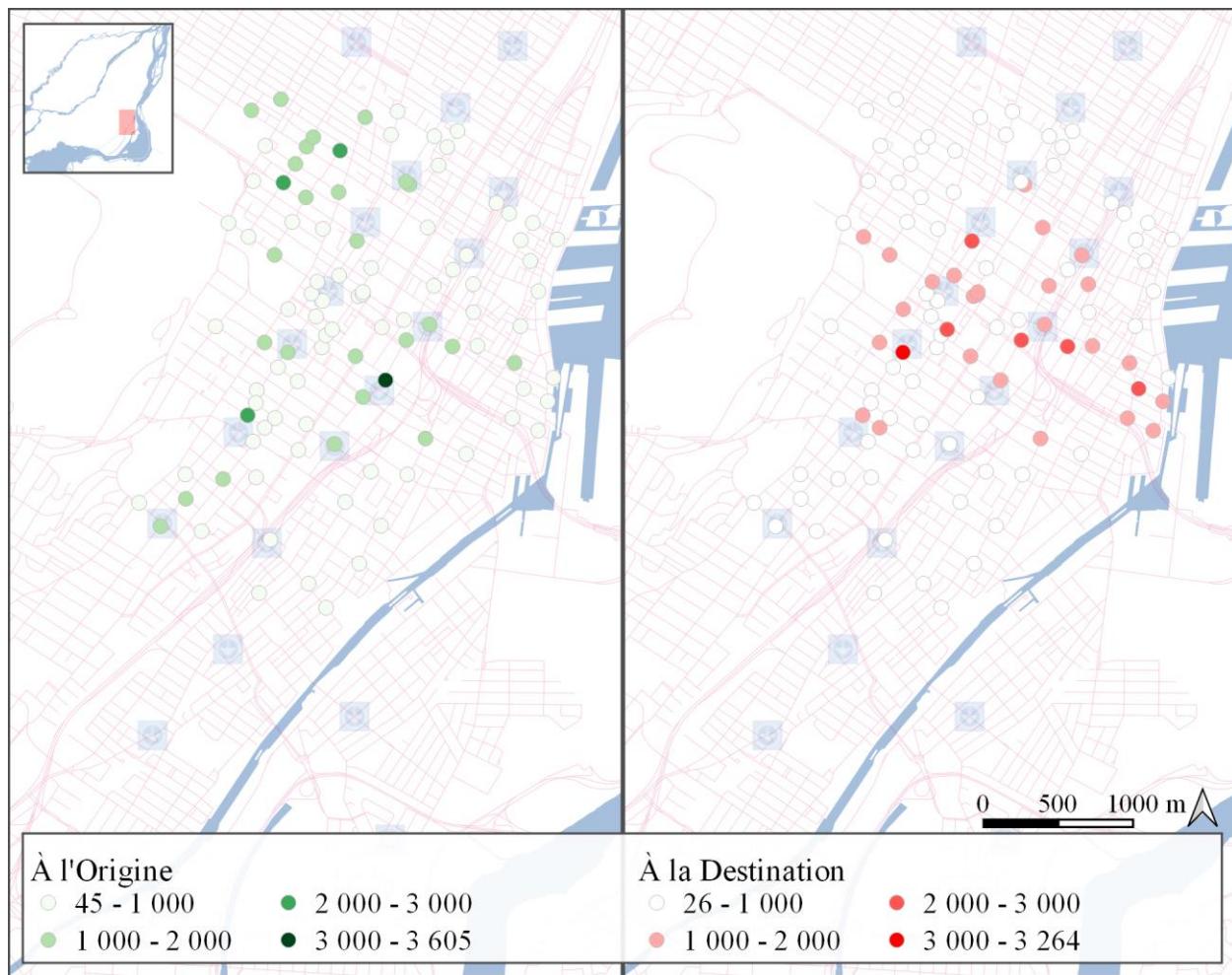


Figure 4.7 : Somme des déplacements aux stations de l'encadré #1 de la matrice O-D 2011

La figure 4.7 illustre donc la somme des déplacements à l'Origine et à la Destination dans l'encadré #1 du Tableau 4.3, soit les 50 stations les plus près du centre-ville. Elle permet de constater l'importance des paires O-D situées dans la région du centre-ville, donc courtes et concentrées, spatialement parlant. La station BIXI la plus importante en fonction de l'Origine de ces déplacements est adjacente à la station de métro Bonaventure, alors que la destination la plus achalandée est à proximité de la station de métro Peel. Il est possible qu'une partie de ces déplacements constitue le premier et le dernier kilomètre d'un déplacement plus important, permettant de lier, par exemple, une station de métro et le lieu de travail.

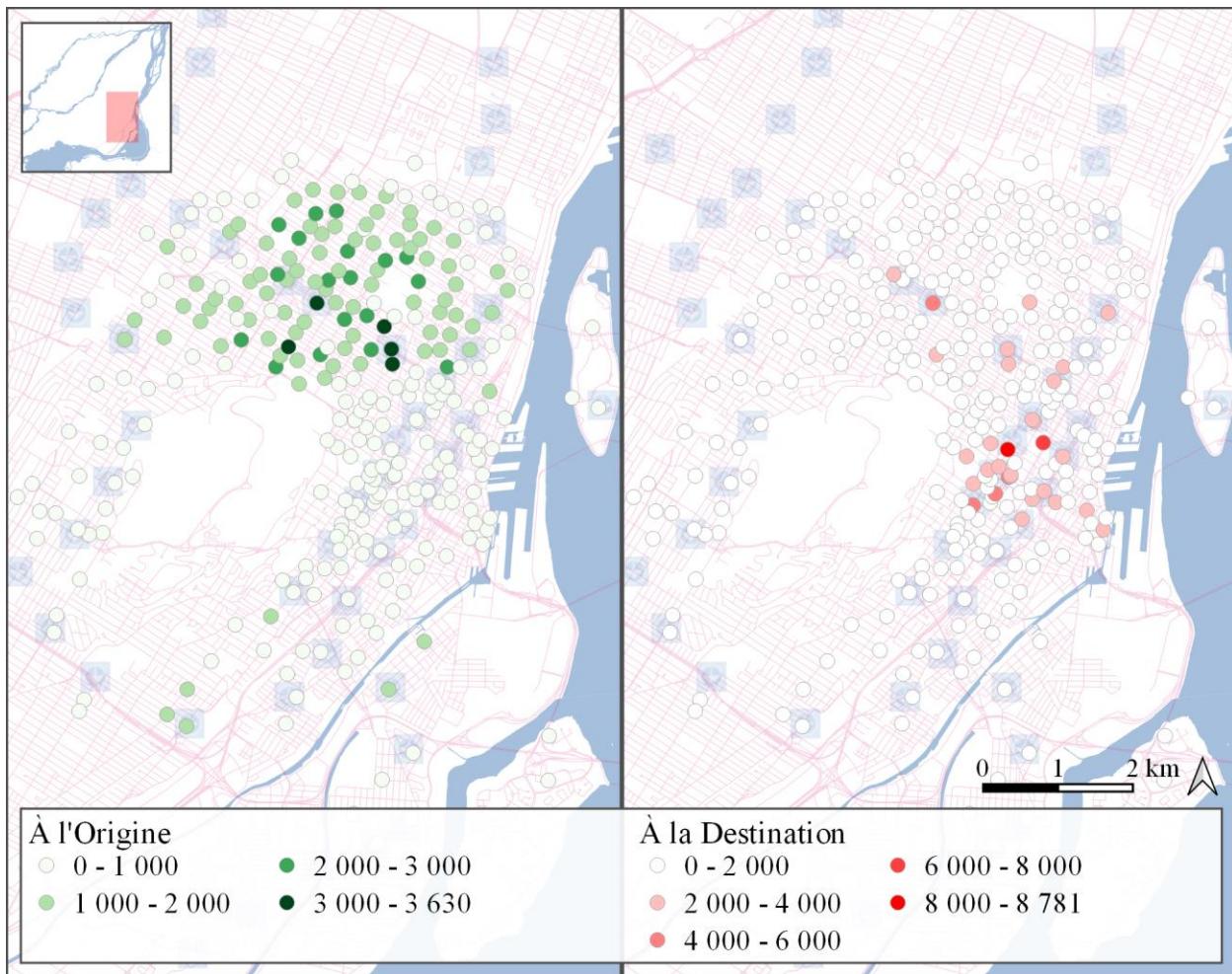


Figure 4.8 : Somme des déplacements aux stations de l'encadré #2 de la matrice O-D 2011

La figure 4.8 représente quant à elle l'encadré #2 de la matrice, soit les déplacements où l'Origine est surtout dans les quartiers centraux, principalement le Plateau Mont-Royal. Contrairement à l'encadré des déplacements précédent, il est possible de constater que les déplacements par station

d'Origine sont, dans ce cas, bien plus nombreux. Or, il faut mentionner qu'il s'agit de 200 stations, ce qui est d'autant plus impressionnant puisque le nombre de stations de Destination en importance est de moins de 30, signifiant ainsi une grande concentration des déplacements, notamment autour du centre-ville et à proximité de stations de métro.

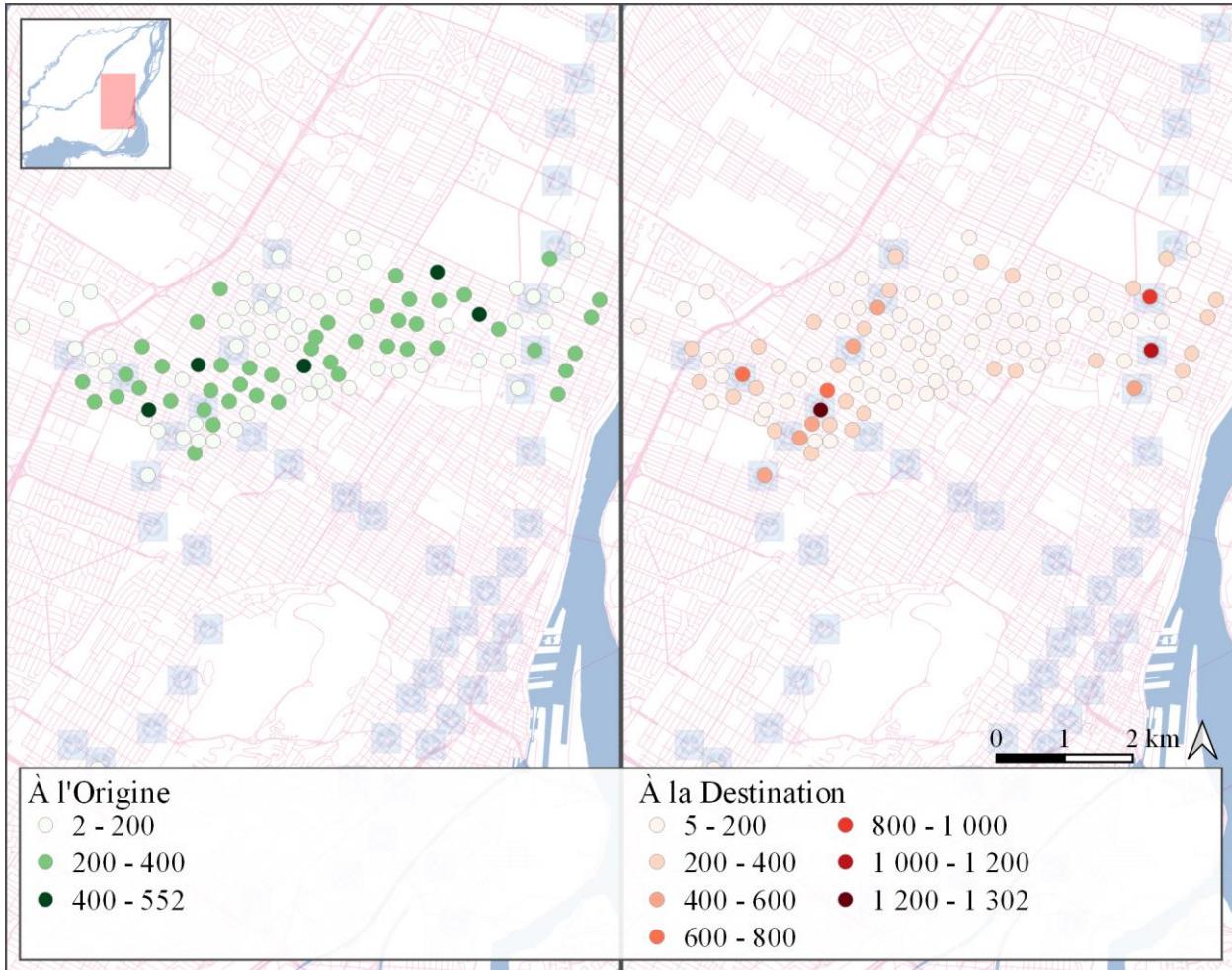


Figure 4.9 : Somme des déplacements aux stations de l'encadré #3 de la matrice O-D 2011

Finalement, l'encadré #3 met en lumière une petite concentration de déplacements, dans le coin inférieur droit de la matrice, illustré à la figure 4.9. Ces déplacements ont des Origines qui sont dispersées dans les quartiers centraux (Rosemont, Villeray, Hochelaga). Encore une fois, les Destinations de ces déplacements sont surtout à des stations BIXI adjacentes à des stations de métro (Jean-Talon, Pie-IX et Viau, entre autres), ce qui permet de supposer une combinaison de modes.

L'analyse de ces trois encadrés de paires O-D permet de visualiser de manière efficace le déséquilibre qui se produit dans le système, comme mentionné pendant la revue de littérature.

Effectivement, alors que les déplacements à l'Origine sont plus dispersés sur le territoire, les Destinations sont, quant à elles, concentrées.

Ensuite, il semble pertinent de spatialiser les paires O-D. Cependant, étant très nombreuses, seules les 20 lignes de désir les plus parcourues sont présentées à la figure 4.10. En guise de comparaison, les déplacements des membres et des utilisateurs occasionnels sont illustrés. Chez les membres, l'Origine ou la Destination semble avoir une station de métro comme point d'attache. De plus, les déplacements sont courts, presque tous inférieurs à un kilomètre, contrairement à ceux effectués par les usagers occasionnels, qui sont plus longs et plus disparates dans l'espace. Peu de lignes de désir sont identiques pour ces derniers et les membres. Il est également intéressant de constater que plusieurs lignes de désir sont identiques chez les membres, peu importe l'année, ce qui n'est pas le cas chez les usagers occasionnels.

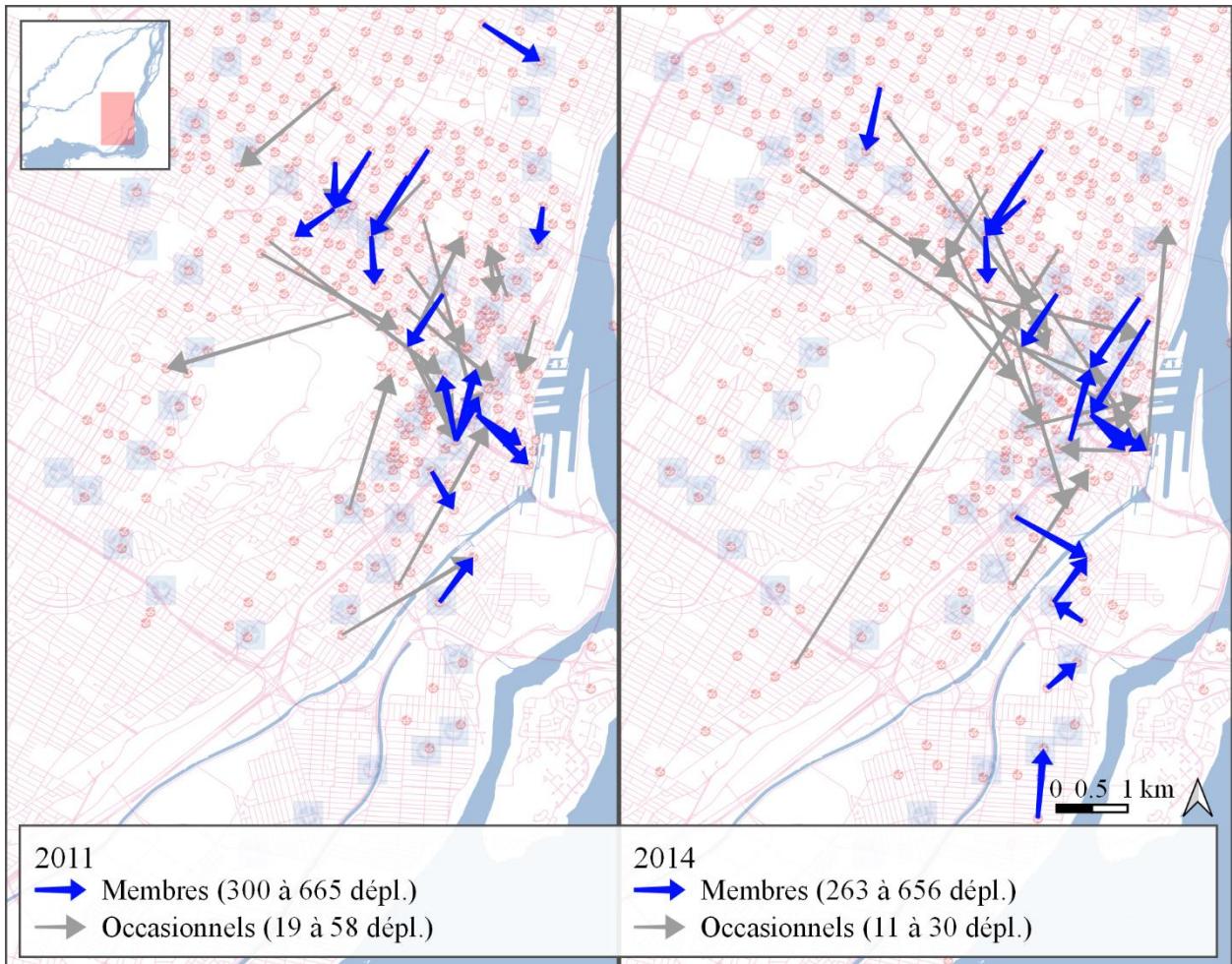


Figure 4.10 : Lignes de désir plus importantes en 2011 et en 2014

Ceci termine donc l'analyse des paires O-D utilisées par les membres. La prochaine section se penche sur la technique utilisée afin de reproduire ces déplacements, soit le modèle gravitaire.

