UNIVERSITÉ DE MONTRÉAL

ANALYSE EXPLORATOIRE DE TRANSACTIONS DE CARTES À PUCES SUR LE RÉSEAU DU MÉTRO DE MONTRÉAL

WILLIAM DOUCET-KOUSSAYA

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)

MAI 2012

© William Doucet-Koussaya, 2012.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ANALYSE EXPLORATOIRE DE TRANSACTIONS DE CARTES À PUCES SUR LE RÉSEAU DU MÉTRO DE MONTRÉAL

Présenté par : DOUCET-KOUSSAYA, William

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

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

M. SAUNIER Nicolas, Ph.D., président

M. CHAPLEAU Robert, Ph.D., membre et directeur de recherche

Mme GIRARD Diane, M.Sc.A., membre

REMERCIEMENTS

L'auteur remercie d'abord la STM d'avoir fourni l'ensemble de données de transactions de validations par cartes à puce, sans lequel la réalisation de ce projet n'aurait pas été possible.

L'auteur remercie le professeur Robert Chapleau pour sa patience, et pour son approche qui lui a permis le développement d'une profonde conscience de soi.

L'auteur tient aussi à remercier ses collègues Félix, Sébastien, Louiselle, Jason et Ablo pour les discussions passionnées et créatives.

Enfin, l'auteur remercie sa famille qui a toujours cru en lui, et en particulier son frère, qui a courageusement enduré de nombreuses remises en question.

RÉSUMÉ

L'implantation des systèmes d'acquisition de données automatisés en transport en commun permet aujourd'hui d'acquérir des données d'achalandage de manière continue. En sondant le réseau de façon passive, c'est-à-dire sans interaction avec l'usager, ces systèmes collectent des données qui offrent une richesse spatiotemporelle que les enquêtes traditionnelles ne permettent pas d'obtenir.

Par ailleurs, les données de service planifié sont maintenant générées dans un format standard dans la plupart des sociétés de transport, soit le format GTFS. Ce format impose la génération de différents fichiers qui contiennent, entre autres, la position des arrêts et stations, et les temps de passage à la minute près de tous les véhicules de transport en commun.

L'objectif de cette recherche est d'examiner le potentiel théorique de l'assemblage de ces différentes sources de données pour le réseau de métro de Montréal, en évaluant l'efficacité de logiciels ouverts, ou « open source ». Les données de transactions par cartes à puce en lien avec le métro présentent certaines caractéristiques qui affineront les analyses. D'une part, l'équipement de perception est fixe en station, ce qui élimine l'incertitude quant à l'emplacement des transactions effectuées dans le métro. Cette caractéristique simplifie les traitements qui doivent être appliqués aux données avant d'en faire l'exploitation. D'autre part, l'empreinte des données de transactions étant à la seconde, il devient possible d'analyser la dynamique par véhicule dans le réseau de métro.

Le projet débute en rappelant les concepts associés à un réseau de métro et à la tarification par carte à puce, et en cernant le rôle des technologies informationnelles en transport. Ensuite, une structure est établie pour associer les données aux objets auxquels elles font référence. Cette étape est fondamentale pour la suite des manipulations, qui consistent à traiter et enrichir les données. Les procédures de traitement et d'enrichissement reposent sur des hypothèses qui elles, dépendent des objets présents dans la base de données, et de leurs attributs.

Une fois traitées et enrichies, les données d'achalandage et de service planifié peuvent être intégrées dans un logiciel de simulation qui permet de visualiser la dynamique des déplacements pour une journée complète de transactions. De plus, la démarche entreprise permet de décrire l'achalandage temporel des stations et des lignes de métro à un niveau de résolution à la seconde, de suivre les déplacements quotidiens associés à un numéro de carte, et de dresser le profil d'achalandage d'un véhicule pour un départ particulier. Ces analyses sont conduites à l'aide de graphiques, de tableaux, et de représentations cartographiques par l'entremise d'un SIG.

Ce projet s'inscrit dans un contexte technologique en constante évolution, et ne doit donc pas être pris comme une entité unique et distincte, mais plutôt comme une pièce d'un ensemble beaucoup plus complexe. En effet, les concepts définis ici sont les fondements de l'analyse de données, mais les moyens techniques utilisés, eux, ne sont pas nécessairement ceux qui seront utilisés d'ici quelques années.

ABSTRACT

The recent implementation of automated data collection systems among transit networks allows continuous data retrieval that provides spatial and temporal load information. Passive surveying of the network using smart card data is done without any interaction with the transit user, and the data gathered by the systems offers a spatial and temporal richness that traditional surveying methods cannot match.

Along with the automated data collection systems is the GTFS. It is the new standard in terms of planned service format within transit authorities. The GTFS requires a certain amount of information that specifies stop spatial coordinates and stop times of each vehicle.

The main objective of this research is to examine the theoretical possibilities that arise from the combination of these different data sources using open source software. Smart card transactions associated with the metro show different characteristics that will facilitate data analysis. On one hand, the automatic fare collection system position is fixed in space. This way, the uncertainty concerning spatial location of transactions is eliminated. This aspect of the data will greatly simplify data processing and enrichment. On the other hand, the precise time stamp associated to the transactions data makes it possible to analyse the dynamics of loadings and alighting's of metro cars, the loads at a particular station or individual itineraries.

The first part of the project consists in a reminder of fundamental concepts tied to a metro network and to smart card attributes, and positions the role of information technologies in the field of public transit. Secondly, it establishes a work structure to define objects linked to the data collected. This step represents the backbone of the following sections concerning data processing and data enrichment. Data processing and data enrichment are based on different hypothesis that depend on the objects defined in the database as well as their attributes.

Once the transactions data is properly processed and enriched, it can be integrated along with planned service data into simulation software. It will allow the visualisation of the usage of the metro network for a whole day of service. In addition, the method deployed in this research allows analysis of the data at different levels of aggregation. Loads can be analysed for a metro line, station, vehicle, or run. These analyses are produced using graphics, tables, and cartographical representations created with GIS software.

This project subscribes in a continuously evolving technological context and must not be taken as a whole, but rather as a piece of a much more complex situation. In fact, the concepts defined in this research are the fundamentals of data analysis in the field of public transit, but the technical instruments used and the way they are used might not be the same in future research.

TABLE DES MATIÈRES

REMERCIEMENTS III

RÉSUMÉ IV

ABSTRACT VI

TABLE DES MATIÈRES VIII

LISTE DES TABLEAUX XII

LISTE DES FIGURES XIV

LISTE DES SIGLES ET ABRÉVIATIONS XVII

LISTE DES ANNEXES XVIII

CHAPITRE 1 INTRODUCTION 1

1.1 Contexte 2

1.1.1 Technologie 2

1.1.2 Population et Déplacements 2

1.1.3 Société de transport 3

1.2 Objectifs 3

1.3 Contenu 4

CHAPITRE 2 FONDEMENTS 6

2.1 Réseau de transport en commun 7

2.1.1 Approche orientée objet 7

2.1.2 Objets associés au métro 8

2.1.3 Planification du service et opération 9

2.1.4 Perception et tarification 11

2.2 La carte à puce en transport en commun 11

2.2.1 Fonctionnement du système 11

2.2.2 Avantages et inconvénients 12

2.2.3 Bénéfices d'utilisation 13

2.2.4 Problématiques de traitement des données 14

2.2.5 Dérivation des points de débarquement 15

2.3 Technologie de l'information 16

2.3.1 Système d'information géographique en transport 16

2.3.2 Google Transit et données GTFS 18

2.4 Modélisation 20

CHAPITRE 3 LE MÉTRO DE MONTRÉAL 21

3.1 Description du réseau 22

3.1.1 Lignes 22

3.1.2 Stations 22

3.1.3 Matériel roulant 22

3.1.4 Perception et tarification 23

3.1.5 Planification du service 24

3.1.6 Coûts d'exploitation 28

3.1.7 Interruptions de service 29

3.2 Expérimentations sur le métro 29

3.2.1 Les pannes de métro et ses clients victimes (Chapleau, 1997) 29

3.2.2 Flux dans le métro de Montréal : concepts et mesures (Piché & Chapleau, 2007) 30

3.2.3 Modélisation espace-temps du métro de Montréal (Blais, 2007) 30

CHAPITRE 4 GESTION DES DONNÉES 31

4.1 Caractérisation de l'ensemble de données de transactions 31

4.1.1 Acquisition des données 31

4.1.2 Objets de la base de données 32

4.1.3 Stockage 34

4.1.4 Cohérence des valeurs observées et cas particuliers 35

4.2 Procédures de traitement et d'enrichissement 37

4.2.1 Procédures de traitement 38

4.2.2 Procédures d'enrichissement 40

4.2.3 Données GTFS : précision à la seconde 45

4.3 Applications 53

4.3.1 Profils de charge : transactions cartes à puce et données GTFS enrichies 53

4.3.2 Usagers touchés par une panne de service 59

CHAPITRE 5 MODÉLISATION DU RÉSEAU DE MÉTRO DE MONTRÉAL 62

5.1 Présentation des logiciels 62

5.1.1 Quantum GIS 62

5.1.2 TRANSIMS 63

5.2 Sommaire de l'expérimentation 67

5.3 ConvertNet : codification du réseau de métro 68

5.3.1 TransimsNet 69

5.3.2 TransitNet 70

5.3.3 ArcNet 72

5.4 RouterRuns : affectation des déplacements 73

5.4.1 Router 74

5.4.2 PlanSum 75

5.4.3 ArcPlan 75

5.4.4 ArcRider 76

5.4.5 Validation 77

5.4.6 Le cas de la ligne 5 78

5.5 Microsimulator : simulation des déplacements 80

5.5.1 PlanSort 81

5.5.2 Microsimulator 81

5.6 Visualize : visualisation de la simulation 82

5.6.1 TRANSIMS Visualizer 82

5.7 Discussion 84

CHAPITRE 6 ANALYSE DES DONNÉES D'ACHALANDAGE 86

6.1 Description de l'échantillon 87

6.2 Profil par ligne 88

6.3 Profil par station 89

6.4 Distribution temporelle des transactions 91

6.4.1 Distribution temporelle selon le numéro de transaction 94

6.4.2 Nombre de transactions par heure/origine 97

6.4.3 Distribution des transactions à la seconde 99

6.5 Discussion 103

CHAPITRE 7 CONCLUSION 105

7.1 Résumé des thématiques abordées 105

7.2 Réflexivité et perspectives 107

BIBLIOGRAPHIE 110

ANNEXES 117

LISTE DES TABLEAUX

Tableau 1.1 - Sommaire des activités de la STM 3

Tableau 2.1 - Objets associés au réseau de métro 8

Tableau 2.2 - Avantages et inconvénients liés à l'utilisation de la carte à puce 13

Tableau 2.3 - Exemples d'analyses à partir des données de transactions 14

Tableau 2.4 - Fichiers requis pour l'utilisation du service Google Transit 19

Tableau 3.1 - Départs par ligne-direction 25

Tableau 3.2 - Information fournie dans les enregistrements du fichier *stop_times.txt* 25

Tableau 3.3 - Caractéristiques opérationnelles calculées à partir des données GTFS (ligne 1) 26

Tableau 4.1 - Décompte des différents objets de la base de données 33

Tableau 4.2 - Support utilisé pour 25 transactions entre 5:48 et 15:55 le 17 mars 2010 36

Tableau 4.3 - Types de produits les plus fréquents lors de transactions uniques 43

Tableau 4.4 - Rang des stations selon le nombre de validations uniques et le nombre total de transactions 44

Tableau 4.5 - Temps de parcours par ligne pour les périodes de pointe et hors pointe selon les données GTFS 47

Tableau 4.6 - GTFS vs Modélisation pour la ligne 5, direction Snowdon 48

Tableau 4.7 - Échantillon de données GTFS modifiées 51

Tableau 4.8 - Provenance des nouvelles heures de départ et d'arrivée selon le type de station 51