## 4.4 Modélisation de la distribution des déplacements

Afin de modéliser les déplacements en vélopartage, des modèles gravitaires sont calibrés à partir des données analysées dans la section précédente. Les modèles de distribution constituent donc la deuxième étape de la PSC et remplissent l'ensemble des cellules de la matrice O-D à partir des marges (section 2.5.1.2). Il est question ici de développer la notion du modèle gravitaire alors que son fonctionnement est présenté en détail dans cette section ainsi que l'ensemble des modèles calibrés selon divers paramètres. Les explications utilisées dans cette section proviennent des mêmes auteurs que pour l'étape de génération, soit Bonnel (2002), McNally (2007) et Ortuzar et Willumsen (2011).

### 4.4.1 Le modèle gravitaire

Le modèle gravitaire se nomme ainsi puisqu'il se base sur la notion de la loi gravitationnelle de Newton où l'attraction entre deux corps est proportionnelle à leur masse. Ce concept a été repris éventuellement dans l'étude des transports et suppose donc que les déplacements entre deux endroits dépendent d'un certain poids, soit la fonction de résistance. Le modèle gravitaire possède plusieurs formes et celle utilisée dans ce cas est présentée à l'équation 4.1 :

$$T_{ij} = \alpha_i \beta_j P_i A_j f(c_{ij}) \quad 4.1$$

où  $T_{ij}$  est le nombre de déplacements effectués de la station  $i$  à  $j$ ;  $\alpha_i$  et  $\beta_j$  des facteurs d'ajustement;  $P_i$  la somme des déplacements produits à la station  $i$ ;  $A_j$  la somme des déplacements attirés à la station  $j$  et  $f(c_{ij})$  la fonction de coût.

#### 4.4.1.1 Les facteurs d'ajustement

Les facteurs d'ajustement permettent d'assurer la double contrainte aux marges, c'est-à-dire la somme des déplacements produits et attirés par chacune des stations. La formulation pour les calculer est la suivante :

$$\alpha_i = \frac{1}{\sum_j \beta_j A_j f(c_{ij})} \quad 4.2$$

$$\beta_j = \frac{1}{\sum_i \alpha_i P_i f(c_{ij})} \quad 4.3$$

Puisque chaque paramètre dépend de l'autre, ils doivent être calculés de manière itérative. D'abord, tous les paramètres  $\alpha_i$  et  $\beta_j$  sont initialisés à 1. La valeur des  $\alpha_i$  est d'abord calculée avec l'équation 4.2. Ensuite, les coefficients  $\beta_j$  sont calculés avec les nouvelles valeurs de  $\alpha_i$  et l'équation 4.3. Ce processus est répété jusqu'à ce que les valeurs des paramètres ne changent plus. Puisque cette mécanique nécessite un temps de calcul important (dépassant facilement les 24 h, selon les variations de la fonction de coût), les facteurs d'ajustements sont arrondis à 20 chiffres après la virgule.

#### 4.4.1.2 La fonction de coût

Cet élément intégré au modèle permet de traduire le coût généralisé pour effectuer un déplacement entre deux stations. Ainsi, chaque paire Origine-Destination se voit dotée d'un coût de friction, un niveau d'attractivité. Cette fonction peut avoir plusieurs formes, mais la version utilisée dans cette recherche est l'exponentielle :

$$f(c_{ij}) = \exp(-b * c_{ij}) \quad 4.4$$

Dans l'équation 4.4,  $c$  est le coût entre les stations  $i$  et  $j$ , qui peut être soit une distance ou un temps, alors que le paramètre  $b$  détermine l'ajustement de la fonction, tel qu'illustré à la figure 4.11.

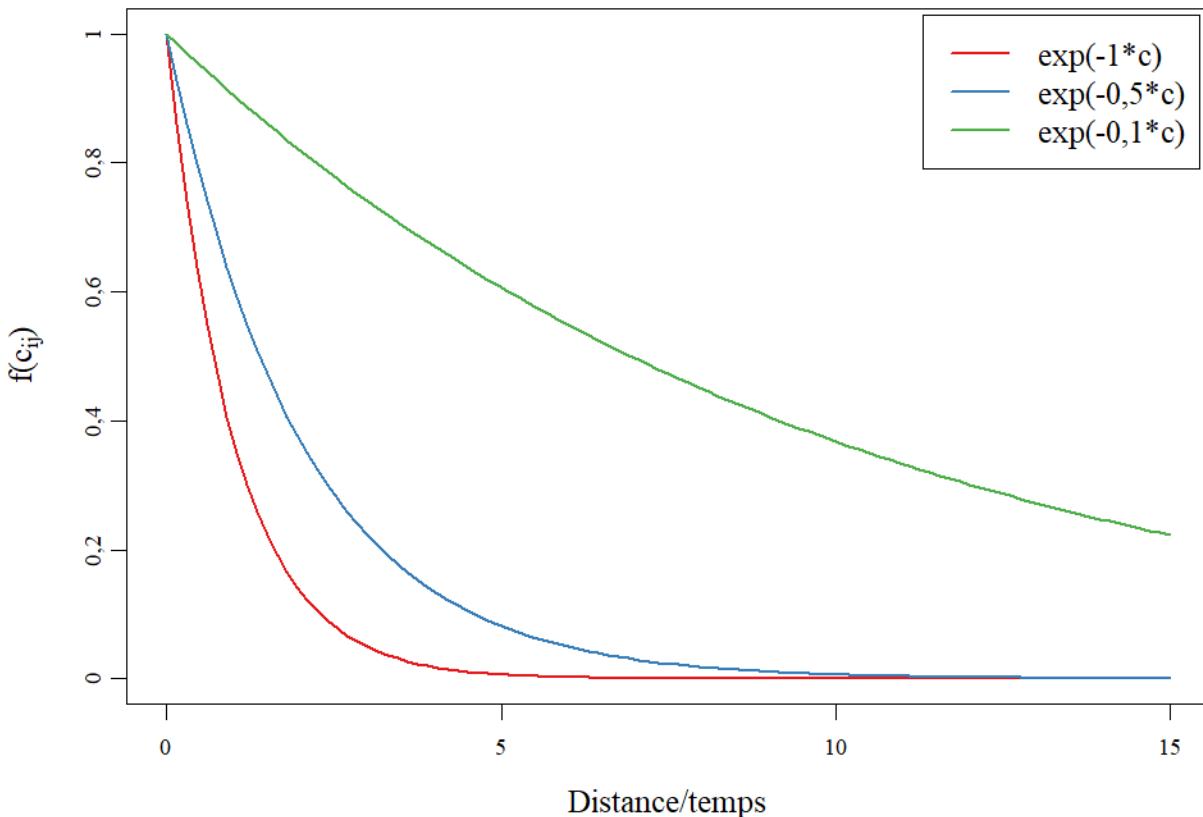


Figure 4.11 : Différentes fonctions de coût

Le modèle de distribution est calibré selon la valeur du coefficient  $b$ . Alors que celui-ci peut être « estimé » à partir de certains tests, il est également possible de le calculer de manière itérative grâce à la distance (ou durée) moyenne des déplacements observés (DMDO). Ce processus est démontré par Hyman (1969) et repris par Ortuzar et Willumsen (2011). La DMDO,  $c^*$ , est calculée avec le nombre de déplacements observés entre chaque paire de stations,  $N_{ij}$  :

$$c^* = \sum_{ij} (N_{ij} c_{ij}) / \sum_{ij} N_{ij} \quad 4.5$$

Cet indicateur permet d'entamer le processus itératif afin de trouver le paramètre  $b$  idéal pour que la distance moyenne des déplacements modélisés (DMDM) s'ajuste sur la DMDO. L'algorithme itératif se déroule selon étapes suivantes :

- 1- à la première itération,  $m = 0$  et un estimé initial de  $b_0$  se fait avec  $1/c^*$ ;

2- cette valeur  $b_0$  permet ensuite de calculer une première matrice de déplacements avec le modèle de distribution (équation 4.1). Par la suite, la DMDM est calculée, soit  $c_0$ , pour ensuite déterminer une meilleure valeur à  $b$  ainsi :

$$b_m = b_0 c_0 / c^* \quad 4.6$$

3- le paramètre  $m$  devient alors  $m + 1$ . À l'aide du dernier paramètre  $b$  ( $b_{m-1}$ ), une nouvelle matrice est calculée, ce qui permet d'obtenir un nouveau DMDM  $c_{m-1}$  et de le comparer avec  $c^*$ . S'ils sont suffisamment rapprochés, selon la proportion désirée ( $\pm 0,05\%$  dans le cas présent), le paramètre  $b_{m-1}$  devient le meilleur estimé, sinon poursuivre avec la prochaine étape ;

4- le calcul d'un meilleur estimé de  $b$  de cette façon :

$$b_{m+1} = \frac{(c^* c_{m-1}) b_m - (c^* - c_m) b_{m-1}}{c_m c_{m-1}} \quad 4.7$$

5- les étapes 3 et 4 sont répétées jusqu'à ce que la DMDM et donc le coût  $c_{m-1}$  soit suffisamment proche de la valeur observée  $c^*$ , soit la DMDO.

Ces deux étapes, soit les facteurs d'ajustements et le paramètre décroissant de la fonction de coût sont effectuées de manière successive pour calibrer le modèle gravitaire. Ainsi, le paramètre  $b$ , selon le type de fonction, permet de bien reproduire la courbe des déplacements observés selon le coût.

#### 4.4.2 Problématique de la diagonale

La diagonale de la matrice possède une valeur de 0. Effectivement, un déplacement ayant la même station d'Origine et de Destination possède une distance et une durée théorique nulle. Cependant, afin de donner un poids à ces paires O-D, il est nécessaire de leur attribuer une valeur fictive, car il est inutile de multiplier le paramètre  $b$  avec 0. Plusieurs tests ont été effectués et il a été décidé de donner la valeur du 1<sup>er</sup> quartile de l'ensemble des coûts à la diagonale. Par exemple, en 2011, 25 % de la distance réseau entre les stations est de 2 752 m et c'est donc cette même valeur qui est insérée dans les cellules où l'Origine et la Destination sont identiques.

### 4.4.3 Développement des modèles

Ainsi plusieurs tests de calibration ont été effectués selon divers paramètres.

- Type d'usager : membres et occasionnels ;
- Temporalité : somme des déplacements en pointe matinale pour l'ensemble de l'année, somme des déplacements en pointe matinale en juillet ;
- Coût : distance réseau sans et avec pentes, durée selon distance réseau sans et avec pentes ;
- Fonction de coût : exponentielle, puissance.

Certaines variantes sont moins performantes que d'autres. Par exemple, les déplacements pour les usagers occasionnels sont beaucoup plus difficiles à calibrer (c'est-à-dire qu'ils nécessitent beaucoup plus de temps) puisque moins nombreux pendant la période matinale. Ainsi, afin de limiter les résultats, seules quelques variantes sont présentées : il s'agit des modèles calibrés sur les déplacements effectués par les membres en pointe matinale lors de jours ouvrables en 2011 et en 2014. Les quatre types de coûts présentés précédemment sont utilisés et présentés dans les résultats et seule la fonction de coût exponentielle est retenue. Les résultats de ces huit modèles sont présentés dans la section suivante.

### 4.4.4 Résultats de la modélisation

Seuls les deux modèles avec la distance réseau entre stations comme coût sont présentés en guise d'analyse approfondie. L'ensemble des résultats des modèles sont comparés à la fin de cette section grâce à cinq indicateurs de performance. À titre indicatif, les deux fonctions de coût calibrées des modèles sont présentées à la figure 4.12 avec le paramètre  $b$  en légende. Les courbes et le paramètre sont identiques ce qui laisse présager peu de différences dans les résultats de la modélisation.

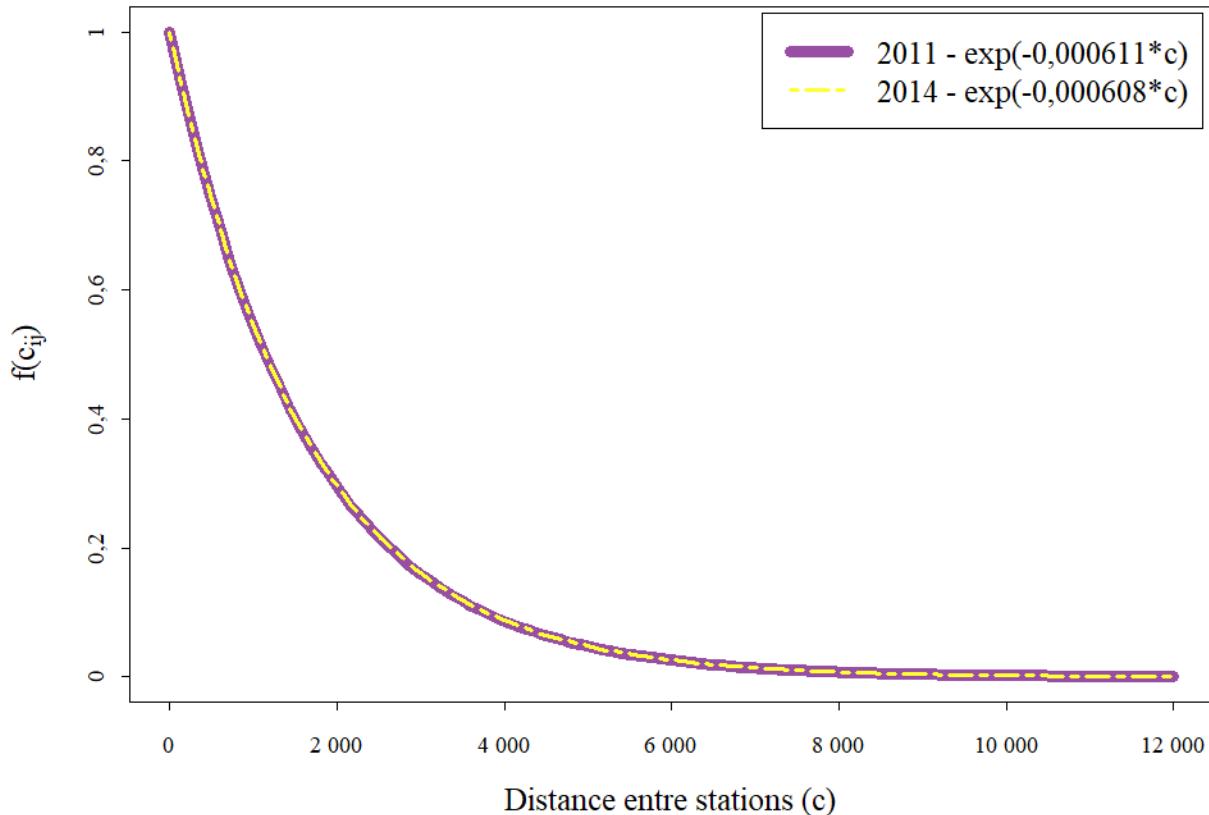


Figure 4.12 : Fonctions de coût calibrées

#### 4.4.4.1 Le concept de résidus

Il est primordial de définir la notion de résidu, qui guide l'analyse des résultats des modèles. Un résidu est la différence entre une valeur obtenue par le modèle et une valeur observée, ce qui est schématiquement représenté à la figure 4.13. Cette équation est utilisée afin de simplifier la représentation des valeurs sous-estimées et surestimées. Le modèle parfait ne comporterait aucun résidu, c'est-à-dire que ceux-ci seraient tous à 0 et la technique reproduirait parfaitement la réalité.