Tableau 5.1 - Répertoire de projet TRANSIMS et son contenu 65

Tableau 6.1 - Résumé des échantillons utilisés 87

Tableau 6.2 Matrice origine-destination par ligne 89

Tableau 6.3 - Les 10 stations les plus fréquentées le 17 mars 2010 90

Tableau 6.4 – Les 10 stations les moins fréquentées le 17 mars 2010 91

Tableau 6.5 - Nombre de transactions par heure sur chaque ligne 93

LISTE DES FIGURES

Figure 1.1 - Schématisation de la méthodologie 4

Figure 2.1 - Exemple de voyage sur la ligne bleue en direction Snowdon selon le service planifié 10

Figure 2.2 - Modèle relationnel simplifié (adapté de : Agard, Morency & Trépanier, 2006) 12

Figure 2.3 - Illustration du processus de dérivation de la station de destination en supposant un déplacement symétrique pour un no. de carte identifié 2 fois dans une journée 15

Figure 2.4 – Interface typique d'un SIG (ex. : QGIS) 17

Figure 2.5 - Base de données relationnelle GTFS (fichiers requis seulement) 19

Figure 3.1 - Plan du métro de Montréal 21

Figure 3.2 - Point de vente de rechargement dans le métro de Montréal 23

Figure 3.3 - Tourniquet du métro de Montréal 23

Figure 3.4 - Temps de parcours ligne verte direction Angrignon 26

Figure 3.5 - Temps interstations GTFS ligne verte direction Angrignon 27

Figure 3.6 - Vitesses GTFS ligne verte direction Angrignon 27

Figure 3.7 - Intervalle ligne verte direction Angrignon 28

Figure 4.1 – Structure d'une base de données de transactions de validation 32

Figure 4.2 - Processus de transaction de validation pour un déplacement métro 32

Figure 4.3 - Durée et nombre d'enregistrements par série 35

Figure 4.4 – Exemple de valeur suspecte 35

Figure 4.5 – Interface du logiciel Stata 38

Figure 4.6 - Distribution horaire des transactions uniques 43

Figure 4.7 Structure des données GTFS 46

Figure 4.8 - Positionnement des points de contrôle 49

Figure 4.9 - Exemple de calcul de coefficient 50

Figure 4.10 - Exemple de positionnement incohérent d'une rame de métro tiré des données GTFS originales et la correction apportée 52

Figure 4.11 Vitesse GTFS modifié ligne verte direction Angrignon 53

Figure 4.12 – Structure des données pour l'énumération des heures de passage 56

Figure 4.13 - Montants par départ sur la ligne bleue en direction Saint-Michel 57

Figure 4.14 - Profil de charge du départ de 16:11:00 en direction Saint-Michel 58

Figure 4.15 - Données de service planifié : nombre de véhicules en service 60

Figure 4.16 - Position des rames à 7 h 30 min sur la ligne 5 en direction Snowdon 60

Figure 5.1 - TRANSIMS Studio 63

Figure 5.2 - Schéma organisationnel de l'utilisation de TRANSIMS 64

Figure 5.3 - Lieux d'activité et couches telles qu'interprétées par TRANSIMS (FHWA TMIP, 2005) 66

Figure 5.4 - Module ConvertNet.py 69

Figure 5.5 - Réseau de métro codifié dans TRANSIMS et illustré avec QGIS 72

Figure 5.6 - Station Berri-UQAM codifiée dans TRANSIMS et illustrée avec QGIS 72

Figure 5.7 - Module RouterRuns.py 74

Figure 5.8 - Profil de charge du métro de Montréal pour la journée du 17 mars 2010 76

Figure 5.9 - Comparaison des résultats obtenus par simulation à ceux obtenus manuellement 78

Figure 5.10 - Matrice des temps de parcours pour un déplacement sur la ligne bleue 79

Figure 5.11 - Matrices OD des distances interstations 80

Figure 5.12 - Module Microsimulator.py 81

Figure 5.13 - Module Visualize.py 82

Figure 5.14 TRANSIMS Visualizer 83

Figure 5.15 - Illustration de l'évolution du profil de charge dans le temps avec TRANSIMS 84

Figure 6.1 - Imbrication des requêtes spatio-temporelles (tiré de Klein, 2007) 86

Figure 6.2 - Niveau de résolution et étendue de l'analyse par ligne 88

Figure 6.3 – Les 10 stations les plus fréquentées 90

Figure 6.4 - Nombre de transactions selon l'heure du jour 92

Figure 6.5 Profil de charge à différents moments de la journée 94

Figure 6.6 - Transactions par heure selon le numéro de transaction 95

Figure 6.7 - Durée des activités selon le numéro de séquence de l'activité 97

Figure 6.8 - Achalandage temporel à Berri-UQAM 98

Figure 6.9 - Provenance des usagers dont la destination est Berri-UQAM selon l'heure du jour . 99

Figure 6.10 - Transactions à la seconde à la station Vendôme 100

Figure 6.11 - Transactions à la seconde à la station Henri-Bourassa 101

Figure 6.12 - Transactions à la seconde à la station Crémazie 102

Figure 6.13 - Distribution corrigée à la station Crémazie 103

LISTE DES SIGLES ET ABRÉVIATIONS

AMT	Agence Métropolitaine de Transport
CMM	Communauté Métropolitaine de Montréal
GTFS	General Transit Feed Specification
RMM	Région Métropolitaine de Montréal
RMR	Région Métropolitaine de Recensement
SIG	Système d'information géographique
SIG-T	Système d'information géographique en transport
SIG-TOO	Système d'information géographique en transport orienté-objet
STI	Système de Transport Intelligent
STM	Société de Transport de Montréal

LISTE DES ANNEXES

Annexe A : Procédures de traitement des données dans STATA……...……117

Annexe B : Commandes STATA……119

Annexe C : Paramètres des fichiers de contrôle TRANSIMS……....120

Annexe D : Fichiers input TRANSIMS……..…133

CHAPITRE 1 INTRODUCTION

La planification du transport en commun a entrepris un virage technologique depuis l'intégration des systèmes d'acquisition de données automatisés. Plusieurs sociétés de transport à travers le monde ont équipé leur infrastructure de GPS, de compteurs de passagers, et de systèmes de perception par cartes à puce. Les réseaux de transport en commun peuvent maintenant être sondés quotidiennement, sans avoir à passer par les méthodes traditionnelles d'enquêtes de déplacements.

À la base, ces systèmes ont été installés dans le but d'évaluer et de valider les recettes du réseau de transport en commun. Rapidement, il a été reconnu que les données récoltées pouvaient aussi servir à quantifier précisément l'achalandage, la productivité du service, et la ponctualité des véhicules.

Ces nouvelles technologies permettent de récolter des ensembles de données riches en information. Cependant, le traitement de ceux-ci doit être réalisé systématiquement avant de procéder à leur analyse. Le nombre d'enregistrements contenus dans la base de données, la période couverte par une série d'enregistrements, et les points d'enregistrements sont des caractéristiques fondamentales à l'analyse d'une base de données. Elles permettent de cerner les dimensions et le cadre spatio-temporel de la situation observée. La particularité des données de transactions de validation effectuées dans un réseau de métro est qu'elles sont faites à des points fixes dans l'espace, ce qui allège les traitements à appliquer aux données.

Une fois la logique d'enregistrement des transactions (de validation) reconstituée, les données peuvent subir différents traitements afin d'être enrichies. De plus, différents ensembles de données peuvent être jumelés afin de raffiner les analyses qui en découlent. Enfin, les données peuvent être intégrées à un logiciel de modélisation afin d'en valider les procédures de calcul.

Ainsi, le défi méthodologique de l'analyse des systèmes de transport, défini par Manheim (1979), prend toute sa pertinence : « The methodological challenge of transportation systems analysis is to conduct a systematic analysis in a particular situation which is valid, practical, and relevant, and which assists in clarifying the issues to be debated. » La nature totalement désagrégée des données et la quantité de données disponibles favorisent une grande diversité d'analyses pouvant être conduites. C'est la rigueur avec laquelle les données sont approchées qui sera déterminante de la justesse et de la précision de ces analyses.

À ce titre, plusieurs procédures de traitement et d'analyse ont été développées et testées sur des ensembles de données de transport en commun. De plus, divers outils sont disponibles pour permettre la manipulation, l'illustration, et la modélisation des données. À cet effet, on retrouve les systèmes d'information géographique (SIG), les logiciels de statistique, et les logiciels de modélisation.

1.1 Contexte

1.1.1 Technologie

L'arrivée des technologies d'acquisition de données automatisée s'est faite progressivement dans les différentes autorités organisatrices de transport à partir de la fin des années 1990. En l'occurrence, la carte à puce s'est vue devenir le standard en matière de mode de perception. On l'utilise sur les réseaux de transport de plusieurs villes à travers le monde, notamment : Londres, Séoul, Sydney, Santiago, New York et Montréal. C'est en 2009 que le projet d'implantation du système de perception par carte à puce s'est complété sur le réseau de transport en commun de la région métropolitaine de recensement (RMR) de Montréal.

1.1.2 Population et Déplacements

La population de la RMR compte 3 940 000 personnes selon l'enquête origine-destination 2008 (Agence métropolitaine de transport, 2008). Cette même enquête révèle que 18,5% des 8 000 000

déplacements quotidiens effectués sur le territoire de la RMR sont complétés en transport en commun. Compte tenu du fait que la carte à puce est à ce jour, avec le paiement en espèce, l'unique mode de perception du réseau de transport en commun, tous les déplacements effectués sur le réseau sont comptabilisés automatiquement. Ainsi, il devient possible de comparer l'ordre de grandeur des résultats de l'enquête OD et des données de transactions recueillies quotidiennement.

1.1.3 Société de transport

La Société de Transport de Montréal (STM) agit à titre d'exploitant du réseau de transport en commun de la ville de Montréal. La STM est responsable de la gestion, de la planification, de l'exploitation et de l'entretien de (STM, 2012a, STM, 2012c, et STM, 2011c) :

Tableau 1.1 - Sommaire des activités de la STM

Réseau de surface (bus)	**Réseau sous-terrain (métro)**
1696 bus	756 voitures
Environ 8500 arrêts	68 stations
213 lignes	4 lignes
9 042 employés (2011)	
Budget annuel de 1,227 G $ (2012)	

1.2 Objectifs

C'est dans ce contexte technologique et opérationnel que s'inscrit cette recherche qui vise à définir des méthodes d'exploitation de données de transactions effectuées dans le réseau de métro de Montréal en :

- identifiant la littérature qui se rapporte à la thématique des cartes à puce;
- définissant les objets qui constituent le réseau de métro;
- proposant une approche systématique face à la gestion et à la manipulation des données de service planifié et de transactions du métro dans le but d'en faire l'exploitation;

- développant des procédures d'extraction de données d'achalandage, puis en intégrant ces données dans des analyses types;
- procédant à la codification et à la modélisation du métro de façon dynamique à l'aide d'outils de simulation;

1.3 Contenu

Ce projet se divise en 4 parties qui se rejoignent toutes en certains points : la composition d'un réseau de métro et les caractéristiques spécifiques du réseau de métro de Montréal, la présentation des bases de données et des différentes méthodes de raffinement et d'exploitation, la modélisation du métro de Montréal, et l'analyse des résultats. La Figure 1.1 illustre les relations entre les différentes sections du projet de recherche.

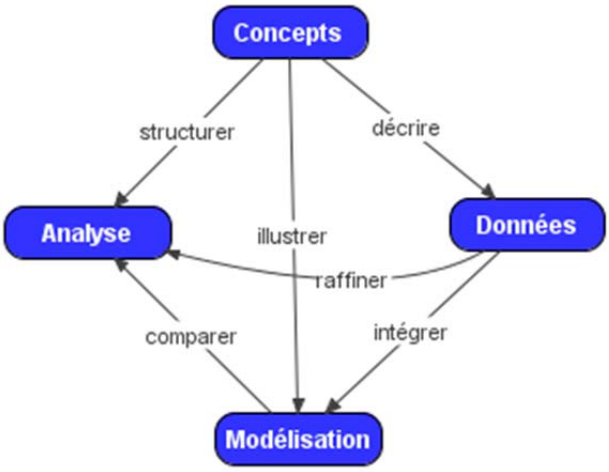


Figure 1.1 - Schématisation de la méthodologie

Le chapitre 2 est une introduction aux notions de base qui seront réutilisées tout au long de ce projet. Les thématiques abordées dans ce chapitre sont : les réseaux et les objets qui les décrivent, les données de cartes à puce, les technologies de l'information, et la modélisation. Une revue de la littérature publiée sur ces thématiques sera dressée dans le but d'en résumer les connaissances actuelles.