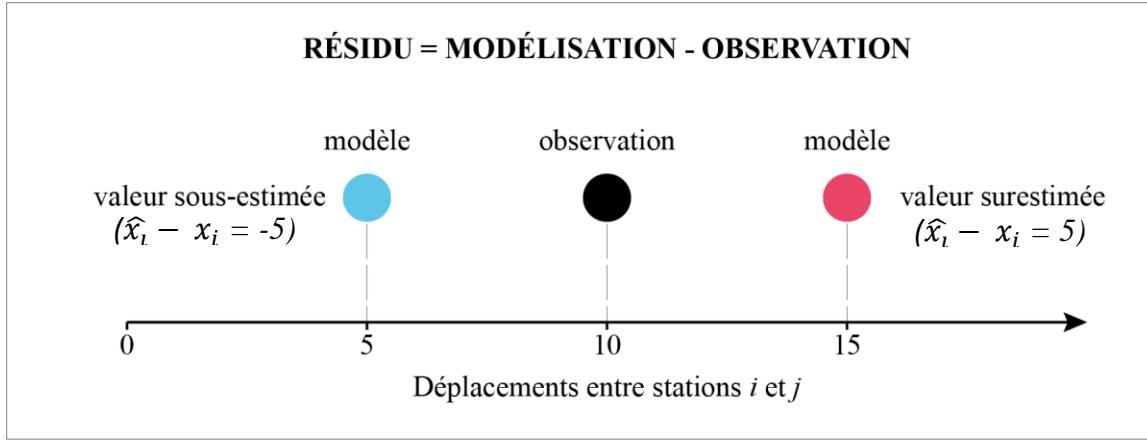


Figure 4.13 : Schéma des types de résidus

#### 4.4.4.2 Analyse des résidus des modèles 2011 et 2014

À première vue, les résultats des modèles reproduisent bien les cas observés, comme illustrés dans la figure 4.14. En revanche, il est possible de constater que les modèles ne réussissent pas à atteindre le nombre cumulatif de déplacements observés dans les deux années.

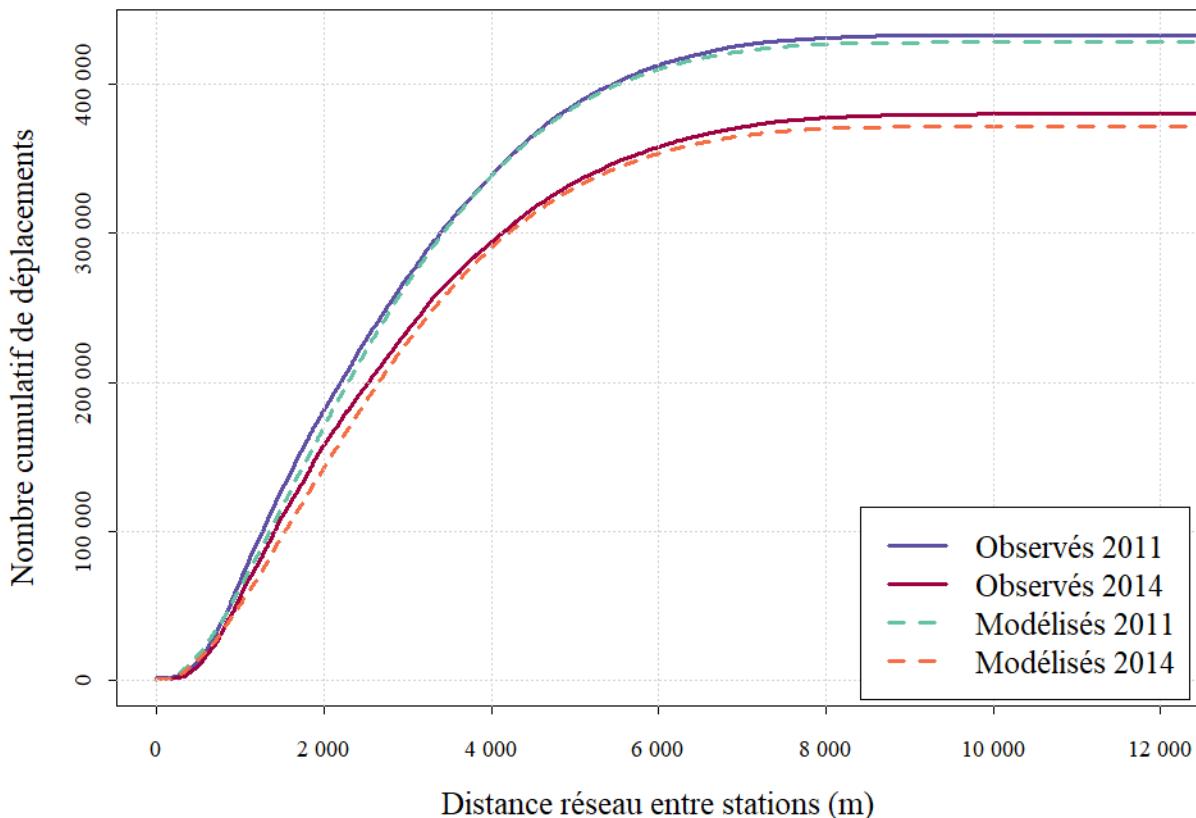


Figure 4.14 : Cumulatif des déplacements observés et modélisés

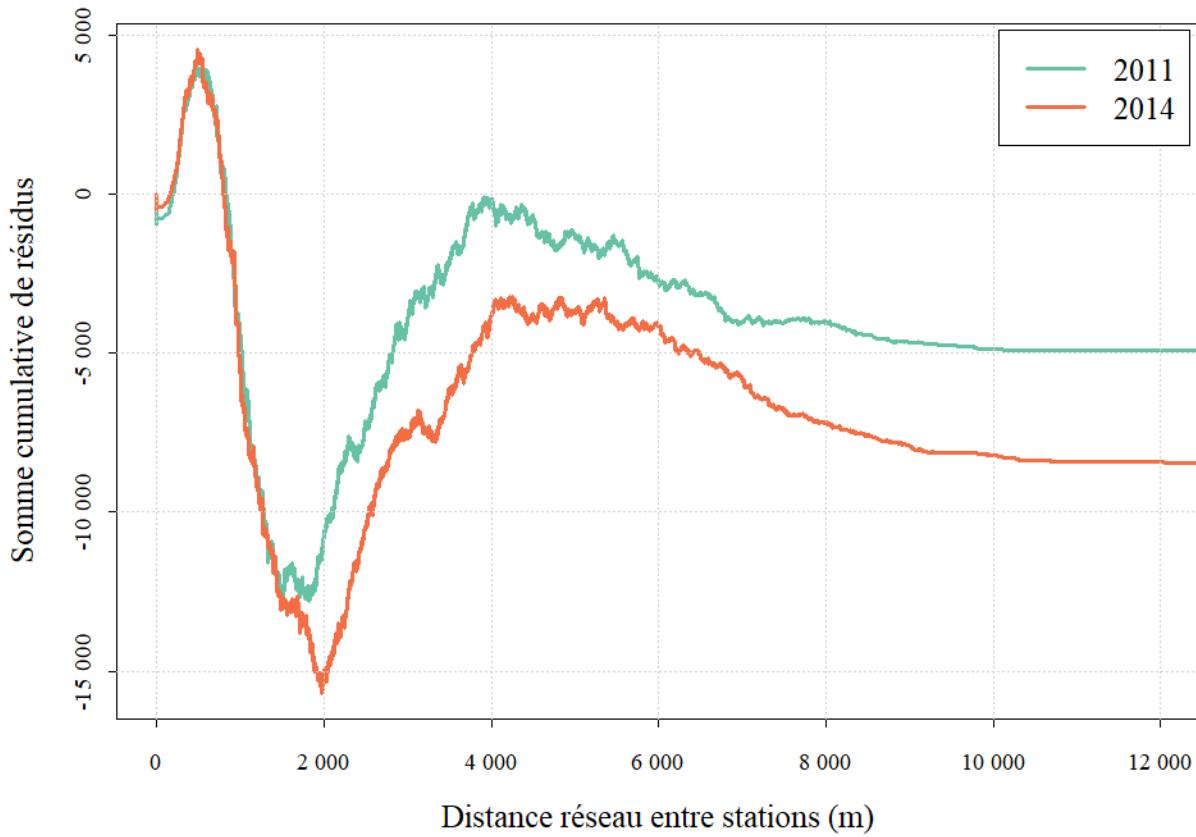


Figure 4.15 : Cumulatif des résidus des modèles 2011 et 2014

Ceci est plus visible en observant la courbe de la somme cumulative des résidus des deux modèles, à la figure 4.15. Les valeurs modélisées tendent à surestimer les déplacements lorsque la distance entre deux stations est très courte (sous 1 km) et l'inverse se produit par la suite, jusqu'à un certain plafonnement négatif.

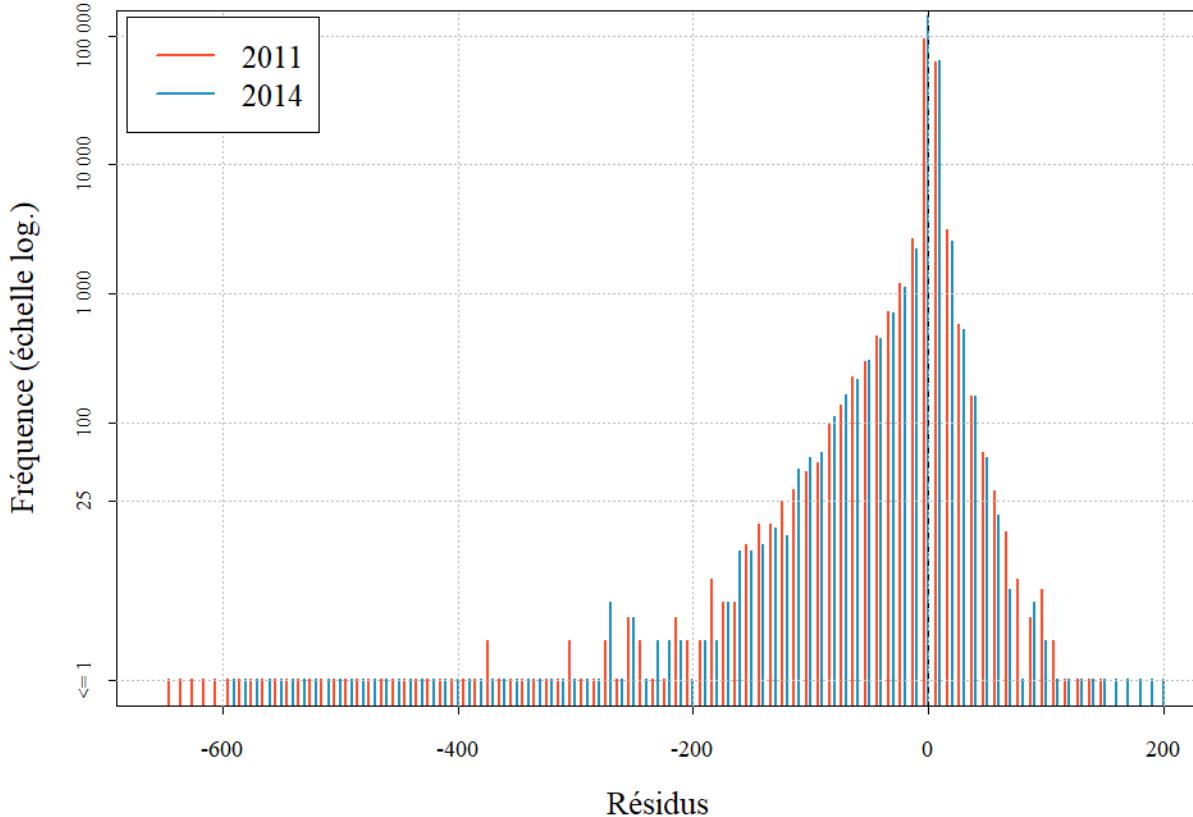


Figure 4.16 : Distribution des résidus (nombre de déplacements prédis)

De manière similaire, la figure 4.16 illustre la distribution de ces résidus pour les deux modèles sous analyse. Il est possible de constater une grande concentration de résidus autour de 0, ce qui est positif, et il est important de souligner que l'axe des ordonnées suit une échelle logarithmique. La figure démontre également que les modèles ont une certaine tendance à sous-estimer les déplacements.

Ces différentes tendances sont présentées sous forme statistique au tableau 4.4. Il est possible d'y constater une grande variabilité dans les résidus selon l'écart-type et encore plus pour la somme de ceux aux stations de Destination (écart-type de 38,01 et 33,14). Il est possible que ce grand écart des marges modélisées par rapport aux déplacements observés soit dû à l'arrondissement des facteurs d'ajustement  $\alpha_i$  et  $\beta_j$  du modèle gravitaire, puisque ces facteurs devraient théoriquement garantir la somme des flux modélisés aux stations sur la matrice observée. La somme importante de résidus absolu s'explique simplement par le grand nombre d'observations alors que le 3<sup>e</sup> quartile de ceux-ci se situe à 1 et le maximum à 552.

Tableau 4.4 : Statistiques sommaires des résidus

|  | 2011       |         |            | 2014       |         |            |
|--|------------|---------|------------|------------|---------|------------|
|  | Somme      | Moyenne | Écart-type | Somme      | Moyenne | Écart-type |
| <b>Résidus</b>                         | -4 930,00  | -0,03   | 9,46       | -8 450,00  | -0,04   | 8,03       |
| <b>Résidus absolus</b>                 | 454 064,00 | 2,77    | 9,05       | 430 732,00 | 2,04    | 7,77       |
| <b>Résidus totaux à l'Origine</b>      | -4 930,00  | -12,17  | 9,02       | -8 450,00  | -18,41  | 8,14       |
| <b>Résidus totaux à la Destination</b> | -4 930,00  | -12,17  | 38,01      | -8 450,00  | -18,41  | 33,14      |

Ainsi, il est nécessaire de diagnostiquer davantage les résidus et de les mettre en relation avec d'autres éléments afin de tenter de repérer certaines tendances. En ce sens, la somme des résidus à l'Origine et à la Destination du modèle 2011 est représentée graphiquement à la figure 4.17 et ce, en fonction de la distance de la station par rapport au centre-ville. Le modèle a nettement plus de difficulté à bien prédire les déplacements effectués vers des stations à proximité du centre-ville.

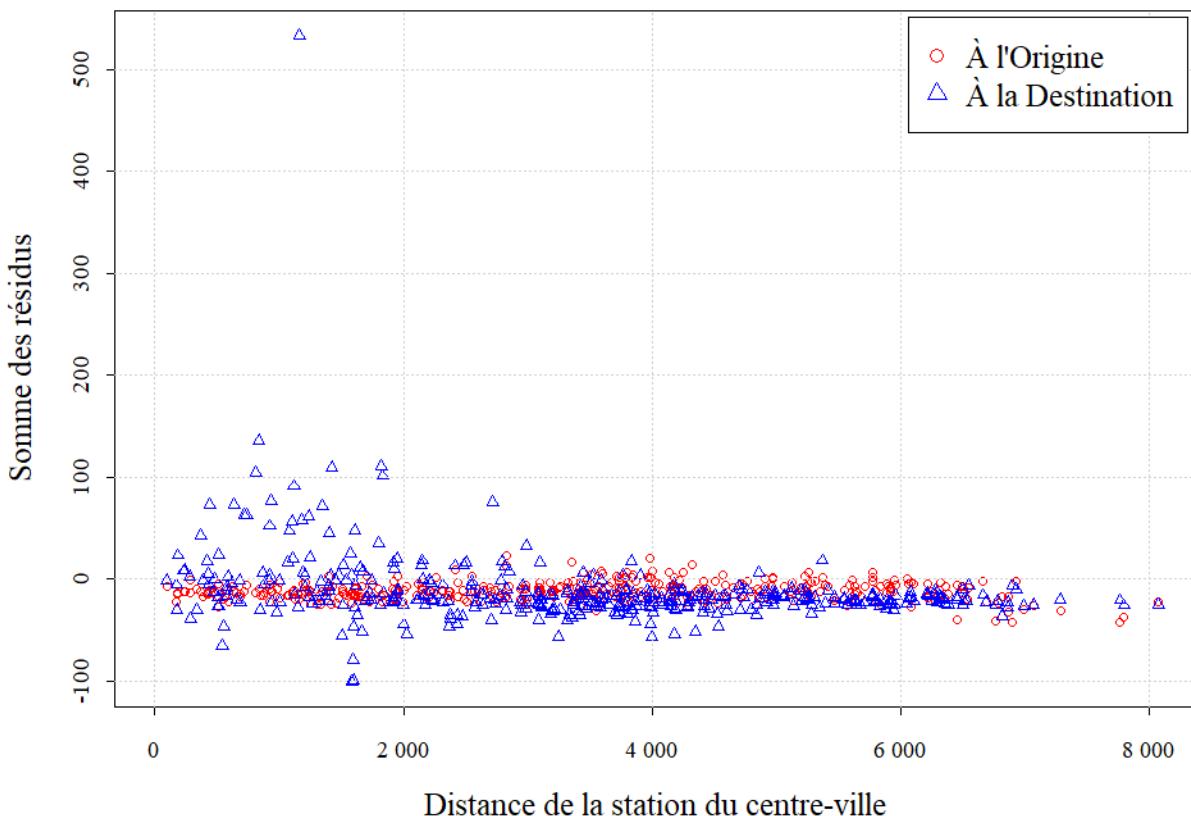


Figure 4.17 : Somme des résidus du modèle 2011 selon la distance de la station au centre-ville

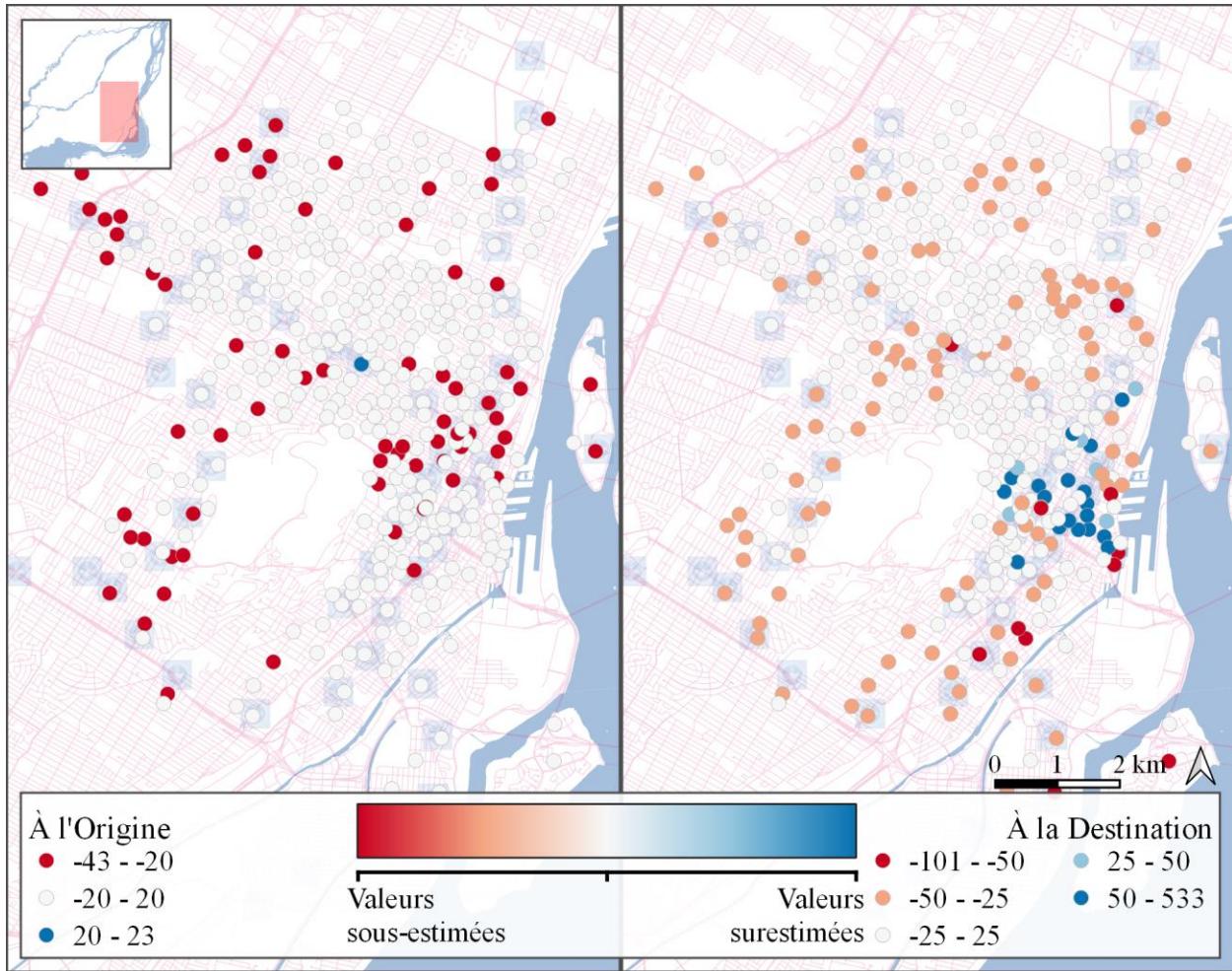


Figure 4.18 : Spatialisation de la somme des résidus du modèle 2011

C'est également ce qu'il est possible de voir grâce à la spatialisation des résidus à la figure 4.18 et figure 4.19. Les modèles sous-estiment les déplacements à l'Origine pour plusieurs stations, mais surestiment très peu. À l'inverse, il y a davantage de variabilité pour les stations de Destination puisque les résidus sont plus importants, qu'ils soient positifs ou négatifs. En ce sens, il faut raffiner le modèle pour traduire l'importance des stations près du centre-ville.

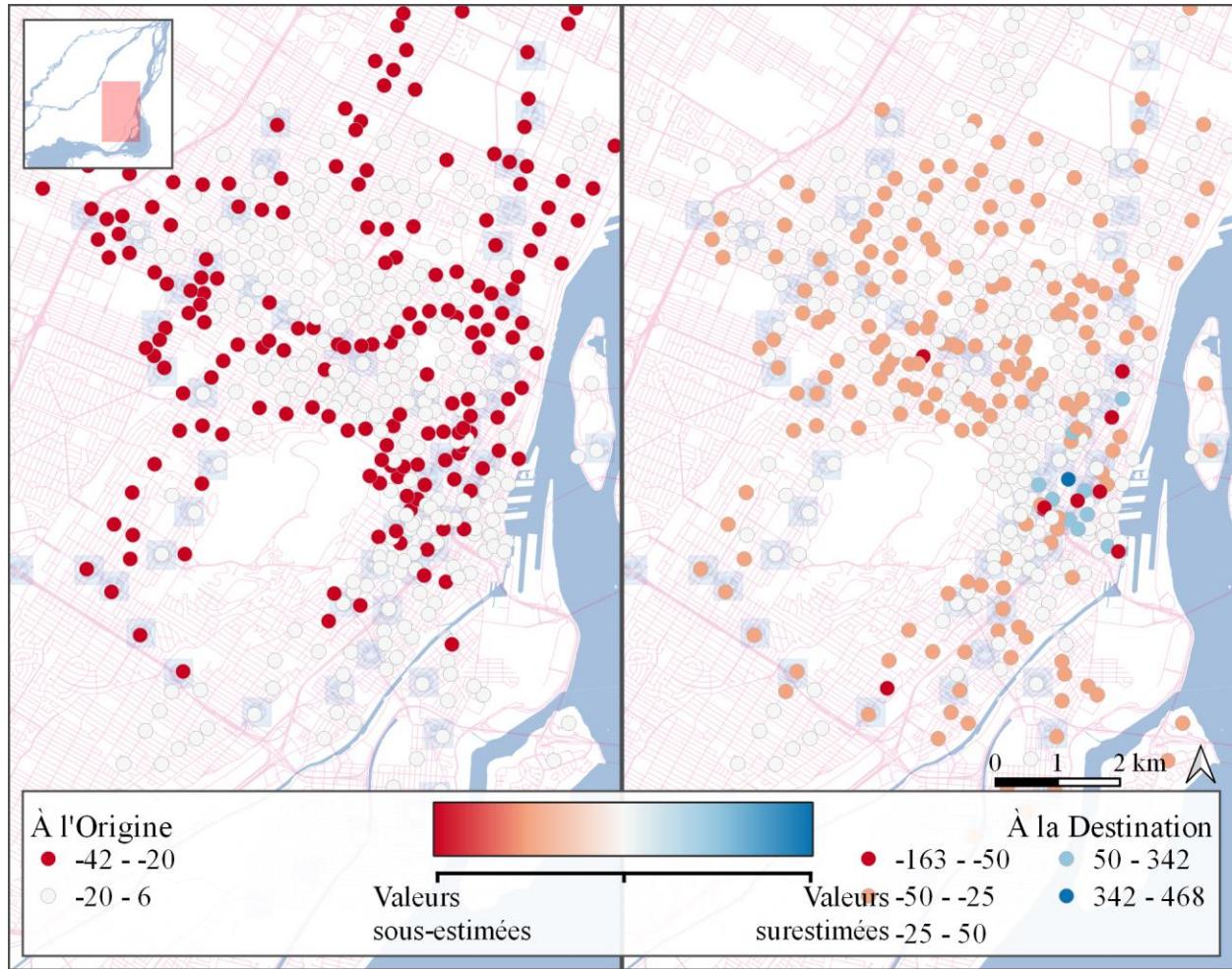


Figure 4.19 : Spatialisation de la somme des résidus du modèle 2014

Les paires O-D les moins bien estimées par les modèles sont représentées à la figure 4.20, soit les 20 paires les plus surestimées et les plus sous-estimées. À première vue, les lignes de désir près de certaines stations de métro font partie de celles moins bien modélisées. Ceci s'explique nécessairement parce que c'est souvent ces mêmes paires O-D qui attirent le plus de déplacements, tels que vus précédemment. Il est aussi intéressant de noter que les lignes de désir les plus surestimées ont plusieurs Origines et Destinations communes et ce, pour les deux années, alors que les paires O-D les plus sous-estimées sont plutôt distinctes.

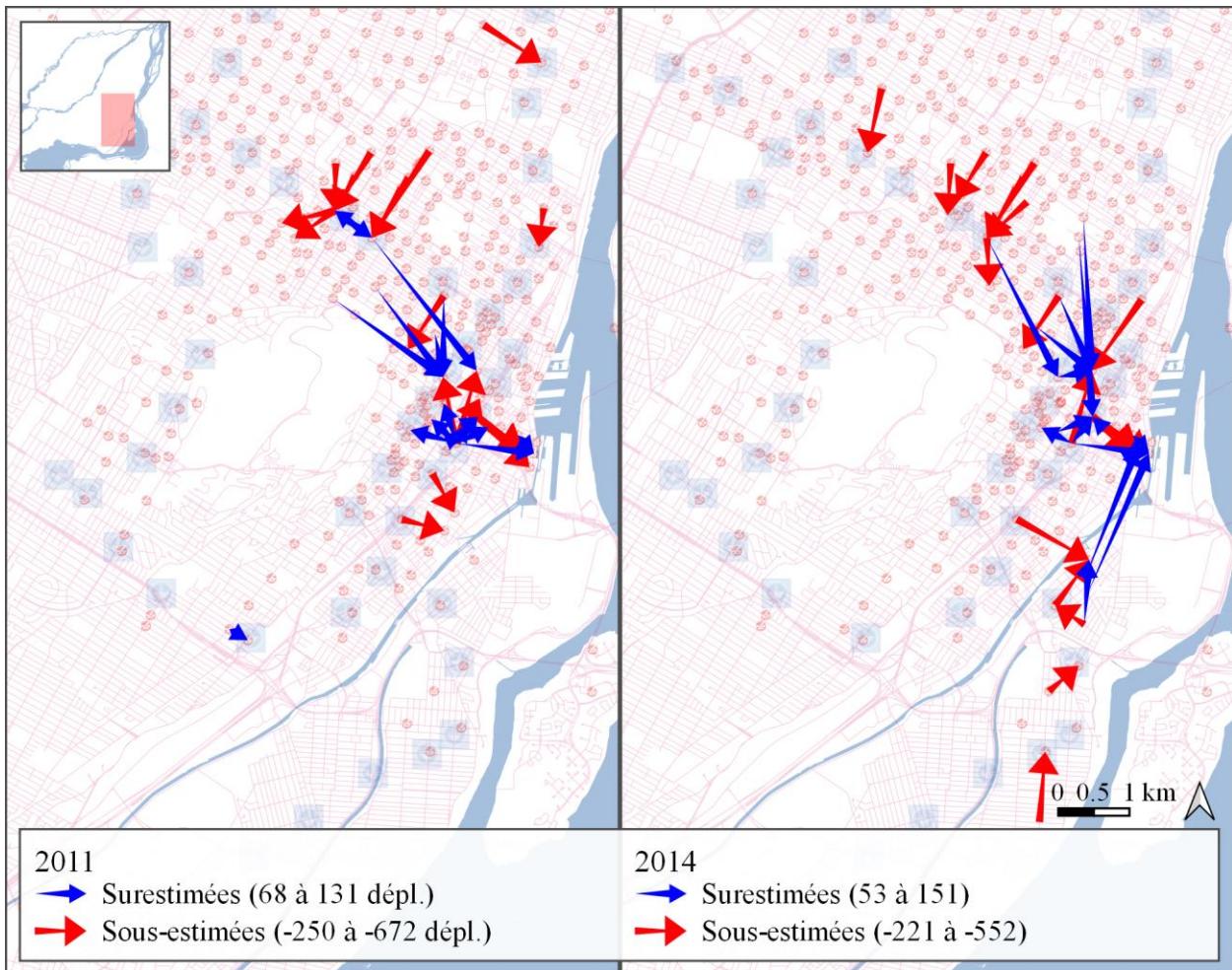


Figure 4.20 : Spatialisation des paires O-D les moins bien estimées

Puisqu'il est très difficile de visualiser l'ensemble des paires O-D, une typologie des paires O-D selon la capacité des stations est utilisée. Cette typologie se base sur la capacité de la station à l'Origine et à la Destination que celle-ci soit petite, moyenne ou grande. Pour ce faire, chaque station dont la capacité est inférieure au 33<sup>e</sup> percentile de l'ensemble de la capacité des stations est

considérée comme étant petite, et celles dont la capacité est supérieure au 66<sup>e</sup> percentile sont réputées grandes; les autres sont ainsi des stations avec capacité moyenne. Chaque paire O-D est donc classée au sein de neuf groupes, selon la taille de la station à l'Origine et à la Destination, comme montré au tableau 4.5. Pour chaque groupe, il contient une brève description de la typologie, le pourcentage de paires O-D, les déplacements observés et modélisés.

Tableau 4.5 : Distribution des déplacements en pourcentage selon le type de paires O-D

| Gr. | Cap. O → D<br><i>P = petite</i><br><i>M = moyenne</i><br><i>G = grande</i> | 2011          |                   |                    | 2014          |                   |                    |
|-----|--|---------------|-------------------|--------------------|---------------|-------------------|--------------------|
|     |  | Paires<br>O-D | Dépl.<br>observés | Dépl.<br>modélisés | Paires<br>O-D | Dépl.<br>observés | Dépl.<br>modélisés |
| 1   | P → P  | 22,71         | 9,25              | 9,01               | 15,90         | 6,03              | 6,05               |
| 2   | P → M  | 6,59          | 4,16              | 4,19               | 10,25         | 7,03              | 6,83               |
| 3   | P → G  | 18,36         | 25,50             | 25,59              | 13,72         | 19,21             | 19,17              |
| 4   | M → P  | 6,59          | 2,72              | 2,48               | 10,25         | 5,13              | 4,75               |
| 5   | M → M  | 1,91          | 1,63              | 1,44               | 6,61          | 5,40              | 5,36               |
| 6   | M → G  | 5,33          | 9,90              | 10,32              | 8,85          | 16,04             | 16,48              |
| 7   | G → P  | 18,36         | 7,54              | 7,33               | 13,72         | 6,47              | 6,11               |
| 8   | G → M  | 5,33          | 4,27              | 4,36               | 8,85          | 7,36              | 7,38               |
| 9   | G → G  | 14,84         | 35,02             | 35,29              | 11,85         | 27,33             | 27,87              |

Il est possible de voir une évolution non seulement dans les déplacements de 2014 par rapport à 2011, mais également dans le nombre de paires O-D. Effectivement, la capacité des stations du système est mieux balancée dans la deuxième période observée puisque les variations entre les paires sont moins grandes.