Le chapitre 3 présente un portrait du métro de Montréal : lignes et stations, perception et tarification, matériel roulant, planification et coûts d'exploitation. Ces détails permettront de se définir un ordre de grandeur du métro de Montréal, et son importance au niveau du transport en commun à Montréal.

Le chapitre 4 a trait à la gestion de données. L'intégration des technologies embarquées en transport en commun permet la génération quotidienne de très grands ensembles de données. Ceux-ci ne peuvent être utilisés que si des méthodes de gestion, de traitement et d'analyse de données efficaces sont développées. De plus, les ensembles de données peuvent provenir de différentes sources. C'est pourquoi, dans le but de profiter au maximum du potentiel des données, il est primordial de pouvoir joindre ces données entre elles. Pour ce faire, il est requis de se familiariser avec les données, et d'utiliser des outils qui permettront de traiter les données dans des temps raisonnables vu la taille des bases de données.

Le chapitre 5 présente la démarche de modélisation du métro de Montréal à l'aide du logiciel TRANSIMS. Les résultats obtenus dans ce chapitre sont comparés à ceux obtenus au chapitre précédent pour valider le fonctionnement du logiciel. Par ailleurs, TRANSIMS permet la visualisation dynamique des déplacements. En chargeant une demande dans le logiciel, il est possible de voir la variabilité de la charge de chaque rame dans le temps. Toutes les étapes pour arriver à ce résultat sont décrites, et les paramètres utilisés sont fournis en annexe.

Le chapitre 6 couvre les différentes pistes d'analyses qui peuvent être conduites en assemblant les différents ensembles de données. Les dimensions spatiales et temporelles sont explorées, auxquelles est combinée la caractérisation du type de transaction. Les résultats sont présentés sous forme de tableaux, de graphiques et de cartes géoréférencées.

CHAPITRE 2 FONDEMENTS

Les systèmes de collecte de données automatisés, dont le système de perception par carte à puce, forment un sous-ensemble de la famille des systèmes de transport intelligent (STI). La définition des STI est donnée par Bergeron (2008) :

> Système interactif de collecte, de traitement et de diffusion d'information appliqué aux transports, basé sur l'intégration des technologies de l'information et de la communication aux infrastructures et aux véhicules utilisés, de manière à améliorer la gestion et l'exploitation des réseaux de transport et des services aux utilisateurs qui y sont associés.

Changements dans les paradigmes

La définition des STI telle que formulée par Bergeron (2008), est une introduction aux changements dans les paradigmes auxquels les autorités organisatrices de transport font face. En effet, comme l'explique Chu (2010) :

> Récemment, des changements aux paradigmes remodèlent ce processus (celui d'interaction entre l'offre et la demande). D'une part, les opérateurs de transport en commun adoptent une orientation axée sur la performance et le client. (…) D'autre part, l'avancement des systèmes automatiques de collecte de données et leur adoption par les opérateurs génèrent une abondance de données dans un environnement où la collecte de données était auparavant limitée par le coût et la main d'œuvre des méthodes manuelles.

Les sociétés de transport se sont constituées un environnement riche en données grâce aux STI. La rentabilisation de systèmes tels que la perception par carte à puce réside d'abord dans la réduction de la fraude, mais aussi dans l'évaluation automatique des recettes, et dans la caractérisation du réseau. En ce sens, la littérature propose différentes approches pour profiter, entre autres, du niveau de résolution des données, de leur caractère spatio-temporel, et du fait qu'elles soient enregistrées en continu. Conjuguées à l'application des concepts de réseau de transport en commun, les données permettent d'avoir une meilleure connaissance de l'utilisation du service dans le temps et l'espace. Notamment, dans le réseau de métro de Montréal, les transactions par cartes à puce enregistrées tous les jours permettent de connaître le profil d'utilisation de chaque station dans le temps, et ce, à la seconde près. Il est donc possible, contrairement à l'approche traditionnelle d'une journée moyenne, de planifier le service en tenant

compte des événements spéciaux, des conditions climatiques, ou d'autres conditions particulières en différents points de la ville (Zhao, 2004).

Ce chapitre énonce d'abord les concepts liés à la définition d'un réseau de transport en commun. Par la suite, les implications de l'utilisation de la carte à puce comme mode de perception dans un réseau de métro y sont présentées : fonctionnement, avantages et inconvénients, potentiel d'utilisations, traitement des données et méthodes de dérivations d'information implicite. Puis, les différentes applications des technologies de l'information disponibles et utiles à la planification du transport en commun sont explorées. Enfin, un survol des approches de modélisation du transport en commun est effectué.

2.1 Réseau de transport en commun

Un réseau correspond à un ensemble de liens, connectés par des nœuds, selon une certaine géométrie. La géométrie d'un réseau de transport (en commun) réfère aux différentes composantes de type nœud et lien qui définissent son « apparence » et sa structure (Chapleau, 2009). Le service de transport en commun est assuré par des véhicules qui circulent le long des lignes (liens), qui elles, sont constituées d'arrêts, ou de stations (nœuds). La flotte est opérée par l'autorité de transport responsable du territoire desservi. L'usager se voit payer un montant, fixe ou variable, pour accéder au service de transport en commun. Le service est disponible à chaque jour de la semaine, selon un service défini, et peut être offert la nuit selon un service réduit.

2.1.1 Approche orientée objet

La gestion de grands ensembles de données appelle à l'emploi d'une structure définie pour systématiser la classification des concepts. L'approche orientée objet en transport en commun offre un cadre structurant aux données et aux procédures de traitement (Trépanier et al., 2004). Les 4 classes de l'approche orientée objet en transport en commun sont les suivantes (Trépanier, Chapleau, 2001a) :

- les objets *statiques* possèdent une localisation fixe dans le temps et l'espace. Leur rôle est de décrire le territoire et servir de balise à l'espace utilisable pour des fins de transport. Quelques exemples : les générateurs de déplacements, les codes postaux, les secteurs de recensement et les zones.
- les objets *dynamiques* sont les acteurs du transport. Ces objets « décident » et contribuent à leur mouvement. Ils représentent tantôt un groupe de personnes (ménages et collections de personnes), un objet mobile (train, automobile) ou une masse déplacée (telle qu'une marchandise).
- les objets *cinétiques* sont les descripteurs du mouvement. Quelques exemples : le déplacement, la chaîne de déplacements, le trajet, la ligne d'autobus.
- les objets *systémiques* sont des groupes d'objets interreliés. Ils peuvent être opérationnels (réseau routier, réseau de métro), informationnels (enquêtes, recensements) ou multifonctionnels (centre-ville, banlieue).

2.1.2 Objets associés au métro

Les objets associés aux réseaux de métro sont les suivants :

Tableau 2.1 - Objets associés au réseau de métro

Objet	Classe	Définition
Réseau	Systémique	Un réseau (de métro) est un regroupement de lignes connectées ou qui se croisent entre elles, et qui sont coordonnées pour offrir un service profitable pour les usagers et pour assurer l'efficacité opérationnelle (Vuchic, 2005).
Ligne	Cinétique	Une ligne est l'infrastructure à géométrie fixe sur laquelle les véhicules circulent pour fournir le service, opérant selon un horaire prédéterminé (Vuchic, 2005).
Station	Statique	Une station est une installation spécifique pour l'embarquement et le débarquement, l'attente et le processus de transfert des usagers. Elles peuvent contenir différentes composantes destinées à l'usager telles que des quais, une mezzanine, des escaliers et l'équipement de perception, et d'autres destinées à l'opération des véhicules, telles que des plateformes de retournement et d'entreposage (Vuchic, 2005).

Tableau 2.1 - Objets associés au réseau de métro (suite)

Station de correspondance	Statique	Une station de correspondance joint 2 ou plusieurs lignes, permettant ainsi aux usagers de passer de l'une à l'autre des lignes connectées (Vuchic, 2005).
Terminus	Statique	Un terminus est la dernière station d'une ligne (Vuchic, 2005).
Édicule	Statique	Entrée d'une station de métro, généralement située au niveau de la rue (OQLF, 2011).
Tourniquet	Statique	Dispositif rotatif qui gère l'accès des personnes à une zone tarifée (OQLF, 2011).
Quai	Statique	Construction surélevée facilitant l'accès des voyageurs et le chargement et le déchargement des véhicules sur rails (OQLF, 2011).
Rame	Dynamique	Suite de (voitures) traînées par un même moteur. On réserve le terme « wagon » (freight car) à tout véhicule transportant des marchandises et celui de « voiture » (passenger car) à tout véhicule transportant des voyageurs. (OQLF, 2011)
Voiture	Dynamique	Voiture, remorque d'automotrice ou automotrice destinée au transport des voyageurs, comportant ou non un ou plusieurs compartiments ou emplacements spéciaux pour les bagages, les colis, la poste, etc. (OQLF, 2011).

2.1.3 Planification du service et opération

La planification du réseau de métro consiste à déterminer l'intervalle adapté à la desserte du réseau, c'est-à-dire le temps maximal entre les passages des rames pour satisfaire à la demande du réseau. L'intervalle est adapté aux variations d'achalandage temporelles, qui dépendent du jour (semaine, samedi, dimanche), et de l'heure de la journée. De plus, le service de transport en commun est révisé 4 fois par année pour tenir compte des variations d'achalandage saisonnières. C'est en fonction de ces contraintes que sont déterminées les heures de passages de chaque rame de métro, à chacune des stations du réseau de métro. La progression d'une rame de métro, pour un voyage dont le départ est planifié à 7h48, est illustrée en exemple à la Figure 2.1.

Les rames de métro, pour la plupart, fonctionnent par pilotage automatique. Un centre de contrôle permet généralement la surveillance des stations à distance et la gestion de la circulation des rames sur le réseau.

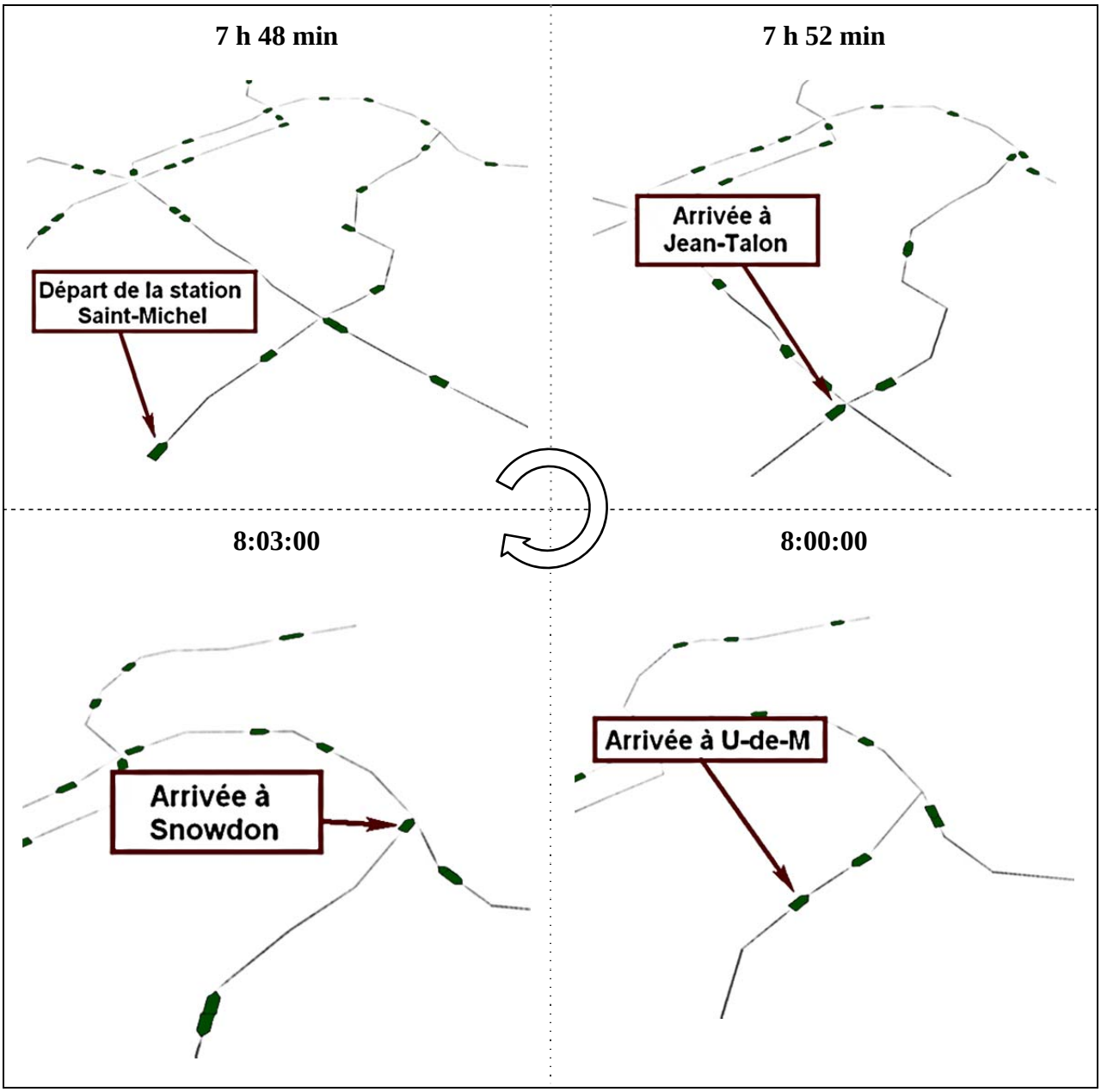


Figure 2.1 - Exemple de voyage sur la ligne bleue en direction Snowdon selon le service planifié