Ensuite, les déplacements qui se dirigent vers une grande station (groupes 3, 6 et 9) sont les plus nombreux et ce, en 2011 et en 2014. Les déplacements où l'Origine et la Destination comprennent une grande capacité (groupe 9) sont particulièrement nombreux, dépassant le tiers des observations en 2011 et 27,34 % en 2016. De plus, la proportion des déplacements modélisés dans chaque groupe est similaire à ce qui est observé.

Dans la même optique, les résidus sont également analysés selon cette typologie au tableau 4.6. Les valeurs sont absolues et les 0 sont exclus. Le tableau permet de constater que, encore une fois, les résidus les plus importants et qui ont davantage de variation sont ceux aux groupes 3, 6 et particulièrement 9, soit les paires O-D qui comprennent au moins une grande station.

Tableau 4.6 : Statistiques sommaires des résidus absolu (excluant 0) selon la capacité

| Gr. | Cap. O → D<br><i>P = petite</i><br><i>M = moyenne</i><br><i>G = grande</i> | 2011       |         |            | 2014       |         |            |
|-----|--|------------|---------|------------|------------|---------|------------|
|     |  | Somme      | Moyenne | Écart-type | Somme      | Moyenne | Écart-type |
| 1   | P → P  | 50 277,00  | 3,17    | 6,69       | 30 574,00  | 3,17    | 7,09       |
| 2   | P → M  | 21 713,00  | 3,92    | 8,10       | 31 683,00  | 4,15    | 10,25      |
| 3   | P → G  | 118 570,00 | 5,67    | 11,18      | 82 783,00  | 5,31    | 10,73      |
| 4   | M → P  | 14 656,00  | 3,51    | 9,21       | 24 679,00  | 3,65    | 9,58       |
| 5   | M → M  | 7 573,00   | 4,76    | 9,00       | 24 389,00  | 4,43    | 10,33      |
| 6   | M → G  | 42 376,00  | 6,86    | 12,86      | 68 265,00  | 6,11    | 13,52      |
| 7   | G → P  | 38 168,00  | 3,87    | 8,61       | 29 800,00  | 3,81    | 8,98       |
| 8   | G → M  | 20 733,00  | 5,08    | 11,59      | 31 256,00  | 4,83    | 11,20      |
| 9   | G → G  | 139 998,00 | 8,44    | 18,17      | 107 303,00 | 7,41    | 15,50      |

Un exercice similaire a été effectué en classant les paires O-D selon la distance des stations par rapport au centre-ville. La classification a été faite de la même manière (avec le 33<sup>e</sup> et 66<sup>e</sup> percentile de l'ensemble des distances au centre-ville) en neuf groupes, selon une distance rapprochée, moyenne, ou éloignée au centre-ville. Alors que les statistiques sommaires des résidus selon cette typologie sont présentées seulement à l'annexe F, il est intéressant de souligner que les plus grandes fluctuations dans ce cas sont dans les groupes avec une station près du centre-ville. Ces deux typologies réitèrent donc la difficulté pour le modèle gravitaire dans sa forme basique de reproduire les déplacements vers de grandes stations et à proximité du centre-ville. Il y a nécessairement possibilité d'améliorer le modèle en incluant certaines de ses informations.

#### **4.4.4.3 Analyse de l'environnement spatial autour des stations des paires O-D**

Ainsi, en guise de procédure exploratoire, une tentative d'analyse spatiale des résidus est effectuée grâce à des variables comme l'usage du sol à l'Origine et à la Destination des déplacements dans une régression linéaire multiple. Les différents usages du sol constituent une partie des variables indépendantes dans le modèle et proviennent de la CMM, ce qui a été expliqué dans les modèles de génération à la section 3.2.1.2. Cet usage, pour les déplacements de 2011, est celui de 2012, alors que pour les déplacements de 2014, les données d'utilisation du sol sont de la même année. La somme de la superficie des différents usages du sol suivants est calculée en kilomètre carré dans une zone tampon de 400 m autour de chaque station : résidentiel de 1 logement; résidentiel de plus de 1 logement (somme des usages de 2 à 4 logements, 5 à 24 logements et 25 logements et plus); industriel et institutionnel économique et non économique. D'autres types d'usages qui risqueraient de traduire des éléments typiques qu'on retrouve au centre-ville comme commercial et bureau sont trop corrélés avec d'autres variables ce qui signifie qu'ils ne sont pas considérés. L'utilisation du sol est également accompagnée d'autres informations, soit la capacité, le nombre de stations avoisinantes et le nombre de stations de métro (dans un rayon de 400 m) ainsi que l'altitude.

En guise de variable dépendante, le ratio des déplacements observés sur les déplacements modélisés est utilisé. Si le nombre de déplacements observés pour une paire O-D est positif et que le nombre de déplacements modélisés est nul, ce ratio est porté à 1. Lorsque les déplacements observés et modélisés sont nuls, le ratio indiqué est 0. Un ratio entre 0 et 1 signifie que le modèle surestime les déplacements, alors qu'un ratio supérieur à 1 signifie une sous-estimation.

Finalement, le logarithme négatif de ce ratio est utilisé afin que sa distribution ressemble à une courbe normale et puis normalisé entre 0 et 1 afin de n'avoir que des valeurs positives.

Tableau 4.7 : Coefficients de régression linéaire des résidus

|                         | 2011  | 2014                          |
|-------------------------|---|-------------------------------|
| <i>À l'Origine</i>      | Intercept   | -0,50*** (0,005)              |
|                         | Capacité  | <b>0,001***</b> (0,0001)      |
|                         | Nombre de stations avoisinantes                   | <b>0,002***</b> (0,0003)      |
|                         | Nombre de stations de métro                       | <b>-0,004***</b> (0,001)      |
|                         | Élévation (m)                                     | <b>-0,0001**</b> (0,0000)     |
|                         | Usage résidentiel (1 logement, km <sup>2</sup> )  | 0,0000 (0,0000)               |
|                         | Usage rés. (plus d'un logement, km <sup>2</sup> ) | <b>0,0001***</b> (0,0000)     |
|                         | Usage industriel (km <sup>2</sup> )               | <b>0,0001**</b> (0,0000)      |
|                         | Usage inst. économique (km <sup>2</sup> )         | <b>-0,0001***</b> (0,0000)    |
|                         | Usage inst. non-écon. (km <sup>2</sup> )          | 0,0001* (0,0001)              |
| <i>À la Destination</i> | Capacité  | <b>0,001***</b> (0,0001)      |
|                         | Nombre de stations avoisinantes                   | <b>0,001***</b> (0,0002)      |
|                         | Nombre de stations de métro                       | <b>0,01***</b> (0,001)        |
|                         | Élévation (m)                                     | -0,0000 (0,0000)              |
|                         | Usage résidentiel (1 logement, km <sup>2</sup> )  | <b>-0,0002***</b> (0,0000)    |
|                         | Usage rés. (plus d'un logement, km <sup>2</sup> ) | <b>-0,0001***</b> (0,0000)    |
|                         | Usage industriel (km <sup>2</sup> )               | -0,0001* (0,0000)             |
|                         | Usage inst. économique (km <sup>2</sup> )         | <b>-0,0001***</b> (0,0000)    |
|                         | Usage inst. non-écon. (km <sup>2</sup> )          | -0,0000 (0,0001)              |
|                         | R <sup>2</sup>                                    | 0,03                          |
| <i>Significativité</i>  |   | *p<0,1 ; **p<0,05 ; ***p<0,01 |

Les résultats de la régression linéaire multiple sont présentés au tableau 4.7. Il est possible d'y voir l'impact significatif (en gras lorsque 95 % et plus) de nombreuses variables. Un coefficient négatif signifie une augmentation de la sous-estimation alors que l'inverse s'explique par un coefficient positif. Le nombre de stations de métro est la variable qui semble avoir l'impact le plus important de toutes sur l'estimation, négatif à l'Origine et positif à la Destination, constant selon l'année. Le nombre de stations avoisinantes tend également à faire augmenter l'estimation, tout comme l'usage résidentiel de plus de 1 logement à l'Origine. À la Destination, cette variable fait diminuer l'estimation. L'ensemble de ces variables constitue une piste de solution d'une manière qui pourrait améliorer le modèle de distribution dans des itérations subséquentes.

## 4.4.5 Validation des modèles

Après l'analyse des résidus, il est question d'utiliser des indicateurs de performance afin de valider les modèles. Le modèle gravitaire a été calibré sur les déplacements de deux années, soit 2011 et 2014. Afin de valider ces modèles, le paramètre  $b$  respectif de chaque modèle a été appliqué aux déplacements observés en l'an 2016. Quelques indicateurs de validation qui sont utilisés afin de comparer les résultats des modèles sont d'abord présentés avant de conclure avec les calculs des différents modèles.

### 4.4.5.1 Indicateurs de validation

La revue de littérature révèle l'existence de nombreux indicateurs permettant de valider un modèle de distribution par rapport aux flux observés. Les cinq indicateurs retenus dans le cas présent proviennent des auteurs suivants : Thomas (1977); Smith et Hutchinson (1981); Sherali, Narayanan et Sivanandan (2003); Ricci (2005) et Nazem, Trépanier et Morency (2013). Les formules des indicateurs sont présentées ci-dessous, où  $x_i$  est le nombre de déplacements observés pour la paire O-D  $i$ ,  $\hat{x}_i$  est le nombre de déplacements modélisés pour cette même paire O-D,  $N$  le nombre de paires O-D et  $T$  le nombre total de déplacements observés :

- le coefficient de détermination ( $R^2$ ) a été présenté comme indicateur de validation de la régression linéaire à l'équation 3.6. Celui-ci varie entre 0 et 1, et plus la valeur est élevée, plus le modèle prédit bien les valeurs. Lorsque ce coefficient est négatif, ceci signifie que le modèle ne suit pas une tendance linéaire et que le  $R^2$  n'est pas nécessairement approprié en tant que coefficient pour évaluer la modélisation. De ce fait, cet indicateur n'est pas le plus approprié pour évaluer ces modèles, mais il est présent dans la littérature donc il est calculé tout de même ;
- la statistique PHI =  $\sum(x_{ij} \left| \log \frac{x_i}{\hat{x}_i} \right|)$ ;
- la moyenne des résidus absolus normalisée (MRA) =  $\frac{\sum_i |x_i - \hat{x}_i| / N}{T / N}$ ;
- l'écart-type des résidus quadratiques normalisé (ERQ) =  $\frac{\sqrt{\sum_i (x_i - \hat{x}_i)^2 / N}}{x_{max} - x_{min}}$ .

- le pourcentage mal attribué (PMA) =  $\frac{50}{T} \sum_i |x_i - \hat{x}_i|$ . Cet indicateur, en pourcentage, identifie le nombre de flux modélisés qui devraient être réassignés à d'autres paires O-D pour reproduire les flux observés.

Outre le  $R^2$ , une valeur plus faible pour l'ensemble des indicateurs signifie un meilleur modèle.

#### 4.4.5.2 Résultats des modèles selon les indicateurs

Les indicateurs sont calculés pour chaque variante des modèles, soit selon les quatre types de coûts : distance réseau avec et sans dénivelé; durée de déplacements sur distance réseau avec et sans dénivelé. Les autres paramètres ne sont pas modifiés, à savoir, les déplacements effectués par les membres lors de la pointe matinale des jours ouvrables avec une fonction de coût exponentielle. Ces modèles calibrés sont ensuite appliqués sur les déplacements observés en 2016 et les indicateurs sont calculés également. Le tout est présenté au tableau 4.8.

Règle générale, les modèles calibrés en 2011 et 2014 appliqués aux déplacements de 2016 fonctionnent moins bien que sur les années de référence. Ceci est logique puisque la calibration permet une plus grande cohérence avec la distribution des déplacements. Les paramètres des deux années de référence qui utilisent une distance réseau en tant que coût, appliqué à l'an 2016, fonctionnent de manière quasi identique selon l'ensemble des indicateurs. En revanche, lorsque le dénivelé est inclus dans le calcul du coût entre les stations dans les modèles, ceux-ci ne sont plus performants. Il se peut que cet élément fonctionne mieux avec les déplacements en 2014 et en 2016, ce qui corrobore l'hypothèse de stabilisation du système.

Finalement, le tableau 4.9 consiste à présenter la répartition des résidus en pourcentage selon les modèles. Ainsi, sont notés le nombre de résidus à 0 (l'absence de résidus), ceux qui sont positifs (valeurs surestimées), négatifs (valeurs sous-estimées), et ceux qui se retrouvent entre -10 et 10. Pour l'ensemble des modèles, l'absence de résidus varie de 48 à 60 %, alors que les modèles 2016 qui se basent sur les paramètres calibrés en 2011 fonctionnent mieux que ceux-ci, alors que l'inverse est vrai pour les modèles 2014. Les modèles tendent à sous-estimer les résidus, ce qui a été mentionné maintes fois pendant l'analyse des résultats, et ce, pour 31 à 40 % des déplacements. Contrairement à ce qui a été révélé par les indicateurs, les métriques de coût avec dénivelé semblent fonctionner davantage lorsque les paramètres de 2011 sont appliqués sur les observations de 2016 qu'avec ceux de 2014.

Tableau 4.8 : Résultats des modèles selon les indicateurs

|                      | Distance réseau |       |                |                | Distance réseau avec dénivelé |       |                |                | Durée selon distance réseau |       |                |                | Durée selon distance réseau avec dénivelé |       |                |                |
|----------------------|-----------------|-------|----------------|----------------|-------------------------------|-------|----------------|----------------|-----------------------------|-------|----------------|----------------|---|-------|----------------|----------------|
| Ind.                 | 2011            | 2014  | 2016<br>(2011) | 2016<br>(2014) | 2011                          | 2014  | 2016<br>(2011) | 2016<br>(2014) | 2011                        | 2014  | 2016<br>(2011) | 2016<br>(2014) | 2011                                      | 2014  | 2016<br>(2011) | 2016<br>(2014) |
| <b>R<sup>2</sup></b> | 0,295           | 0,265 | 0,265          | 0,266          | 0,296                         | 0,221 | -0,457         | 0,161          | 0,294                       | 0,266 | 0,267          | 0,267          | 0,294                                     | 0,210 | -1,125         | 0,135          |
| <b>PHI</b>           | 931,0           | 917,5 | 1200,2         | 1196,1         | 943,2                         | 950,7 | 175,0          | 1 109,8        | 935,0                       | 913,4 | 1 192,0        | 1 192,0        | 895,4                                     | 863,1 | 150,7          | 978,7          |
| <b>MRA</b>           | 1,049           | 1,133 | 1,131          | 1,131          | 1,050                         | 1,213 | 1,313          | 1,229          | 1,050                       | 1,132 | 1,131          | 1,131          | 1,060                                     | 1,240 | 1,392          | 1,254          |
| <b>ERQ</b>           | 0,014           | 0,012 | 0,012          | 0,012          | 0,014                         | 0,013 | 0,016          | 0,012          | 0,014                       | 0,012 | 0,011          | 0,011          | 0,014                                     | 0,013 | 0,020          | 0,012          |
| <b>PMA</b>           | 52,5            | 56,7  | 56,7           | 56,5           | 52,5                          | 60,67 | 65,6           | 61,4           | 52,4                        | 56,7  | 56,5           | 56,5           | 53,0                                      | 62,0  | 69,6           | 62,7           |

Tableau 4.9 : Pourcentage de résidus selon les modèles

|                   | Distance réseau |      |                |                | Distance réseau avec dénivélé |      |                |                | Durée selon distance réseau |      |                |                | Durée selon distance réseau avec dénivélé |      |                |                |
|-------------------|-----------------|------|----------------|----------------|-------------------------------|------|----------------|----------------|-----------------------------|------|----------------|----------------|---|------|----------------|----------------|
| Rés.              | 2011            | 2014 | 2016<br>(2011) | 2016<br>(2014) | 2011                          | 2014 | 2016<br>(2011) | 2016<br>(2014) | 2011                        | 2014 | 2016<br>(2011) | 2016<br>(2014) | 2011                                      | 2014 | 2016<br>(2011) | 2016<br>(2014) |
| <b>0</b>          | 48              | 60   | 56             | 56             | 48                            | 57   | 58             | 53             | 48                          | 60   | 56             | 56             | 49  | 56   | 59             | 53             |
| < 0               | 40              | 32   | 34             | 34             | 39                            | 35   | 32             | 38             | 39                          | 31   | 34             | 34             | 39  | 35   | 31             | 38             |
| > 0               | 12              | 9    | 9              | 9              | 12                            | 9    | 10             | 9              | 12                          | 9    | 9              | 9              | 12  | 9    | 10             | 9              |
| <b>[-10 - 10]</b> | 94              | 96   | 95             | 95             | 93                            | 95   | 94             | 94             | 94                          | 96   | 95             | 95             | 93  | 95   | 94             | 94             |

#### **4.4.6 Limites et perspectives**

La modélisation de la distribution des déplacements avec le modèle gravitaire étudié semble pertinente. Toutefois, il est important de mentionner les limites de cette étape de la PSC, plus particulièrement celles qui sont pertinentes à la présente étude, telles que présentées par Bonnel (2002).