2.1.4 Perception et tarification

Dans un réseau de transport en commun, la perception se divise en deux principales composantes : la tarification et le mode de perception. D'une part, la tarification peut être fixe, avec un seul prix peu importe la destination, ou bien variable, où la distance parcourue détermine le tarif déboursé. Le prix peut varier directement selon la distance, le nombre de correspondances permises, ou encore être délimité par des zones géographiques. Le tarif varie aussi selon le type d'utilisateur. En effet, un étudiant bénéficiera la plupart du temps d'un rabais sur le tarif régulier.

D'autre part, le mode de perception correspond au support utilisé pour payer son entrée dans le réseau. Il peut s'agir, entre autres, d'argent comptant, d'un ticket magnétique, d'une correspondance, ou d'une carte à puce. Cette dernière nous intéresse plus particulièrement dans le projet actuel.

2.2 La carte à puce en transport en commun

La définition d'une carte à puce, dans un contexte de transport en commun, est établie par Bagchi & White (2005) de la façon suivante :

> Smart cards are similar in look and size to credit cards and can be used in place of traditional fare media such as paper tickets and magnetic stripe cards. Each smart card can be identified by a unique serial number. The cards can be registered to a given individual, or they can be anonymous. On these cards can be placed electronically a range of fare options such as travelcards or stored value (a monetary amount credited to the card which is debited as and when journeys are made).

2.2.1 Fonctionnement du système

Dans un réseau de transport en commun utilisant la carte à puce comme mode de perception, l'usager doit d'abord se procurer sa carte dans un point de vente équipé pour identifier et initialiser les cartes. Par la suite, il doit l'alimenter en achetant son titre dans un point de vente de rechargement. Lors de son entrée dans le réseau de métro, l'usager doit valider son titre à

l'appareil prévu à cet effet, situé dans le tourniquet. La transaction est acceptée ou refusée selon 2 critères : la date d'expiration et la compatibilité du titre. L'information relative à la transaction, soit la date et l'heure, le type d'entrée, le tarif, le numéro de carte et le code de refus, est par la suite acheminée vers un serveur, qui lui, stocke les données. Le système de tarification du réseau de métro de Montréal est à paiement unique : seule l'identification de la station d'embarquement est enregistrée dans le serveur. La Figure 2.2 illustre le modèle relationnel simplifié d'une base de données typique.

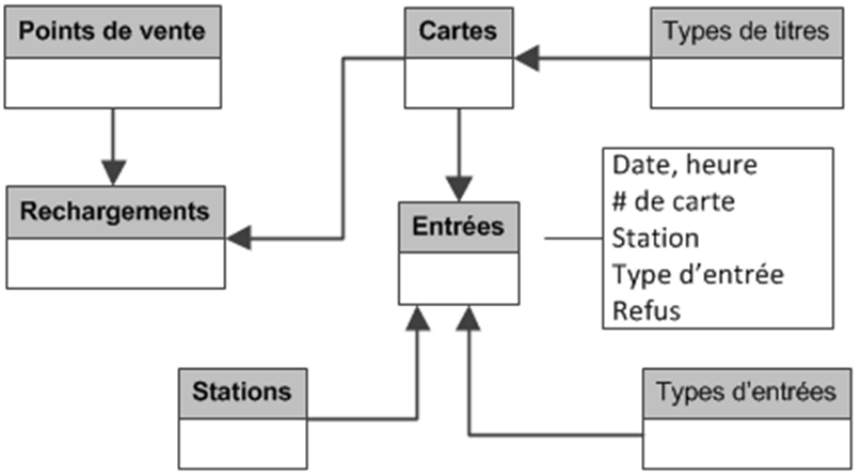


Figure 2.2 - Modèle relationnel simplifié (adapté de : Agard, Morency & Trépanier, 2006)

2.2.2 Avantages et inconvénients

Certains des avantages et inconvénients liés à l'utilisation de la carte à puce comme mode de perception en transport en commun par rapport au mode de paiement traditionnel sont identifiés par Vuchic (2005) :

Tableau 2.2 - Avantages et inconvénients liés à l'utilisation de la carte à puce

Avantages	Inconvénients
• Réduction de la fraude • Diminue l'intervention humaine dans la collecte des données • Génère des statistiques d'achalandage complètes	• Nécessite une infrastructure plus complexe (valideuses, délimitations de zones tarifaires) • Investissement initial élevé

Une fois le système implanté, les données de transactions peuvent être analysées quotidiennement, et à très faible coût comparativement à une enquête sur les déplacements (Zhao, 2004).

2.2.3 Bénéfices d'utilisation

Brevetée il y a plus de 40 ans (Schelfer, Procaccino, 2002), la carte à puce a progressivement été introduite dans les sociétés de transport à partir de la fin des années 1990 (Blythe, 2004). Aujourd'hui, plusieurs sociétés de transport à travers le monde ont adopté ce mode de perception. D'une part, les usagers du service peuvent enregistrer tous leurs titres de transport sur une seule et même carte. D'autre part, Bagchi et White (2005) identifient les possibilités, pour l'opérateur du réseau, qui émergent de l'introduction de cette technologie en transport en commun :

1. Accès à un large volume de données sur les trajets individuels;
2. Association des données de transactions à un usager;
3. Observation des transactions des mêmes usagers sur de longues périodes de temps;
4. Identification de la clientèle la plus régulière.

Utsonomiya, Attanucci & Wilson (2006) identifient les différentes analyses rendues possibles avec les ensembles de données générés, pour lesquelles un exemple a été relevé :

Tableau 2.3 - Exemples d'analyses à partir des données de transactions

Analyse	Exemple de la littérature
Exploitation informationnelle des données de transactions par cartes à puce	Trépanier, Chapleau, Tranchant (2005)
Analyses comportementales des usagers à partir des données de transactions par cartes à puce	Agard, Morency, Trépanier (2006)
Habitudes de déplacements des usagers à partir de données de cartes à puce	Chapleau & Chu, (2007)

2.2.4 Problématiques de traitement des données

Les possibilités d'analyses qui émergent de l'utilisation de la carte à puce en transport en commun sont nombreuses. Toutefois, plusieurs problématiques liées au traitement des données restent présentes, telles que les erreurs systématiques et l'absence de point de destination (Trépanier, Chapleau, Tranchant, 2005). De plus, certaines erreurs liées à l'utilisation de l'équipement de perception (cartes, valideuses, serveurs) peuvent affecter la qualité des données. Des transactions peuvent être manquantes (Okamura et al., 2004) et (Utsunomiya, 2006) ou mal identifiées dû à de l'équipement défectueux. Les données de transaction de cartes à puce ne peuvent donc pas être utilisées directement. En effet, pour apprécier le plein potentiel des données disponibles, il faut d'abord passer par des procédures de traitement et d'enrichissement (Chu, Chapleau, 2008). L'intervention humaine lors de l'extraction des données est une autre source d'erreur à laquelle il faut porter attention. Les traitements auxquels les données seront exposées doivent donc être systématiques pour s'assurer de la conformité des données aux besoins des analyses.

2.2.5 Dérivation des points de débarquement

La dérivation des points de débarquement est une préoccupation que l'on retrouve dans les réseaux à validation unique, comme dans le métro de Montréal. En effet, puisque la validation du titre se fait seulement à l'entrée du réseau, le point de descente n'est pas connu. Cependant, Barry et al. (2002) proposent une méthode de dérivation des stations de destination appliquée au réseau de métro de la ville de New York (NYCTA). Cette méthode est basée sur deux hypothèses indépendantes l'une de l'autre :

- La station de destination du déplacement métro actuel d'un usager est la même que la station d'origine du déplacement métro suivant;
- La station de destination du dernier déplacement métro de la journée d'un usager est la même que la station d'origine du premier déplacement de la journée.

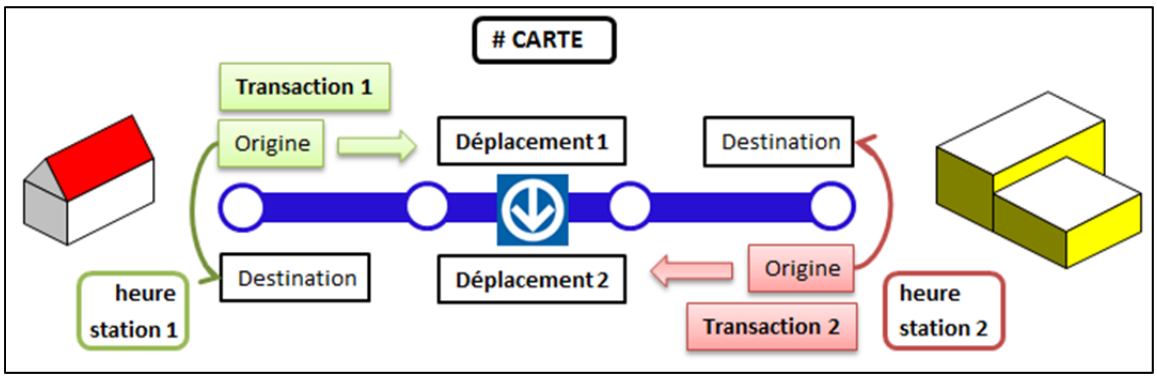


Figure 2.3 - Illustration du processus de dérivation de la station de destination en supposant un déplacement symétrique pour un no. de carte identifié 2 fois dans une journée

Seuls les déplacements métro sont pris en compte pour reconstruire les itinéraires avec la méthode présentée. D'autres recherches se penchent sur la reconstitution des matrices origine-destination des déplacements bimodes, entre autres : Zhao (2004) sur le réseau de transport en commun de Chicago, et Munizaga, Palma & Fischer (2011) sur le réseau de Santiago.

2.3 Technologie de l'information

Les ensembles de données avec lesquels les planificateurs et analystes en transport doivent maintenant travailler sont de dimensions trop grandes pour être traités manuellement. Il est donc nécessaire d'adapter l'utilisation des outils disponibles aux besoins actuels, afin de ne pas s'engager dans des dépenses de temps inutiles. Les systèmes d'information géographique, les tableurs, et les logiciels de modélisation permettent d'appliquer différents traitements aux données, et ce, de façon automatisée.