D'abord, il y a la question des déplacements où la station à l'Origine et à la Destination est la même, soit la diagonale des matrices O-D. Effectivement, il a été question dans la section 4.4.2 de la distance fictive attribuée à ces cellules, car elles ne peuvent être laissées à 0. Or, ce coût est le même, peu importe l'emplacement de la station. Pourtant, les stations qui sont situées à proximité d'activités importantes et d'attrait touristiques variés risquent d'avoir un achalandage différent que des stations en périphérie situées dans un quartier résidentiel, par exemple. En ce sens, ce coût fictif ne permet pas de bien reproduire la réalité et a nécessairement un effet sur l'ensemble de la distribution des déplacements modélisés.

Ensuite, il est question d'une distance ou d'une durée en tant que seule métrique permettant de déterminer la distribution des flux. Effectivement, le modèle gravitaire tente de caler la fonction de coût sur la distribution des déplacements observés. Or, d'autres éléments entrent nécessairement en compte dans le choix d'une Destination que la distance à parcourir et ceci peut différer pour chaque personne. Une des pistes de solutions réside dans l'intégration de certains facteurs sociodémographiques ou encore de la qualification du réseau au modèle. Une discussion sur cette perspective est présentée dans la conclusion du chapitre. De plus, la fonction de coût peut prendre plusieurs formes alors qu'il n'est question que de la forme exponentielle dans le cas présent. Or, alors que des tests exploratoires ont été effectués avec d'autres types de courbes, soit la forme puissance et la forme combinée, cette dernière respectant davantage la distribution des déplacements selon la distance parcourue, les résultats préliminaires étaient similaires ou moins pertinents que ceux obtenus avec la forme exponentielle. Il est probable que de meilleurs résultats soient obtenus avec une variante de la fonction de coût. Cependant, calibrer ces fonctions est un long processus ce qui diminue la possibilité d'expérimenter plusieurs approches et paramètres.

Puis, comme c'était le cas dans les limites de l'étape de génération, les paramètres calés sont fixes dans le temps. Or, il y a de fortes chances qu'avec de nouvelles configurations du réseau, en plus

de l'évolution externe des lieux d'activités, la distribution des déplacements change également. Cet élément explique pourquoi les modèles ont été validées avec les données observées en 2016.

Finalement, la qualité des données utilisées possède également un impact sur les modèles. Heureusement, les systèmes de VLS permettent d'avoir une précision importante de l'ensemble des déplacements effectués avec ce mode. Un certain tri est réalisé avec la base de données des déplacements et il est possible d'avoir un bon portrait de la réalité. Comme le modèle gravitaire se base sur la distance ou la durée de déplacement entre chaque station, il vaut la peine de mentionner à nouveau que cette métrique est estimée. L'hypothèse est que l'ensemble des membres opte pour le chemin le plus court, ce qui est impossible à savoir. Ceci a donc une incidence sur la vraie distance ou durée parcourue pour chaque déplacement observé. La problématique de la capacité des stations joue également un impact. Une station pleine est problématique pour un usager, ce qui a nécessairement un impact sur la distance ou durée qui aurait été parcourue dans le cas contraire.

## 4.5 Conclusion du chapitre

Ce chapitre a présenté un portrait des déplacements des usagers de vélopartage. Plus précisément, il a été question des lignes de désir des membres BIXI lors de la pointe matinale des jours ouvrables. Ceci a révélé une grande concentration de la Destination des déplacements, vers le réseau de transport collectif, ou plus particulièrement des stations de métro, et vers des lieux d'activités importantes tel le centre-ville. La matrice O-D, et sa représentation spatiale, ont mis en lumière certaines dynamiques qui existent en vélopartage pendant la période étudiée.

Ensuite, le modèle gravitaire a été expliqué, qui permet de reproduire la distribution des déplacements grâce à une fonction de coût. Les divers paramètres qui le composent ont été détaillés.

L'analyse des résidus s'est révélée plutôt complexe. Outre ce qui a été présenté dans le document, de nombreuses explorations et transformations de données ont été effectuées. Non seulement est-ce que le but est d'expliquer les résidus, afin d'en déterminer la cause, mais la nature d'un modèle de distribution double le processus : certaines valeurs prédictives sont sous-estimées et d'autres sont surestimées. Il pourrait être pertinent d'agrégner les stations, par secteur de recensement par exemple, ce qui pourrait améliorer les résultats et rendre le traitement des résultats plus facile.

Effectivement, la revue de littérature sur la distribution des déplacements en vélopartage est mince et les études sur le sujet selon d'autres modes se basent sur une quinzaine de zones, alors qu'il est question de plus de 400 stations dans le cas présent.

Outre ceci, l'analyse des résidus démontre que la technique utilisée permet tout de même de bien reproduire les déplacements observés, mais que certaines paires O-D possèdent des relations particulières que le modèle simple ne peut reproduire efficacement. Une piste d'analyse spatiale a donc été présentée, ce qui permet d'énumérer quelques perspectives de recherche particulière à considérer.

Effectivement, il est possible d'intégrer au modèle gravitaire le facteur  $K$ , qui permet d'ajuster la modélisation selon des dynamiques particulières entre certaines paires O-D. Alors que la littérature suggère de fixer ce paramètre le moins possible, il pourrait être pertinent de l'intégrer en tant que coefficient qui traduit des éléments spatiaux ou même des caractéristiques du réseau BIXI. Il serait question d'insérer l'usage du sol, la capacité ou encore la proximité au centre-ville des stations et ce, à l'Origine et à la Destination.

Une autre piste de recherche est d'utiliser la fonction de coût combinée qui permet de mieux caler la distribution des déplacements selon la distance parcourue. La majorité des efforts ont été mis sur la calibration d'une fonction de coût à un seul paramètre et il pourrait être certainement pertinent de tenter de calibrer ce type de fonction qui en comprend deux. Des tests préliminaires ont été effectués sans succès, plus de temps est donc nécessaire.

Finalement, il pourrait être intéressant de tester un autre type de modèle de distribution, soit celui d'opportunité. Ce dernier se base sur un classement des stations effectué a priori. Les opinions dans la littérature sur le sujet sont partagées, tout comme sur le modèle gravitaire, mais il pourrait être intéressant de le tester dans le contexte de déplacements d'une station vers une autre.

## CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS

Ce chapitre vient conclure le mémoire, qui désirait approfondir les connaissances sur vélopartage et particulièrement la grande variabilité spatio-temporelle qui existe dans son utilisation. Il est donc question de conclure en présentant un retour sur la recherche. Une synthèse de celle-ci permet de rappeler des éléments révélés lors des analyses de la demande de déplacements aux stations ainsi que la distribution spatiale de ceux-ci. De plus, les contributions des différentes étapes de modélisations sont rappelées. En guise de fin, il est question des limites de la recherche qui n'ont pas été évoquées auparavant avant de terminer avec la présentation des perspectives de recherche sur le vélopartage.

### 5.1 Synthèse de la recherche et contributions

D'abord, une mise en contexte, qui se sépare en deux grandes parties, a permis de faire état de plusieurs aspects du vélopartage. La première partie se penche sur le vélopartage en tant que mode de transport. Elle place ce mode de déplacement dans un contexte urbain par rapport à d'autres modes de transport. Un survol de l'évolution des systèmes de VLS a été effectué ensuite afin de comprendre les différents enjeux intrinsèques à ce mode ainsi que les pistes de développement technique dans l'avenir. Puis, le fonctionnement ainsi que les composantes clés ont été décrites pour le réseau montréalais BIXI, qui constitue également la source des données pour les analyses et modélisations subséquentes. La deuxième partie de la contextualisation s'est penchée sur la planification des transports puisque le mémoire a une portée qui tend vers une planification stratégique. Il y est donc question du vélopartage en tant que sujet d'étude permettant de traduire la mobilité des usagers et la relation qui existe entre celle-ci et des tendances d'utilisation. En ce qui concerne la demande de déplacements aux stations, la revue de littérature scientifique a démontré une propension qui existe entre l'usage de différents systèmes de vélopartage, le temps et l'espace. Cette variabilité temporelle doit être réduite afin de mieux cerner une utilisation plus régulière alors que plusieurs variables permettant de décrire le territoire se sont révélées comme ayant un impact significatif. Les études concernant la distribution spatiale des déplacements se sont avérées bien moins nombreuses, mais révèlent néanmoins l'importance de la distance entre l'Origine et la Destination et le moment dans la journée. La présentation de la PSC et de ses quatre étapes a permis de présenter brièvement la méthodologie utilisée dans la modélisation. Les deux

premières étapes sont utilisées dans ce mémoire dans une perspective stratégique, soit la génération et la distribution des déplacements.

Afin de modéliser la génération des déplacements, une analyse temporelle a permis de constater une grande variabilité dans l'usage selon le moment dans l'année et le jour dans la semaine. De plus, un simple taux moyen des flux sortants à chaque station a révélé une grande disparité d'utilisation selon l'emplacement de celles-ci. Ainsi, il est question de modéliser ces déplacements grâce à une régression linéaire multiple, qui a été présentée dans le détail. Celle-ci se base sur plusieurs variables indépendantes et permet de voir l'effet qu'elles ont sur une variable dépendante, dans ce cas, les flux sortants et entrants moyens quotidiens aux stations. Les variables indépendantes, au nombre de 15, traduisent l'espace dans lequel s'insèrent les stations grâce à des zones tampons de 400 m, et sont regroupées en quatre catégories : sociodémographique; environnement bâti et usage du sol; offre de transport; offre du réseau BIXI. Ces dernières ont été retenues en fonction de la revue de littérature. La technique a donc permis de révéler que des éléments comparables aux diverses études sur le sujet ont un impact sur les flux, notamment la densité de population, l'âge, les emplois, le passage des autobus aux arrêts, la présence d'infrastructures cyclables et la capacité des stations BIXI. En guise de validation, la méthode a été reproduite sur des déplacements d'une autre année et démontre que l'effet de certaines variables est instable, ce qui signifie que la recherche doit être approfondie en ce sens.

La modélisation de la distribution des déplacements s'est attardée quant à elle sur les membres du réseau BIXI et leur utilisation en pointe matinale lors de jours ouvrables. Une analyse de la distance et de la durée des déplacements a été présentée selon différentes métriques. L'analyse des données a également permis de constater une grande concentration de la Destination des déplacements pendant la période observée. De plus, la spatialisation de la matrice O-D a été utilisée pour mettre en lumière la probable combinaison du vélopartage avec l'utilisation du métro, tout comme la spatialisation des paires O-D les plus fréquentées. Ces lignes de désir ont donc été modélisées grâce au modèle gravitaire, qui attribue une valeur attractive à toutes les paires O-D en fonction de la distance qui les sépare. Une calibration du modèle permet d'obtenir un paramètre qui ajuste la fonction de coût à la distribution des déplacements selon la distance ou la durée. Les résultats de cette modélisation sont prometteurs, malgré quelques erreurs importantes. Celles-ci sont surtout dues aux stations qui produisent et attirent de nombreux déplacements, que le modèle ne peut pas reproduire efficacement. Une brève analyse spatiale a soulevé l'impact de certains éléments sur les

résidus ce qui constitue une piste à approfondir afin d'améliorer les modèles. Finalement, l'utilisation d'indicateurs pour valider les paramètres calibrés sur une autre année a démontré une constance dans les résultats. De plus, pour la totalité des modèles, au moins 93 % des résidus absolu sont inférieurs à 10, ce qui est prometteur, considérant qu'il s'agit d'une somme annuelle qui est modélisée.

## 5.2 Limites

La modélisation est un procédé qui tente de refléter un phénomène observé. Pour ce faire, elle utilise des données, indicateurs et variables qui tentent de traduire la réalité. Or, cette représentation est nécessairement limitée puisqu'elle ne peut prendre en compte les relations complexes qui dépassent le cadre de la recherche. Outre ceci, des limites plus définies viennent baliser le cadre de cette recherche.

D'abord, la méthodologie utilisée dans cette recherche emprunte les deux premières étapes de la PSC, qui elle-même possède des limites en tant qu'outil de planification des transports (section 2.5.1.5). De plus, chaque étape comprend également ses propres limites, présentées en détail dans chaque partie de modélisation, celle sur la génération à la section 3.2.5 et celle sur la distribution à la section 4.4.6.

Au-delà de ces limites déjà abordées, il est important de mentionner que l'analyse et la modélisation effectuée dans ce mémoire se basent sur des données et ne sont donc pas exemptes d'erreurs. Par exemple, les informations sur les déplacements effectués par les usagers BIXI ne prennent pas en compte ceux qui ne sont pas effectués si une station est vide, ou encore modifiés, lorsqu'une station est pleine. Cet élément est plus complexe à considérer et une meilleure redistribution des vélos en temps réel permet de minimiser ces événements, ce qui relève de la planification opérationnelle.

Ceci permet également de ramener le concept de temporalité. Les analyses et modèles présentés dans ce mémoire sont effectués dans une optique stratégique, ce qui explique pourquoi il est question d'une base annuelle. Or, l'utilisation des systèmes de VLS varie nécessairement selon le moment dans la journée, le jour dans la semaine ainsi que la période dans l'année, ce qui a été expliqué et démontré maintes fois dans ce document.

Finalement, il faut rappeler que dans le contexte de ce document, un déplacement signifie la location d'un vélo à la station d'Origine jusqu'à son retour à la station de Destination. Or, en réalité,

ceci peut constituer une partie d'un déplacement, ou de plusieurs déplacements qui pourrait former une chaîne. Ce faisant, de nombreux éléments peuvent avoir des impacts à court, moyen et long terme sur l'utilisation d'un réseau de vélopartage. Il est nécessaire de garder une vision globale du territoire qui peut avoir une influence significative sur la performance et l'utilisation de BIXI.

### **5.3 Perspectives**

De nombreux défis restent à venir en ce qui concerne l'étude et la planification des systèmes de vélopartage. Effectivement, ce mode de transport est aujourd'hui planétaire et pertinent dans l'optique où la demande de déplacements est de plus en plus multimodale.

D'abord, les modèles de génération présentés sont pertinents, mais certains éléments ne sont pas stables selon les années modélisées. Il semble donc judicieux de suivre l'évolution de ces impacts sur plusieurs années de références. Ceci permettrait d'assurer une constance dans l'utilisation de variables indépendantes qui expliquent la demande aux stations, soit sur le plan de la production ou de l'attraction de déplacements.

En ce qui concerne les modèles de distribution, l'intégration de certaines variables pourrait nécessairement améliorer la qualité prédictive du modèle gravitaire. Cette piste devrait être approfondie. De plus, alors que les déplacements effectués par les membres ont été utilisés, il serait également pertinent de tester avec l'ensemble des usagers et de comparer les résultats, puisque l'utilisation du réseau par les usagers occasionnels a nécessairement une influence sur les déplacements réalisés par tous, notamment en raison de la capacité des stations. D'autres temporalités d'utilisation pourraient également être observées, que ce soit la pointe du soir, la fin de semaine, mensuelle, etc. Une dernière piste d'amélioration à souligner est celle concernant la forme de la fonction de coût qui détermine l'attractivité d'une paire O-D donnée. Effectivement, la littérature tend vers une forme combinée qui allie puissance et exponentielle puisque celle-ci se conforme davantage à la distribution des déplacements selon la distance de ceux-ci. Or, puisqu'aucun paramètre ne provient d'autres études empiriques, une calibration est nécessaire. Ce processus requiert un temps de calcul important dû aux nombreuses possibilités de paires O-D. L'ajout d'un second paramètre à calibrer entraîne par le fait même un temps considérable à réaliser, mais semble très prometteur.

Finalement, il y aurait lieu de développer un modèle intégré qui pourrait prendre en compte ces diverses améliorations, soit un outil permettant de modéliser la génération et la distribution des déplacements. Cet outil pourrait être testé avec d'autres types de systèmes partagés comme l'autopartage et il permettrait également d'évaluer des scénarios de stations. Effectivement, la force du système BIXI est d'être modulable. Une planification grâce à un modèle intégré permettrait d'optimiser le réseau pour les usagers tout comme pour l'opérateur.

## BIBLIOGRAPHIE

- Active healthy communities. (2017). PedShed analysis. Tiré de <http://www.activehealthycommunities.com.au/plan/gis-analysis/walking-cycling-pedshed-analysis/>
- Agence QMI. (2017, 25 octobre). 4,4 millions de déplacements: Saison record pour BIXI. *Le journal de Montréal*. Tiré de <http://www.journaldemontreal.com/2017/10/25/44-millions-de-deplacements-saison-record-pour-bixi>
- Bérubé, N. (2017, 12 mai). Les BIXI, plus nombreux que jamais. *La Presse*. Tiré de <http://www.lapresse.ca/actualites/montreal/201705/12/01-5097243-les-bixi-plus-nombreux-que-jamais.php>
- BIXI Montréal. (2018a). Comment utiliser BIXI ? Tiré de <https://www.bixi.com/fr/comment-utiliser-bixi>
- BIXI Montréal. (2018b). *Point d'ancrage* [Photographie]. Tiré de <https://www.bixi.com/fr/comment-utiliser-bixi>
- BIXI Montréal. (2018c). *Station BIXI* [Photographie]. Tiré de <https://www.bixi.com/fr/tarifs>
- BIXI Montréal. (2018d). Tarifs. Tiré de <https://www.bixi.com/fr/tarifs>
- Bonnel, P. (2002). *Prévision de la demande de transport*. (Thèse de doctorat, Université Lumière - Lyon II, Lyon, France). Tiré de <https://tel.archives-ouvertes.fr/tel-00268919>
- Borgnat, P., Robardet, C., Rouquier, J.-B., Abry, P., Flandrin, P., & Fleury, E. (2010). Shared bicycles in a city: A signal processing and data analysis perspective. *Advances in Complex Systems*, 14(03), 415-438. doi:10.1142/S0219525911002950
- Buck, D., & Buehler, R. (2012). *Bike lanes and other determinants of Capital Bikeshare trips*. Communication présentée à 91st Annual Meeting of the Transportation Research Board, Washington, D.C.
- Communauté métropolitaine de Montréal. (2018). Données géoréférencées. Tiré de <http://cmm.qc.ca/donnees-et-territoire/observatoire-grand-montreal/produits-cartographiques/donnees-georeferencees/>
- Daddio, D. W. (2012). *Maximizing bicycle sharing: An empirical analysis of Capital Bikeshare usage*. (Mémoire de maîtrise, University of North Carolina at Chapel Hill). Tiré de <https://cdr.lib.unc.edu/record/uuid:5fded6fa-eed9-46b0-a121-2c5578dcb899>
- DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of Public Transportation*, 12(4), 41-56. doi:<https://doi.org/10.5038/2375-0901.12.4.3>
- Faghih-Imani, A., & Eluru, N. (2015). *Analysing destination choice preferences in bicycle sharing systems: An investigation of Chicago's Divvy system*. Communication présentée à 94th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Faghih-Imani, A., Eluru, N., El-Geneidy, A. M., Rabbat, M., & Haq, U. (2014). *How does land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal*. Communication présentée à 93rd Annual Meeting of the Transportation Research Board, Washington, D.C.

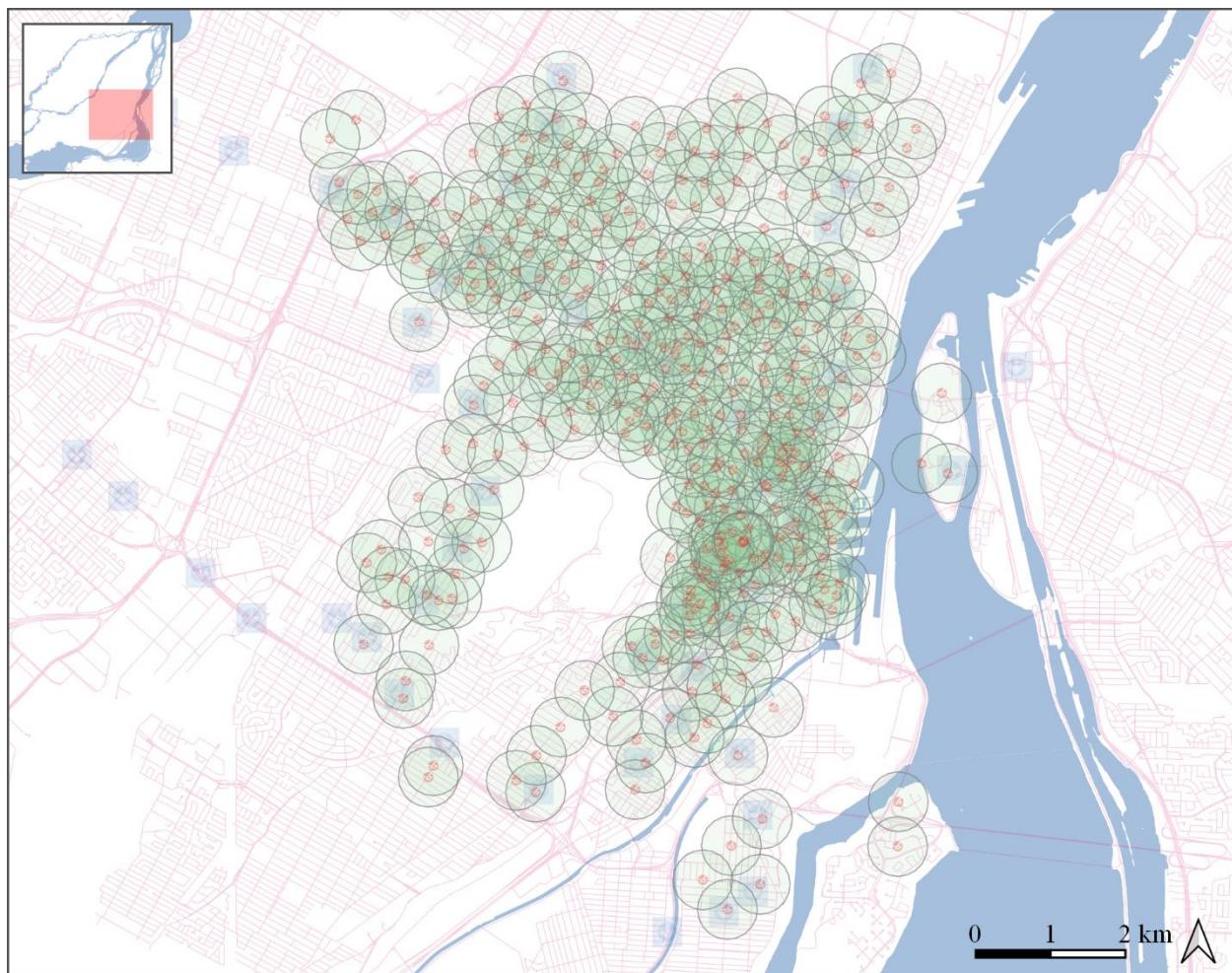
- Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews*, 36(1), 92-113. doi:<https://doi.org/10.1080/01441647.2015.1033036>
- Fishman, E., Washington, S., & Haworth, N. L. (2013). Bike share: a synthesis of the literature. *Transport reviews*, 33(2), 148-165. doi:10.1080/01441647.2013.775612
- Ford, C. (2017). The Wilcoxon Rank Sum Test. *University of Virginia Library*. Tiré de <http://data.library.virginia.edu/the-wilcoxon-rank-sum-test/>
- Froehlich, J., Neumann, J., & Oliver, N. (2009). *Sensing and predicting the pulse of the city through shared bicycling*. Communication présentée à 21st International Joint Conference on Artificial Intelligence, Pasadena, CA.
- Fuller, D., Gauvin, L., Kestens, Y., Daniel, M., Fournier, M., Morency, P., & Drouin, L. (2011). Use of a new public bicycle share program in Montreal, Canada. *American Journal of Preventive Medicine*, 41(1), 80-83. doi:10.1016/j.amepre.2011.03.002
- González, F., Melo-Riquelme, C., & de Grange, L. (2016). A combined destination and route choice model for a bicycle sharing system. *Transportation*, 43(3), 407-423. doi:<https://doi.org/10.1007/s11116-015-9581-6>
- Hampshire, R. C., & Marla, L. (2012). *An analysis of bike sharing usage: Explaining trip generation and attraction from observed demand*. Communication présentée à 91st Annual Meeting of the Transportation Research Board, Washington, D.C.
- Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100-108. doi:10.2307/2346830
- Hyman, G. M. (1969). The calibration of trip distribution models. *Environment and Planning A*, 1(1), 105-112. doi:10.1068/a010105
- Luxen, D., & Vetter, C. (2011). *Real-time routing with OpenStreetMap data*. Communication présentée à Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Chicago, Illinois. doi:<https://doi.org/10.1145/2093973.2094062>
- Mahmoud, M. S., El-Assi, W., & Habib, K. N. (2015). *Effects of built environment and weather on bike sharing demand: Station level analysis of commercial bike sharing in Toronto*. Communication présentée à 94th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Mateo-Babiano, I., Bean, R., Corcoran, J., & Pojani, D. (2016). How does our natural and built environment affect the use of bicycle sharing? *Transportation Research Part A: Policy and Practice*, 94(Supplement C), 295-307. doi:10.1016/j.tra.2016.09.015
- McNally, M. G. (2007). The four-step model. Dans *Handbook of Transport Modelling* (p. 35-53).
- Meddin, R., & DeMaio, P. (2018). The bike-sharing world map. Tiré de [https://www.google.com/maps/d/viewer?ll=15.114552725691297%2C42.890623000000005&spn=143.80149%2C154.6875&hl=en&msa=0&z=1&source=embed&ie=UTF8&om=1&mid=1UxYw9YrwT\\_R3SGsktJU3D-2GpMU](https://www.google.com/maps/d/viewer?ll=15.114552725691297%2C42.890623000000005&spn=143.80149%2C154.6875&hl=en&msa=0&z=1&source=embed&ie=UTF8&om=1&mid=1UxYw9YrwT_R3SGsktJU3D-2GpMU)
- Midgley, P. (2009). The role of smart bike-sharing systems in urban mobility. *Journeys*, 2, 23-31.

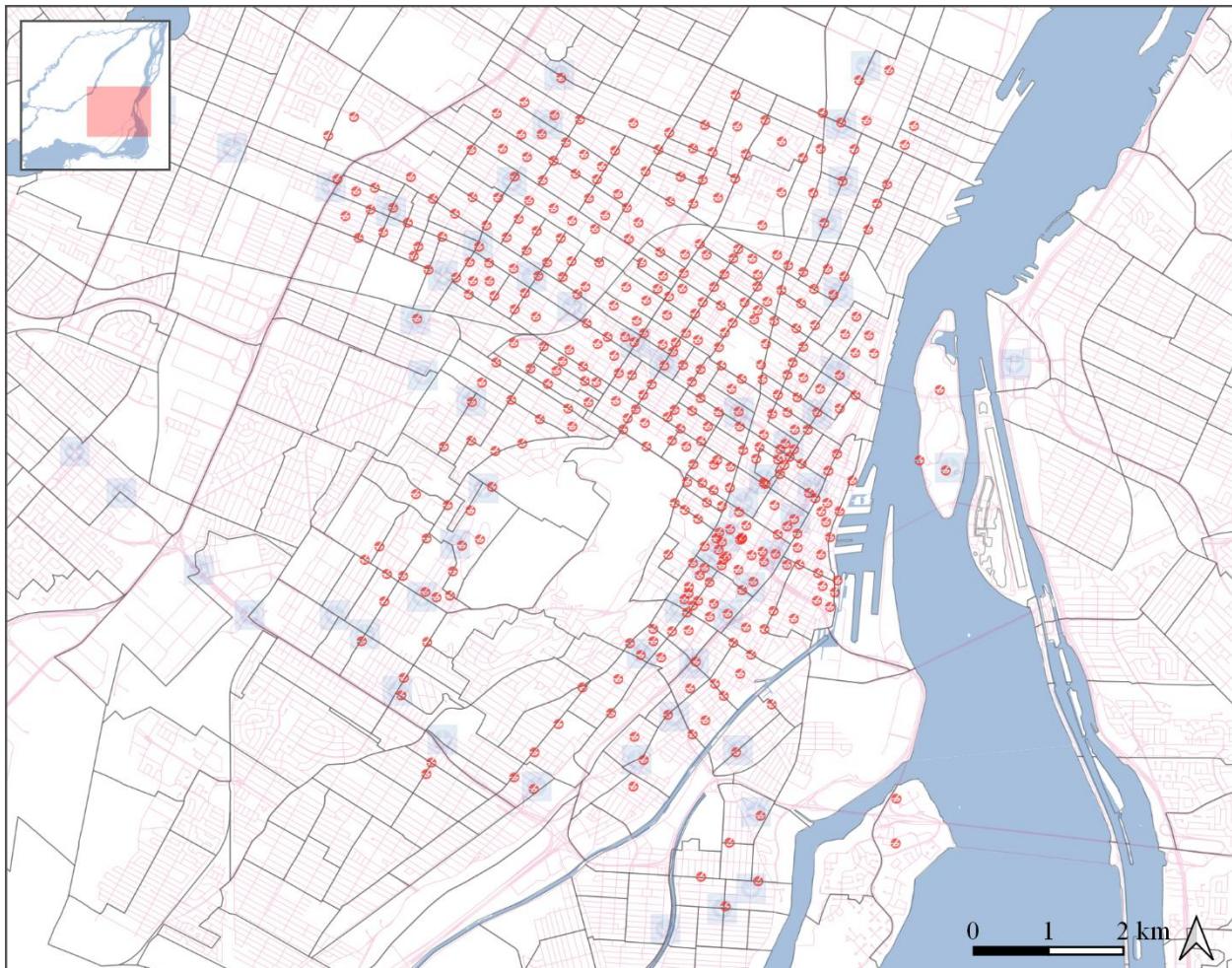
- Mladenović, M. N., & Trifunović, A. (2014). The shortcomings of the conventional four step travel demand forecasting process. *Journal of Road and Traffic Engineering*, 60(1), 5-12. Tiré de [http://www.academia.edu/22281753/The\\_Shortcomings\\_of\\_the\\_Conventional\\_Four\\_Step\\_Travel\\_Demand\\_Forecasting\\_Process](http://www.academia.edu/22281753/The_Shortcomings_of_the_Conventional_Four_Step_Travel_Demand_Forecasting_Process)
- Morency, C., Trépanier, M., Frappier, A., & Bourdeau, J.-S. (2017). *Longitudinal analysis of bikesharing usage in Montreal, Canada*. Communication présentée à 96th annual meeting of the Transportation Research Board, Washington, D.C.
- Morency, C., Trépanier, M., & Godefroy, F. (2011). *Insights into Montreal's bikesharing system*. Communication présentée à 90th annual meeting of the Transportation Research Board, Washington, D.C.
- Morency, C., Trépanier, M., Paez, A., Verreault, H., & Faucher, J. (2016). *Modelling bikesharing usage in Montreal over 6 years*. Communication présentée à 95th annual meeting of the Transportation Research Board, Washington, D.C.
- Nair, R., Miller-Hooks, E., Hampshire, R. C., & Bušić, A. (2013). Large-scale vehicle sharing systems: Analysis of Vélib'. *International Journal of Sustainable Transportation*, 7(1), 85-106. doi:10.1080/15568318.2012.660115
- Nazem, M., Trépanier, M., & Morency, C. (2013). Integrated Intervening Opportunities Model for Public Transit Trip Generation-Distribution. *Transportation Research Record: Journal of the Transportation Research Board*, 2350, 47-57. doi:<https://doi.org/10.3141/2350-06>
- Normandin, P.-A. (2017, 22 avril). Montréal appréhende l'arrivée d'un BIXI chinois « parasite ». *La Presse*. Tiré de [http://plus.lapresse.ca/screens/29256fa1-f841-475d-a30a-5d0d624f5f86%7C\\_0.html](http://plus.lapresse.ca/screens/29256fa1-f841-475d-a30a-5d0d624f5f86%7C_0.html)
- Ortuzar, J. d. D., & Willumsen, L. G. (2011). *Modelling transport* (4<sup>e</sup> éd.). Royaume-Uni: John Wiley & Sons Ltd.
- Quay Communications Inc. (2008). *TransLink public bike system feasibility study*. Tiré de <http://docplayer.net/59852320-Translink-public-bike-system-feasibility-study.html>
- Raux, C., & Zoubir, A. (2016). *Who are bike sharing schemes members and how they travel daily? The case of the Lyon's "Velo 'v" scheme*. Communication présentée à 95th annual meeting of the Transportation Research Board, Washington, D.C. Tiré de <https://halshs.archives-ouvertes.fr/halshs-01193169>
- Ressources naturelles Canada. (2014). API Élévation. Tiré de <http://www.rncan.gc.ca/sciences-terre/geographie/information-topographique/api/17329>
- Ricci, V. (2005). Fitting distributions with R. Tiré de <https://cran.r-project.org/doc/contrib/Ricci-distributions-en.pdf>
- Rixey, R. A. (2013). Station-level forecasting of bikesharing ridership: Station network effects in three U.S. systems. *Transportation Research Record: Journal of the Transportation Research Board*, 2387, 46-55. doi:10.3141/2387-06
- Scarf, P., & Grehan, P. (2005). An empirical basis for route choice in cycling. *Journal of sports sciences*, 23(9), 919-925. doi:10.1080/02640410400023282