2.3.1 Système d'information géographique en transport

Selon Klein (2011), les représentations spatio-temporelles (plus particulièrement le caractère spatial de la mobilité) sont essentielles à la connaissance des structures et des processus qui surviennent quotidiennement dans un environnement urbain. Un des outils qui permettent la construction de telles représentations est le système d'information géographique (SIG). Un SIG est un système qui intègre une interface graphique, un gestionnaire de base de données, des outils de modélisation spatiale et des fonctions de cartographie, de gestion, d'analyse et de diffusion de l'information spatiale et de ses attributs (Dueker & Peng, 2008). Les SIG permettent donc de représenter de l'information qualitative, telle qu'un arrêt de bus ou un bâtiment, et quantitative, tel qu'une vitesse ou un profil de charge, selon un système de coordonnées de référence. Pour chaque objet codifié, il est possible de spécifier des attributs qui seront stockés dans le gestionnaire de base de données. La codification des objets se fait sur des couches pouvant prendre trois formats : point, ligne ou polygone, tel qu'illustré à la Figure 2.4. Ce système de codification, dit vectoriel, permet de changer la résolution de la représentation sans perdre de précision visuelle.

L'évolution des systèmes d'information géographique en transport (SIG-T) se décompose en trois étapes selon Thévenin (2011) : l'étude cartographique bidimensionnelle (nœuds, arcs), le développement d'outils d'aide à la navigation (chemin le plus court, calculs sur les attributs) et la représentation du comportement d'objets discrets (véhicules, personnes). En couplant à ces fonctionnalités l'approche orientée-objet, la représentation des réseaux de transport en commun

se précise davantage. Comme le mentionnent Trépanier & Chapleau (2001b), pour bénéficier pleinement des systèmes d'information géographique en transport, il est possible de lier l'information opérationnelle du réseau de transport en commun au réseau viaire à travers un SIG-T orienté-objet (SIG-TOO). Une fois les relations entre les objets des différents réseaux définies, le SIG-TOO permet de décrire l'information opérationnelle et de considérer les objets dynamiques des ensembles de données intégrés.

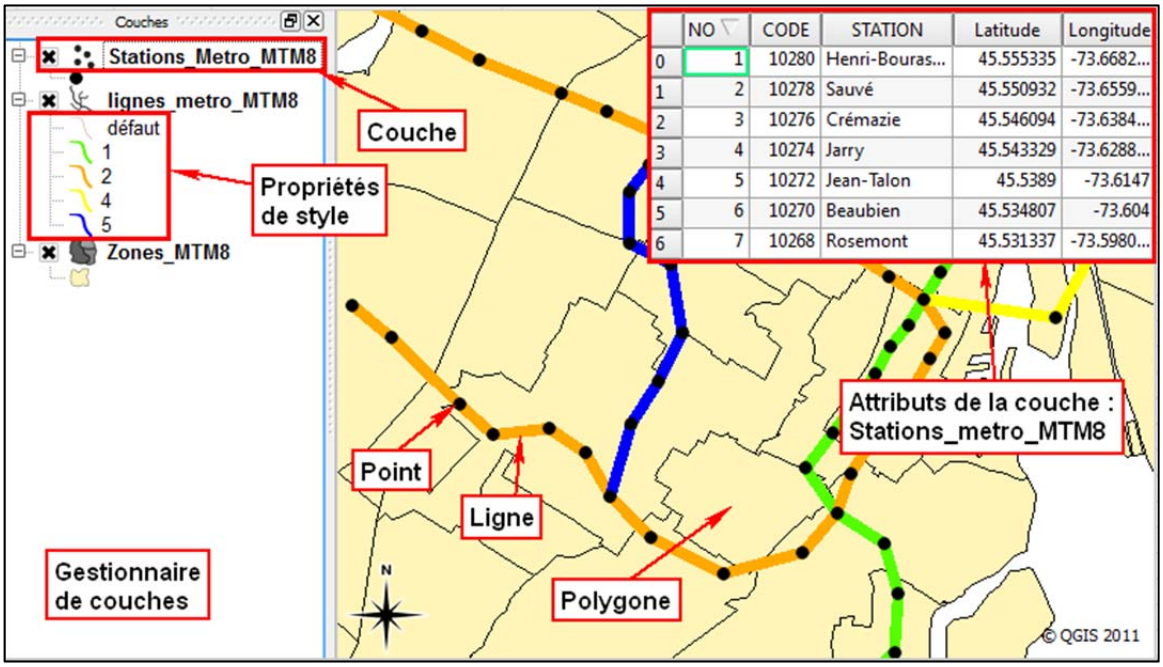


Figure 2.4 – Interface typique d'un SIG (ex. : QGIS)

Les sources de données se sont multipliées depuis l'intégration des systèmes d'information géographique en transport. Aussi, la taille des ensembles de données est de plus en plus grande : ceux-ci sont donc plus lourds à traiter. Toutefois, l'interopérabilité des SIG-T, c'est-à-dire l'ensemble des relations qui existent entre l'interface graphique, les modules de calculs et la base de données, conjuguée à l'approche orientée-objet, permet de tirer profit des différents ensembles de données.

2.3.2 Google Transit et données GTFS

L'usager du transport en commun peut planifier son déplacement en utilisant un calculateur de chemin en ligne. En intégrant les propriétés des SIG-T à une plateforme interactive, le calculateur permet d'évaluer le chemin le plus court entre une origine et une destination spécifiées, selon un mode de transport choisi.

Google Transit (Google Transit, 2012) est un des calculateurs de trajet actuellement disponibles. En 2012, ce sont les usagers de plus de 475 sociétés de transport à travers le monde qui bénéficient de ce service (Google Transit, 2012). Le service se base sur des données de service planifié selon le format *General Transit Feed Specification* (GTFS), développé par Google.

Pour assurer le fonctionnement du service, les sociétés de transports doivent transmettre à Google un certain nombre de fichiers, présentés à la Figure 2.5, décrivant la géométrie du réseau de transport en commun, de même que l'horaire planifié à chaque arrêt et station, dans un format spécifique. À chaque révision du service, les fichiers doivent être retransmis à Google. Les fichiers requis et leur contenu sont décrits au Tableau 2.4 (GTFS Google Code, 2012). De l'information supplémentaire facultative peut être fournie pour préciser d'autres aspects du réseau tels que la tarification et les correspondances possibles.

La force des données GTFS se trouve, d'une part, dans la documentation fournie par Google pour décrire le fonctionnement du système. La codification des fichiers, les données requises, le titre et le format des champs sont tous clairement définis sur le site du GTFS (GTFS Google Code, 2012). D'autre part, les données sont partagées dans un format systématique. En effet, le format GTFS établit un standard pour le partage d'un horaire de transport en commun (Ferris et al., 2010). Toujours selon Ferris et al. (2010), ce procédé permet aux usagers du transport en commun et aux sociétés de transport d'avoir rapidement accès à l'information, et offre la possibilité de développer des applications web ou mobiles.

Tableau 2.4 - Fichiers requis pour l'utilisation du service Google Transit

Fichier	Contenu
agency.txt	Information concernant la société de transport
stops.txt	Liste des arrêts du réseau de transport en commun et leur position géographique en coordonnées lat/long
routes.txt	Identification des lignes
trips.txt	Description des voyages
stop_times.txt	Heures d'arrivée et de départ des véhicules pour chaque arrêt pour la journée de service
calendar.txt	Spécifie la liste à laquelle l'horaire fait référence, sa date de début et de fin, et les journées pour lesquelles le service de transport en commun est offert

Les fichiers utilisés dans le cadre des expérimentations qui suivront décrivent la période de service de janvier 2010, pour un jour moyen de semaine.

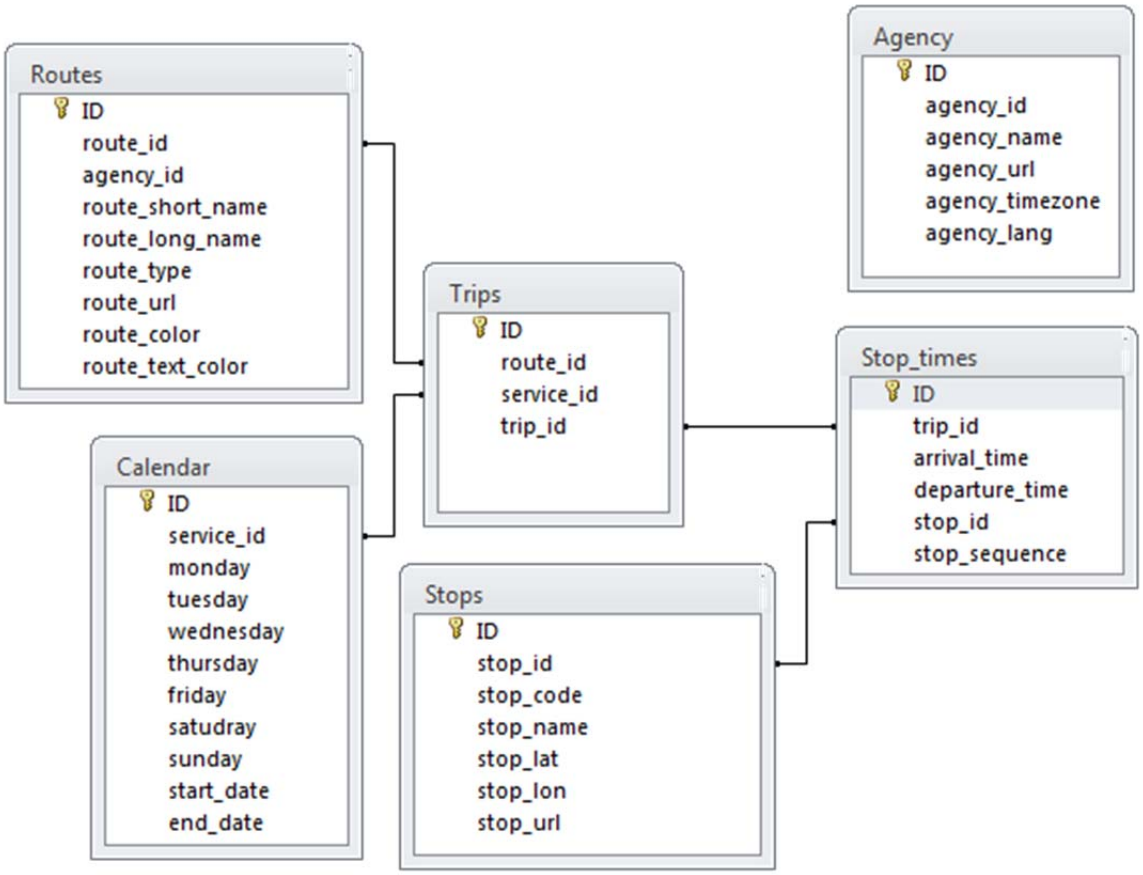


Figure 2.5 - Base de données relationnelle GTFS (fichiers requis seulement)

2.4 Modélisation

La modélisation est une étape incontournable de la prise de décision en transport en commun. La modélisation est une simplification des systèmes et des relations observées en réalité qui, en mettant l'accent sur les relations clés et les données disponibles, permet d'observer les comportements actuels et l'évolution de ceux-ci dans le temps, qu'on intervienne sur le réseau ou pas (Hensher & Button, 2008).

Cette définition fondamentale de Hensher & Button (2008) sera la ligne directrice pour la modélisation du métro de Montréal au chapitre 5. Chacun des aspects de cette définition est lié à une étape de la modélisation : la simplification des systèmes correspond à la codification et à l'interprétation de celle-ci par le logiciel, les relations clés et les données disponibles correspondent aux paramètres, aux intrants et aux extrants du modèle, et les comportements et l'évolution de ceux-ci sont le fruit de l'analyse des fichiers de résultats générés par le logiciel.

CHAPITRE 3 LE MÉTRO DE MONTRÉAL

Suite à son inauguration le 14 octobre 1966, le métro de Montréal est rapidement devenu le mode de transport le plus fiable pour les usagers du transport en commun. Plusieurs facteurs favorisent l'utilisation du métro, entre autres : l'élimination des aléas de circulation ou climatiques, les temps d'attentes courts, les économies de temps dues à la vitesse d'opération, et le développement urbain le long des lignes (Communauté urbaine de Montréal et al. 1983). Ainsi, l'achalandage du métro est passé d'environ 350 000 déplacements par jour dans les années '60-'70 (Communauté urbaine de Montréal et al. 1983), à près de 800 à 900 000 déplacements quotidiens aujourd'hui. Le réseau de métro dans sa forme actuelle est constitué de 68 stations reliées par 4 lignes, tel qu'illustré à la Figure 3.1 (STM, 2011a).