- Shaheen, S. A., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, present, and future. *Transportation Research Record: Journal of the Transportation Research Board*, 2143, 159-167. doi:<https://doi.org/10.3141/2143-20>
- Sherali, H. D., Narayanan, A., & Sivanandan, R. (2003). Estimation of origin–destination trip-tables based on a partial set of traffic link volumes. *Transportation Research Part B: Methodological*, 37(9), 815-836. doi:10.1016/S0191-2615(02)00073-5
- Smith, D., & Hutchinson, B. (1981). Goodness of fit statistics for trip distribution models. *Transportation Research Part A: General*, 15(4), 295-303. doi:[https://doi.org/10.1016/0191-2607\(81\)90011-X](https://doi.org/10.1016/0191-2607(81)90011-X)
- Statistique Canada. (2015). Secteur de recensement. Tiré de <http://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo013-fra.cfm>
- Thomas, R. W. (1977). An interpretation of the journey-to-work on Merseyside using entropy-maximising methods. *Environment and Planning A: Economy and Space*, 9(7), 817-834. doi:<https://doi.org/10.1068/a090817>
- Tran, T. D. (2016). *L'usage du vélo en libre-service: Impact de l'environnement socio-économique des stations sur la génération de la demande*. (Thèse de Doctorat, Université Lumière - Lyon II, Lyon, France).
- Tran, T. D., Ovtracht, N., & d'Arcier, B. F. (2015). Modeling bike sharing system using built environment factors. *Procedia CIRP*, 30, 293-298. doi:10.1016/j.procir.2015.02.156
- Vogel, M., Hamon, R., Lozenguez, G., Merchez, L., Abry, P., Barnier, J., . . . Robardet, C. (2014). From bicycle sharing system movements to users: A typology of Vélo'v cyclists in Lyon based on large-scale behavioural dataset. *Journal of Transport Geography*, 41, 280-291. doi:10.1016/j.jtrangeo.2014.07.005
- Vogel, P., Greiser, T., & Mattfeld, D. C. (2011). Understanding bike-sharing systems using data mining: Exploring activity patterns. *Procedia - Social and Behavioral Sciences*, 20, 514-523. doi:<https://doi.org/10.1016/j.sbspro.2011.08.058>
- Wang, X., Lindsey, G., Schoner, J. E., & Harrison, A. (2016). Modeling bike share station activity: Effects of nearby businesses and jobs on trips to and from stations. *Journal of Urban Planning and Development*, 142(1). doi:[https://doi.org/10.1061/\(ASCE\)UP.1943-5444.0000273](https://doi.org/10.1061/(ASCE)UP.1943-5444.0000273)
- Zhang, S., He, K., Dong, S., & Zhou, J. (2016). Modeling the distribution characteristics of urban public bicycle rental duration. *Discrete Dynamics in Nature and Society*, 2016, 9. doi:<https://doi.org/10.1155/2016/8947525>
- Zhang, S., Ji, Y., Sheng, D., & Zhou, J. (2018). *Origin-destination distribution prediction model for public bicycles based on rental characteristics*. Communication présentée à Green Intelligent Transportation Systems: Proceedings of the 7th International Conference on Green Intelligent Transportation System and Safety, Singapore (p. 293-303). doi:[https://doi.org/10.1007/978-981-10-3551-7\\_22](https://doi.org/10.1007/978-981-10-3551-7_22)
- Zhao, J., Deng, W., & Song, Y. (2014). Ridership and effectiveness of bikesharing: The effects of urban features and system characteristics on daily use and turnover rate of public bikes in China. *Transport Policy*, 35, 253-264. doi:10.1016/j.tranpol.2014.06.008

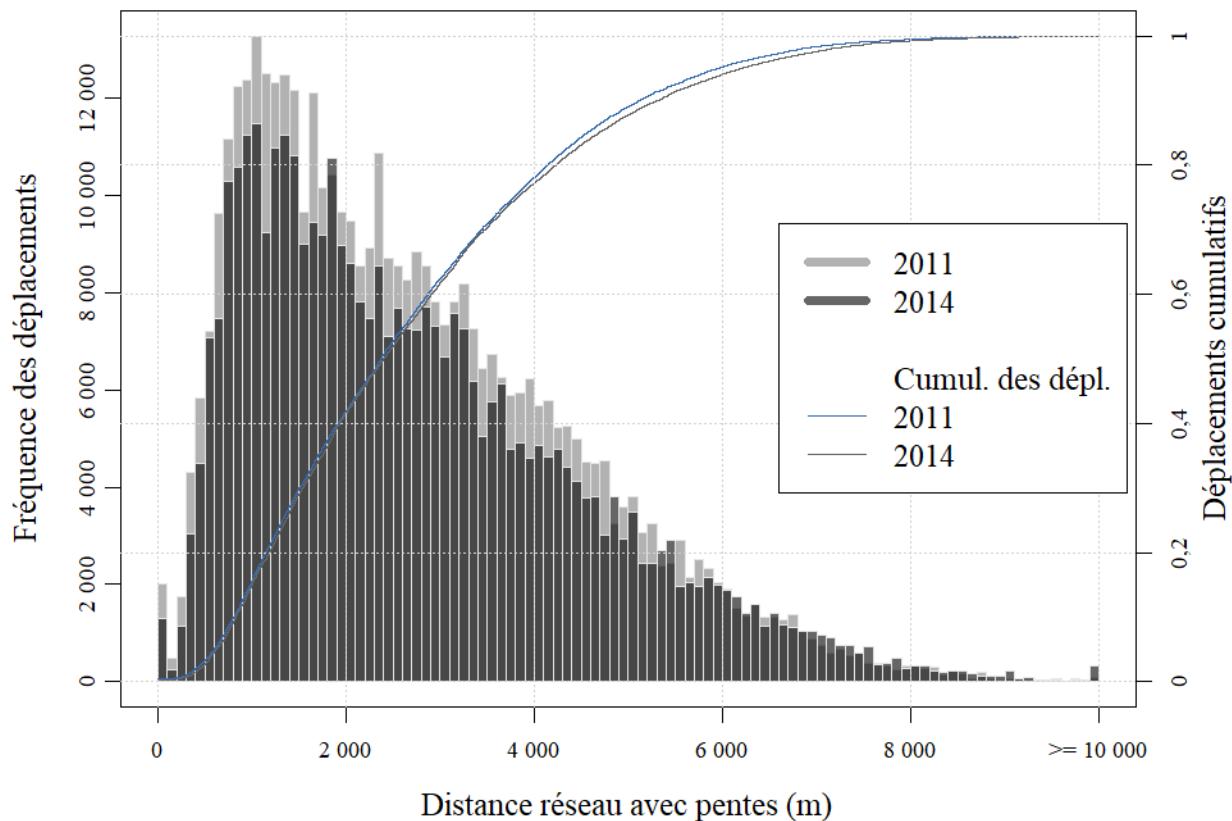
## ANNEXES

### ANNEXE A – ZONES TAMPONS DE 400 M AUTOUR DES STATIONS BIXI EN 2011

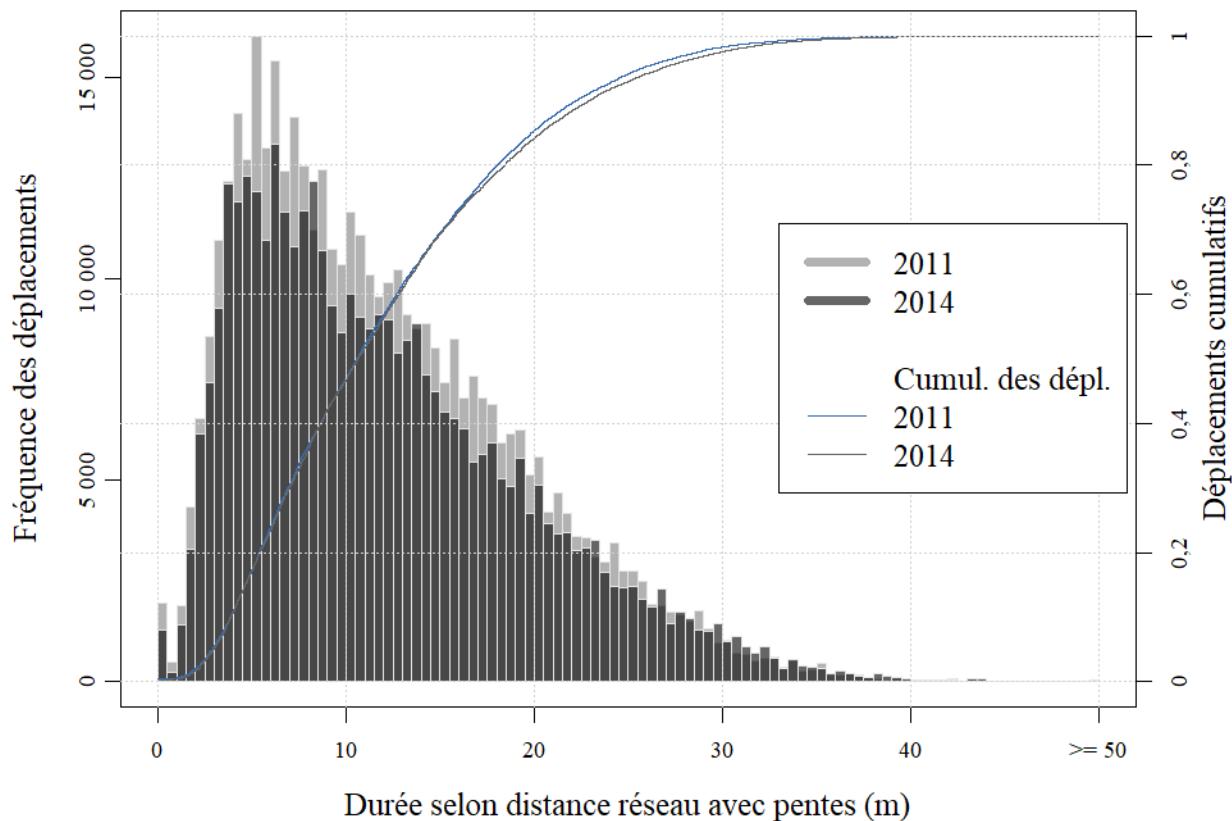


**ANNEXE B – PRÉSENTATION DES SR DU RECENSEMENT DE 2011**

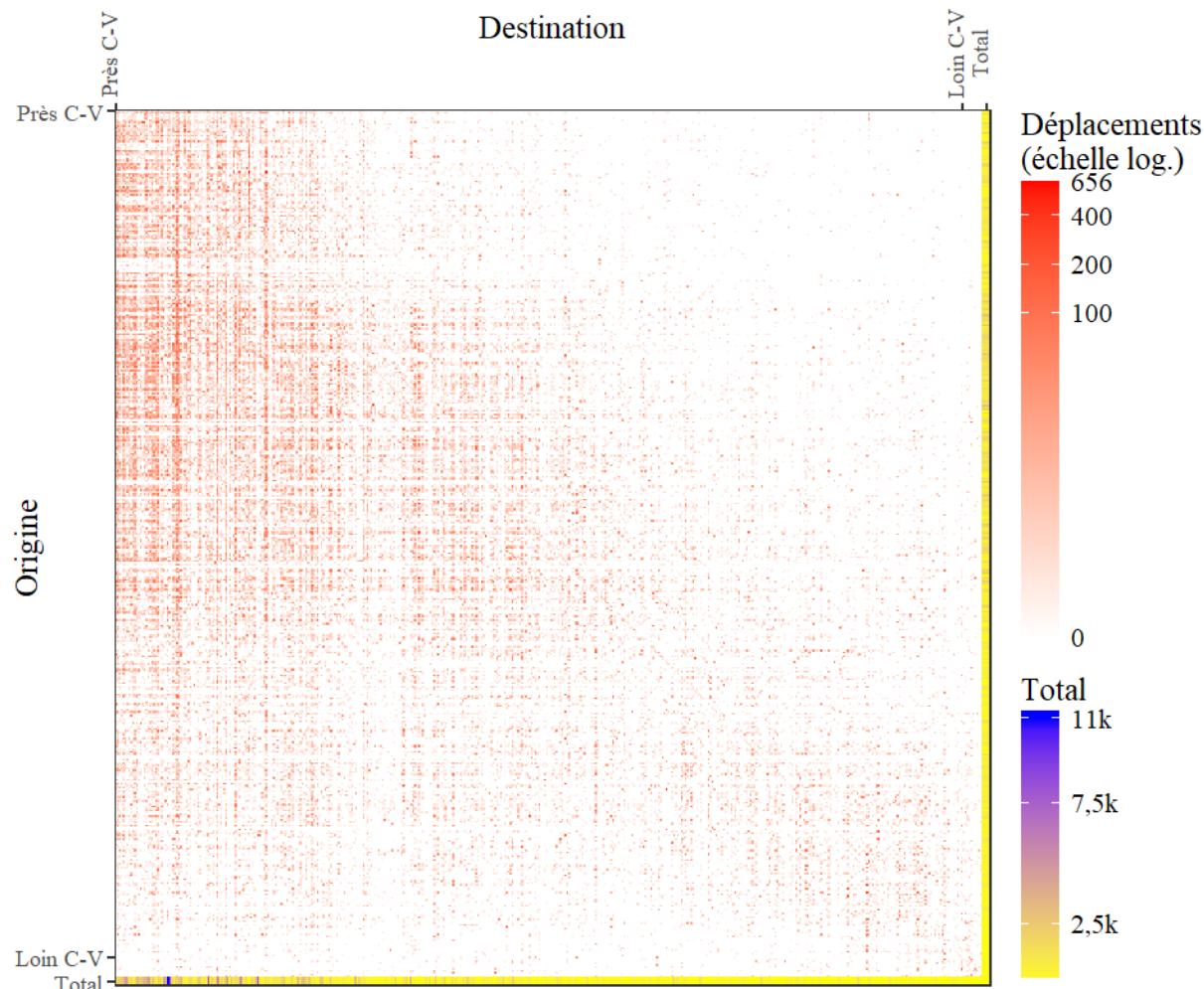
**ANNEXE C – DISTRIBUTION ET CUMULATIF DES DÉPLACEMENTS  
SELON LA DISTANCE RÉSEAU AVEC DÉNIVELÉ**



## ANNEXE D – DISTRIBUTION ET CUMULATIF DE LA DURÉE DES DÉPLACEMENTS SELON LA DISTANCE RÉSEAU AVEC DÉNIVELÉ



**ANNEXE E – MATRICE ORIGINE-DESTINATION DES  
DÉPLACEMENTS DES MEMBRES LORS DE LA POINTE MATINALE  
DES JOURS OUVRABLES EN 2014**



**ANNEXE F – STATISTIQUES SOMMAIRES DES RÉSIDUS ABSOLUS  
(EXCLUANT 0) SELON LA DISTANCE AU CENTRE-VILLE**

|          |       | 2011  |       |         | 2014       |       |         |
|----------|-------|---|-------|---------|------------|-------|---------|
|          |       | Dist. O → D<br><i>P = petite</i><br><i>M = moyenne</i><br><i>G = grande</i> | Somme | Moyenne | Écart-type | Somme | Moyenne |
| <b>1</b> | P → P | 124 027,00  | 8,38  | 17,66   | 138 652,00 | 7,86  | 16,33   |
| <b>2</b> | P → M | 15 745,00   | 2,94  | 6,13    | 17 857,00  | 3,07  | 6,75    |
| <b>3</b> | P → G | 2 679,00  | 2,39  | 5,43    | 2 643,00   | 2,10  | 4,49    |
| <b>4</b> | M → P | 117 921,00  | 7,76  | 12,89   | 105 656,00 | 6,19  | 11,29   |
| <b>5</b> | M → M | 57 271,00   | 5,35  | 14,45   | 53 236,00  | 4,84  | 12,05   |
| <b>6</b> | M → G | 16 577,00   | 2,66  | 5,62    | 15 371,00  | 2,93  | 7,40    |
| <b>7</b> | G → P | 49 380,00   | 3,98  | 7,88    | 40 666,00  | 3,49  | 7,48    |
| <b>8</b> | G → M | 33 615,00   | 3,51  | 7,32    | 26 498,00  | 3,27  | 8,73    |
| <b>9</b> | G → G | 36 849,00   | 3,91  | 9,05    | 30 153,00  | 4,14  | 10,55   |