Figure 3.1 - Plan du métro de Montréal

3.1 Description du réseau

3.1.1 Lignes

Le réseau de métro est constitué de 4 lignes : la ligne 1 (verte), 2 (orange), 4 (jaune) et 5 (bleue), qui regroupent les liens interstations du réseau. Les lignes sont autonomes, c'est-à-dire qu'elles empruntent chacune un tunnel distinct. Cette configuration facilite la régulation du trafic par rapport au nombre de voyageurs à transporter, mais nécessite l'élaboration de stations de correspondance (Communauté urbaine de Montréal et al. 1983). Le territoire desservi par les lignes de métro s'étend de l'arrondissement du Sud-Ouest (à l'ouest) à Mercier-Hochelaga-Maisonneuve (à l'est), et de Laval au nord à Longueuil au sud.

3.1.2 Stations

Le métro contient 68 stations réparties sur les 4 lignes décrites précédemment. Chaque station possède un ou plusieurs édicules, qui permettent l'accès au réseau. On appelle « édicule » d'une station l'accès qui est le plus près du niveau de la rue. Des escaliers fixes et mobiles, et dans certains cas des ascenseurs (aux stations Montmorency, de la Concorde, Cartier, Henri-Bourassa, Berri-UQAM, Lionel-Groulx et Côte-Vertu) donnent accès au niveau inférieur : la mezzanine, où on retrouve généralement les tourniquets. Les autres niveaux sont les transitoires, et enfin, les quais. Les stations de correspondances sont Berri-UQAM (ligne jaune, verte et orange), Jean-Talon (ligne bleue et orange), Snowdon (ligne bleue et orange) et Lionel-Groulx (ligne verte et orange). Les stations de métro sont toutes desservies par au moins 1 ligne d'autobus.

3.1.3 Matériel roulant

Le réseau de métro compte 756 voitures : 333 MR-63 et 423 MR-73 (STM, 2012a). Une rame de métro est constituée de 3, 6, ou 9 voitures. Une rame de 9 voitures peut transporter jusqu'à 1200 passagers, dont 360 assis (STM en bref, 2011). Le remplacement complet du matériel roulant a été confirmé en février 2011 (Corriveau, 2011), alors que le gouvernement du Québec a accordé ce mandat au consortium Bombardier-Alstom.

3.1.4 Perception et tarification

La carte à puce a été implantée sur le réseau de transport en commun de la région métropolitaine de Montréal en 2009. C'est ce mode de perception qui est utilisé dans le métro, les bus et les trains. Selon la STM, la carte à puce peut être gardée par un même usager pendant 2-7 ans, et permettrait de récupérer jusqu'à 20 M$ en fraude annuellement (Mouvement Collectif, 2009). Cependant, une décision a été prise pour limiter la durée de vie des cartes à 4 ans afin d'éviter les défaillances éventuelles (Corriveau, 2012).

Figure 3.2 - Point de vente de rechargement dans le métro de Montréal

Figure 3.3 - Tourniquet du métro de Montréal

La tarification dans le métro de Montréal varie selon le type de titre que l'usager se procure. Un passage coûte 3.00$ à l'unité, tandis qu'un abonnement mensuel coûte 75.50$ au tarif régulier, et 43.75$ au tarif réduit pour étudiants et personnes âgées de 65 ans et plus (STM, 2012b). Il faut noter qu'un abonnement mensuel local ne sera pas utilisable aux stations situées à l'extérieur de l'île. Pour un abonnement mensuel, ces stations requièrent l'achat d'un titre de type « TRAM 3 », dont le prix est fixé par l'AMT.

Des abonnements hebdomadaires, quotidiens, ou pour trois jours sont aussi vendus. Les tarifs varient d'une société de transport à l'autre, et un titre acheté dans une société de transport en particulier ne peut être utilisé dans une autre société de transport, à l'exception du tarif TRAM.

3.1.5 Planification du service

Le service est planifié pour les journées de semaine, samedi, et dimanche séparément. Selon l'horaire disponible sur site de la STM, les intervalles de service sont les plus courts en période de pointe lors des jours de semaine, atteignant les 3 minutes au minimum, et en milieu de journée la fin de semaine, où les intervalles minimaux sont d'environ 6 minutes entre chaque départ. Quelques indicatifs sur le service planifié de la ligne 1 sont fournis au Tableau 3.3, de même qu'aux Figure 3.4, Figure 3.5, Figure 3.6 et Figure 3.7.

Le service de métro est disponible à partir de 5 h 30 le matin, et ce jusqu'à 0 h 30 en semaine et 1 h la fin de semaine.

3.1.5.1 Service planifié en format GTFS

Les intervalles de service du métro de Montréal sont entièrement disponibles sur le site de la STM. Par ailleurs, le service est aussi planifié en format GTFS, tel que présenté à la section 2.3.2. Les fichiers de service planifié sont générés dans un format précis, et transmis à Google. C'est de cette façon que l'outil Google Transit permet de déterminer un itinéraire de transport en commun à Montréal.

L'essentiel du service planifié GTFS est contenu dans deux fichiers, soient stops.txt et stop_times.txt. Ceux-ci décrivent respectivement la position, le nom et le numéro des arrêts de transport en commun, et les temps de passages de chaque véhicule à chaque arrêt et station. Les fichiers ont été traités pour ne tenir compte que des attributs qui concernent le métro.

Ainsi, le fichier *stops.txt*, contient 68 enregistrements qui fournissent la position et le numéro d'identification des stations de métro. Le fichier *stop_times.txt,* lui, contient tous les temps de passages des véhicules aux 68 stations de métro pour un jour moyen de semaine de la période de janvier 2010. Ce fichier est composé de 30 254 enregistrements, qui décrivent les temps de passages de 120 rames pour les 1581 départs planifiés dans la journée.

Dans la définition GTFS, le métro est divisé en 11 lignes-directions, ou tracés. Ainsi, on en compte 2 par lignes, sauf pour la ligne orange, qui se divise en 5 tracés avec des terminus aux stations Montmorency, Henri-Bourassa et Côte-Vertu, tel que présenté au Tableau 3.1.

Tableau 3.1 - Départs par ligne-direction

tracé	ligne	origine	destination	nb départs
1	1	Honoré-Beaugrand	Angrignon	208
2	1	Angrignon	Honoré-Beaugrand	208
3	2	Henri-Bourassa	Côte-Vertu	29
4	2	Henri-Bourassa	Montmorency	3
5	2	Côte-Vertu	Henri-Bourassa	25
6	2	Côte-Vertu	Montmorency	195
7	2	Montmorency	Côte-Vertu	191
8	4	Berri-UQAM	Long.-Université-Sherbrooke	174
9	4	Long.-Université-Sherbrooke	Berri-UQAM	174
10	5	Snowdon	Saint-Michel	187
11	5	Saint-Michel	Snowdon	187
	TOTAL	**-**	**-**	**1581**

Tableau 3.2 - Information fournie dans les enregistrements du fichier *stop_times.txt*

Année / période	Type Jour	Ligne	Véhicule	DépTerminus	arrival_time	departure_time	stop_id	stop_sequence
10J	S	1	1	05:30:00	05:30:00	05:30:00	43	1
10J	S	1	1	05:30:00	05:30:59	05:30:59	42	2
10J	S	1	1	05:30:00	05:33:00	05:33:00	41	3

L'information contenue dans le fichier *stop_times.txt* permet de calculer plusieurs caractéristiques opérationnelles du réseau de métro : nombre de véhicules utilisés, nombre de véhicules en service, nombre de départs, intervalle et fréquence, temps interstation, temps de parcours moyen, etc. La précision des heures d'arrivées aux stations est à la minute, et les heures d'arrivée et de départ sont codifiées comme étant les mêmes. Quelques exemples de caractéristiques calculées sont présentés, pour la ligne 1 (verte), dans le tableau qui suit :

Tableau 3.3 - Caractéristiques opérationnelles calculées à partir des données GTFS (ligne 1)

Caractéristique	***Direction***	***Moyenne***	***Min***	***Max***
Véhicules en service	Toutes	13	0	25
Intervalle (mm:ss)	Ang.	5:33	2:00	12:00
	H-B	5:32	2:00	13:00
Temps de parcours (mm:ss)	Ang.	37:35	36:00	40:00
	H-B	37:34	36:00	39:00

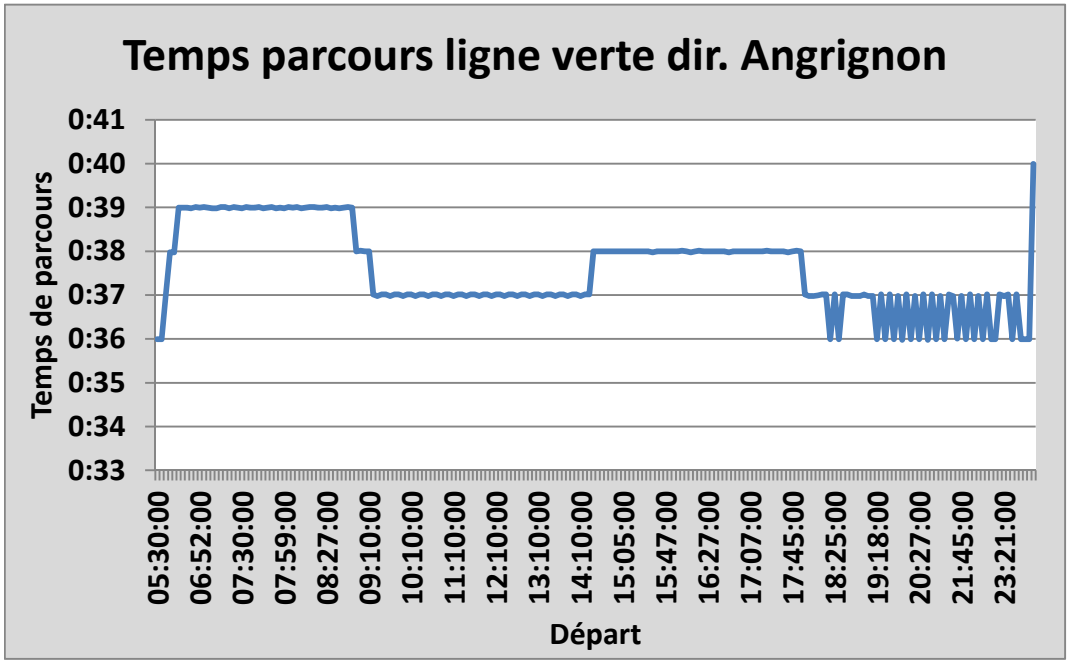


Figure 3.4 - Temps de parcours ligne verte direction Angrignon

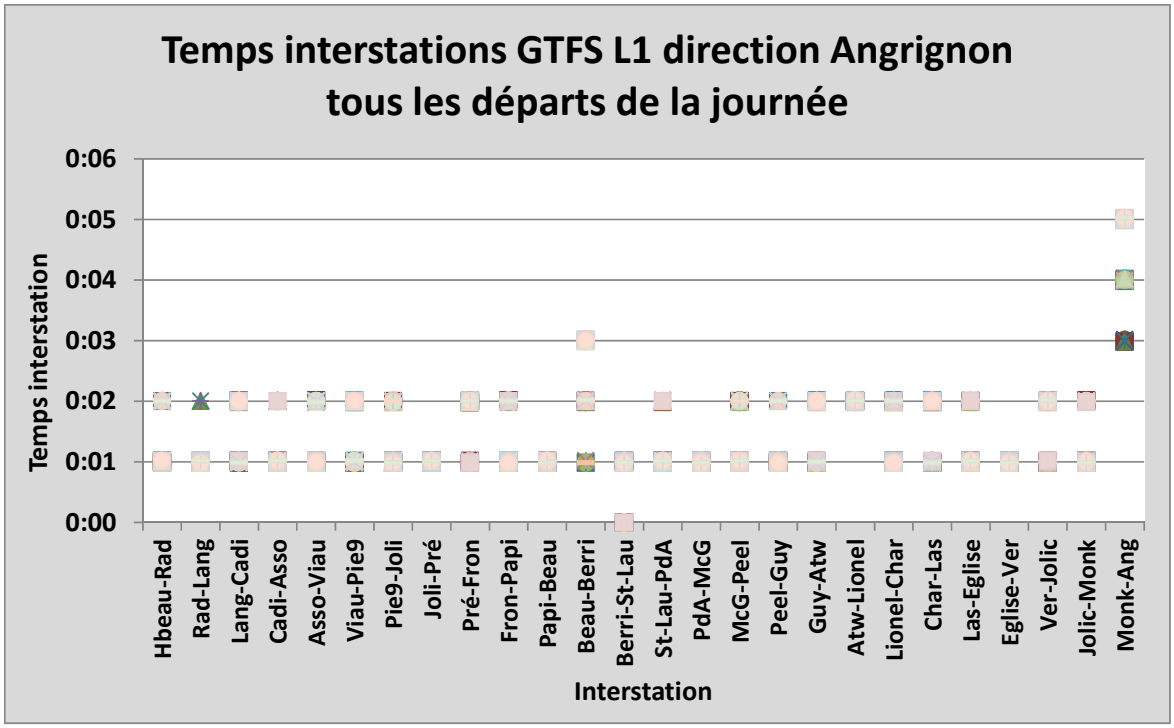


Figure 3.5 - Temps interstations GTFS ligne verte direction Angrignon

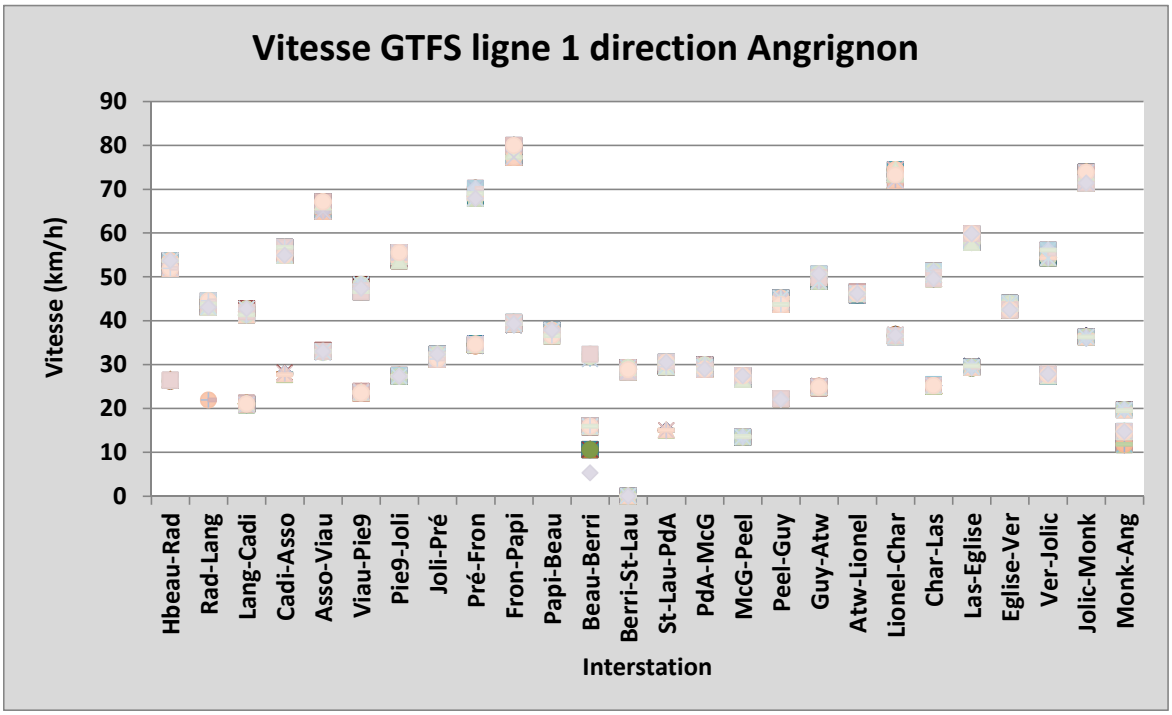


Figure 3.6 - Vitesses GTFS ligne verte direction Angrignon

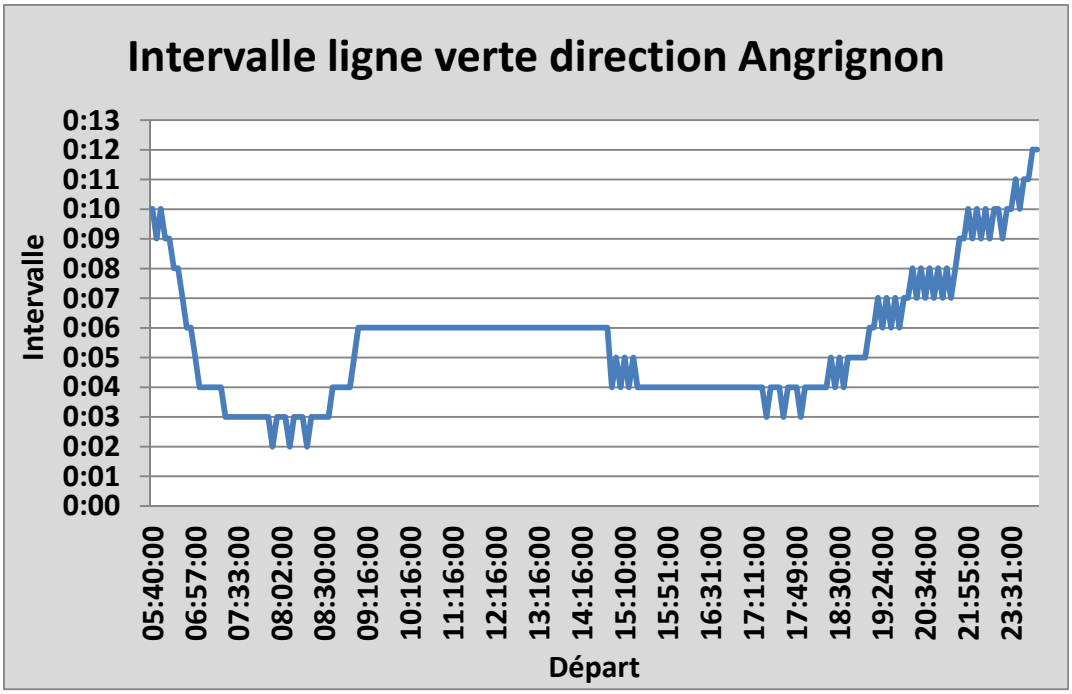


Figure 3.7 - Intervalle ligne verte direction Angrignon

3.1.6 Coûts d'exploitation

La STM fournit son budget annuel sous la forme d'un document *PDF* disponible publiquement sur son site web. Les calculs des coûts d'exploitation du réseau de métro par véhicule*km et par véhicule*heure sont basés sur les chiffres fournis dans ce document. Le budget total de la STM pour 2012 est de 1, 227 G$ (STM, 2012a). Les dépenses d'exploitation liées au métro sont de l'ordre de 25% du budget d'exploitation de 966 M$, soit 243 M$.

Pour l'évaluation des coûts, on utilisera les données réelles de l'année 2010, contenue dans le bilan 2012. En effet, les données nécessaires aux différents calculs (budget réel et distance totale parcourue par les véhicules) y sont présentées. Ainsi, pour 2010, le budget d'exploitation du réseau de métro se chiffre à 215 M$ (STM, 2012a). De plus, on estime à 76,6 millions de km la distance totale parcourue par les 756 voitures de métro, ce qui correspond à environ 100 000 km/voiture par année (STM, 2010). Les coûts d'opération du métro en voiture*km sont donc, pour 2010, de 215 M$/(756 voitures * 100 000 km/voiture) = 2,80 $/km. Cela signifie qu'il en

coûte 2,80$ à la STM pour faire circuler 1 voiture sur 1 km. Par ailleurs, en posant l'hypothèse que la vitesse commerciale moyenne des véhicules est de 35 km/h, on arrive à un total de (2,80 $/voit.*km * 35 km/h) ≈ 100$/voit.*heure : il en coûte 100$ à la STM pour opérer 1 voiture de métro pendant 1 heure.

La moyenne du temps en service par rame de métro est d'environ 6h, et le nombre de départ moyen d'environ 13 selon les données GTFS. Cependant, selon ces mêmes données, on retrouverait 108 rames desservant les lignes 1, 2 et 4, et 12 rames desservant la ligne 5. Ainsi [(108*9) + (12*6)]=1044 véhicules, et non pas 756. Il faut donc considérer que certains véhicules sont réutilisés sur différentes rames, et que la moyenne de 6h par rame s'applique au total de 1044 véhicules.

En juillet 2011, la Presse révélait que les coûts d'entretien du métro de Montréal s'établissaient à 1,25$/voit.*km parcouru pour l'année 2010 (Bisson, 2011). C'est donc l'équivalent de 44% des coûts d'exploitation du métro qui ont été dédiés à l'entretien.

3.1.7 Interruptions de service

Le service de métro a été interrompu à 980 reprises, causant un retard d'au moins 5 minutes à plus de 5,9 M de déplacements en 2011 (STM, 2011c). Les incidents causant ces interruptions ont duré en moyenne 11 minutes. (Bisson, 2012).

3.2 Expérimentations sur le métro

3.2.1 Les pannes de métro et ses clients victimes (Chapleau, 1997)

Cette publication, présentée au congrès annuel de l'AQTR, expliquait comment évaluer la clientèle affectée par une panne de service éventuelle. Les analyses tirées de cette publication sont un prélude à l'analyse des transactions de cartes à puce. En effet, les mêmes procédés

peuvent être utilisés dans ce mémoire, à l'exception des traitements appliqués aux données. Dans la publication de 1997, on se servait des données d'enquête pour conduire les analyses.

3.2.2 Flux dans le métro de Montréal : concepts et mesures (Piché & Chapleau, 2007)

Présentée au 75e congrès de l'ACFAS, cette expérimentation a permis de décrire la dynamique des déplacements faits en métro à Montréal. Des profils de charges par ligne et par station ont été dressés, et des matrices origine-destination ont été développées, encore une fois à partir des données d'enquête O-D. Les concepts sur lesquels se base cette expérimentation sont pratiquement les mêmes que ceux qui seront présentés au chapitre 4 de ce projet. Encore une fois, la différence majeure entre ce projet et la publication de 2007, est la source de données utilisée : les données d'enquêtes O-D vs les données de cartes à puce.

3.2.3 Modélisation espace-temps du métro de Montréal (Blais, 2007)

Ce mémoire de maîtrise, réalisé à l'école Polytechnique de Montréal, présente une méthode de modélisation espace-temps des rames de métro selon des données de cantons. Les résultats obtenus avec la méthode développée décrivent précisément les temps de parcours des rames. C'est pourquoi ils seront réutilisés dans ce projet, dans le but de raffiner les données de service planifié.

CHAPITRE 4 GESTION DES DONNÉES

> Les changements aux paradigmes qui s'exercent actuellement en planification du transport en commun proviennent du fait que l'avancement des systèmes automatiques de collecte de données et leur adoption par les opérateurs génèrent une abondance de données dans un environnement où la collecte de données était auparavant limitée.
>
> *Chu, 2010*

Ainsi, la gestion des données est maintenant une étape incontournable du processus de planification du transport en commun. De plus, la capitalisation sur les technologies de transport est cruciale pour l'industrie du transport en commun : les sociétés qui implémentent ces technologies avec succès voient s'étendre à la fois les champs d'applications des données, et l'ensemble d'individus à qui celles-ci profitent (Center for Urban Transportation Research, 2011).

Ce chapitre regroupe la démarche de caractérisation d'un ensemble de données et l'ensemble des manipulations qui lui seront appliquées. D'abord, les étapes d'acquisition, de stockage et de description d'un ensemble de données de transactions seront explorées. Ensuite, la programmation de procédures de traitement et d'enrichissement sera présentée en parallèle avec les hypothèses sur lesquelles elles reposent. Enfin, différentes applications possibles de l'utilisation des données traitées et enrichies seront illustrées.

4.1 Caractérisation de l'ensemble de données de transactions

4.1.1 Acquisition des données

Le système de perception par cartes à puce permet de recueillir les données de transactions qui se font quotidiennement sur le réseau de transport en commun. L'usager se procure sa carte à un point de vente, et peut la recharger dans les machines prévues à cet effet à chaque station. Lors de tout déplacement en métro, il doit valider sa carte au tourniquet. Se faisant, l'information contenue sur le titre de transport est transmise vers un serveur de stockage. Les données de transactions sont enregistrées aux tourniquets entre 5 h et 2 h, sauf pour certaines stations, où les fichiers sont enregistrés sur une période de 24 h. Les ensembles de données générés

quotidiennement à partir des transactions effectuées dans le métro seulement contiennent environ 900 000 enregistrements. La structure d'une base de données de transactions de validation est illustrée à la Figure 4.1.

Id support	Date	Heure	Réseau	Édicule	Tourniquet	Activité	Refus	Type de tarif
Carte à puce	03/17/2010	08:29:37	2 - STM	10534	32TEM0534MZ0201N101	BJc	0	176/CAMOP
Carte à puce	03/17/2010	15:44:50	2 - STM	10145	32TEM0146MZ0701O106	BJ1	0	176/CAMOP
Ticket	03/17/2010	10:50:08	1 - RC	10288	32PVM0288MZ0251N101	BJ1	0	62/3
Ticket	03/17/2010	19:49:19	2 - STM	10264	32TEM0264PS0201N103	BM	2	856/660O1

Figure 4.1 – Structure d'une base de données de transactions de validation

4.1.2 Objets de la base de données

Les données disponibles permettent de reconstituer le processus de transaction tel qu'illustré à la Figure 4.2. Pour une journée donnée, un usager qui désire se déplacer en métro doit détenir un titre de transport qui correspond à son âge, son occupation et son lieu de résidence. Si son entrée dans le métro est sa première interaction avec le réseau de transport en commun dans un intervalle de 2h, alors il s'agit d'un embarquement. Par contre, si l'usager se rend à la station de métro en utilisant d'abord le bus, son entrée dans le métro est une correspondance. L'usager entre dans le réseau de métro par l'édicule qui lui convient, et valide son titre de transport au tourniquet, pour ensuite accéder au quai de son choix.

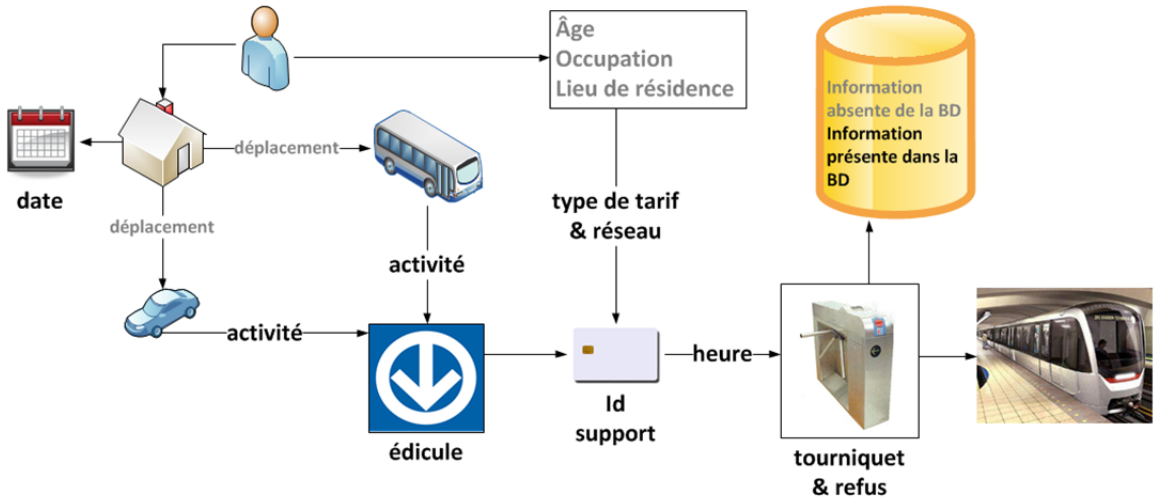


Figure 4.2 - Processus de transaction de validation pour un déplacement métro

Une fois le processus de transaction de validation reconstruit, tous les objets présents dans la base de données peuvent être comptés tel que présenté au Tableau 4.1. Se faisant, les analyses qui seront conduites dans les sections suivantes pourront être validées à partir des résultats obtenus ici. Par exemple, on peut imaginer un lien entre le nombre de tourniquets à une station et la quantité de transactions qui y sont effectuées quotidiennement.

Tableau 4.1 - Décompte des différents objets de la base de données

Nature des données	**Objet de la base de données**	**Cas uniques**
Directes	Id support	421 666
	Réseau	10
	Édicule	88
	Tourniquet	456
	Codes refus	11
	Type de tarif	145
Dérivées	Lignes	4
	Stations	68
	Types de « Id support »	3

Évidemment, les cas uniques associés à certains objets présentés ci-dessus peuvent varier d'une journée de transactions de validation à une autre. L'occurrence des objets dépend de l'utilisation spatio-temporelle du métro, qui elle, peut varier selon les conditions climatiques ou la tenue d'un événement spécial, entre autres. D'autres cas uniques peuvent correspondre à des erreurs dans l'enregistrement des données par l'équipement de perception. Les cas problématiques ou exceptionnels seront traités dans les sections suivantes, afin d'éviter toute confusion dans les analyses.

Les objets peuvent aussi être mis en relation entre eux, afin de pouvoir expliquer les valeurs uniques que chaque objet peut prendre. Par exemple, les tourniquets sont fixes et doivent donc être associés à une seule station. Dans le cas des codes de refus, l'analyse requiert une bonne connaissance des contraintes d'utilisation du réseau. En effet, une validation peut être refusée au tourniquet pour plusieurs raisons, dont la date d'expiration du titre et les réseaux compatibles avec le titre. Pour expliquer les différents codes de refus, il est nécessaire de les croiser avec d'autres objets de la base de données.

4.1.3 Stockage

La base de données de transactions pour une journée correspond à l'ensemble des séries de transactions enregistrées à chaque tourniquet. Les séries comportent en moyenne 346 enregistrements, et ont une durée maximale de 24 heures. La Figure 4.3 présente la durée et le nombre d'enregistrements de chaque série. On observe que les séries ont des nombres d'enregistrements variant de 1 à 422, avec une concentration plus forte près de 400. La description de l'achalandage à un tourniquet est donc contenue dans une ou plusieurs séries, selon l'utilisation du tourniquet. Plus un tourniquet est utilisé, plus il comportera de séries de données. Pour la journée du 17 mars 2010, pour l'ensemble des 456 tourniquets et postes de changeur des 68 stations de métro, on retrouve 2 355 séries de données.

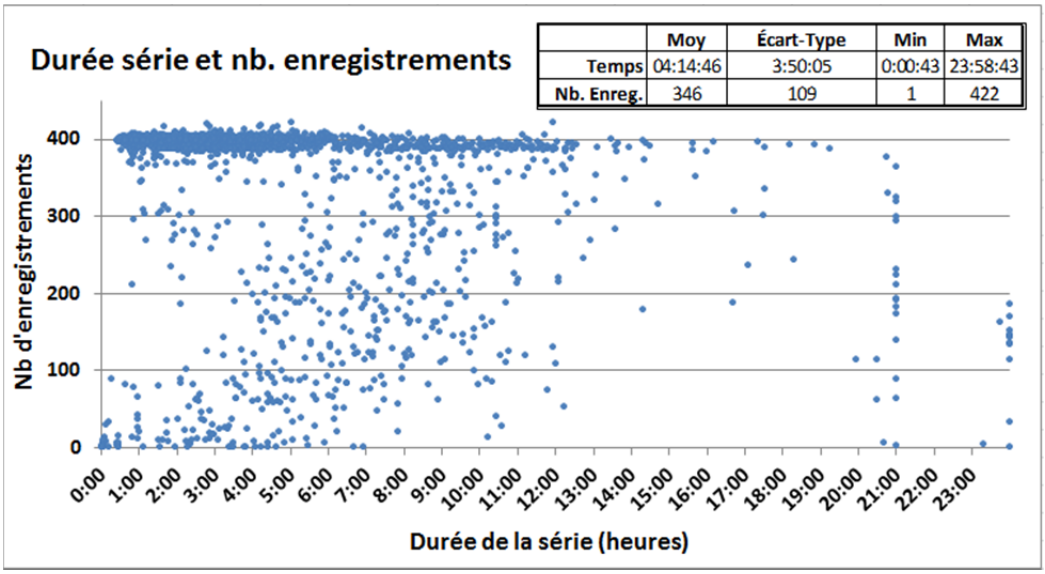


Figure 4.3 - Durée et nombre d'enregistrements par série

4.1.4 Cohérence des valeurs observées et cas particuliers

Une fois que les données ont été mises en relation avec des objets du réseau, il est possible de vérifier la cohérence des valeurs présentes dans la base de données.

4.1.4.1 Tourniquets

En plus de fournir le tourniquet utilisé lors de la validation, la base de données précise le numéro de série de l'équipement de perception contenu dans le tourniquet. En groupant les tourniquets aux numéros de séries de l'équipement de perception, il arrive certains cas où un tourniquet est associé à 2 numéros de séries différents comme à la Figure 4.4. Cette donnée permet de supposer qu'un nouvel équipement de perception pourrait avoir été testé dans ce tourniquet, et que les transactions qui y sont associés ne sont peut-être pas des déplacements.

Count of ligne		
ceb	id_transac	Total
32TEM0534MZ0201N101	_2WK0K7M6R	9310
	_2WK11RGEW	2
32TEM0534MZ0201N102	_2WK0K9OJX	6166
32TEM0534MZ0201N103	_2WK0KB85F	4078

Figure 4.4 – Exemple de valeur suspecte

4.1.4.2 Utilisation multiple d'un même titre

Par ailleurs, il est possible, en observant les valeurs maximales du nombre de transactions par numéro de carte, de voir jusqu'à quelle fréquence un même carte peut être validée dans la même journée. Il suffit de trier les données par numéro de support et par heure, puis d'observer les numéros de support dont les nombres de transactions sont les plus élevés. Le Tableau 4.2 présente le cas d'un même support ayant été utilisé pour compléter 25 transactions dans la même journée, nombre très élevé comparé à la moyenne du nombre de transactions par support, qui est d'environ 1.9.

Tableau 4.2 - Support utilisé pour 25 transactions entre 5:48 et 15:55 le 17 mars 2010

activite	reseau	produit	test	refus	date28	heure28
BJ1	2 - STM	176/CAMOP	Non	0	17 Mar 10	05.48.52
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	06.10.52
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	06.19.00
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	06.52.44
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	07.09.16
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	07.23.07
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	07.30.13
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	07.44.23
BJ1	2 - STM	176/CAMOP	Non	0	17 Mar 10	07.51.53
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	08.44.11
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	08.53.44
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	09.06.35
BM	2 - STM	176/CAMOP	Non	29	17 Mar 10	09.11.11
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	09.21.43
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	09.29.05
BJ1	2 - STM	176/CAMOP	Non	0	17 Mar 10	10.15.35
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	10.26.00
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	10.33.52
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	11.04.34
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	11.16.47
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	11.51.03
BJ1	2 - STM	176/CAMOP	Non	0	17 Mar 10	12.55.39
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	13.45.12
BJ1	2 - STM	176/CAMOP	Non	0	17 Mar 10	15.30.20
BJc	2 - STM	176/CAMOP	Non	0	17 Mar 10	15.55.44

Évidemment, si un nombre de transactions semble anormalement élevé pour un numéro de support, il pourrait s'agir de la carte d'un employé, d'un policier, d'un agent de sécurité, ou d'un musicien du métro. Toutefois, il ne faut pas écarter la possibilité qu'il s'agisse d'une carte frauduleuse.

4.1.4.3 Codes refus

Ensuite, on dénombre un total de 10 codes refus en excluant le code 0, qui correspond à une transaction acceptée au tourniquet. En observant la station à laquelle le code refus se répète le plus souvent, le type de montée, le produit associé à la transaction, et les caractéristiques des transactions précédant et suivant le refus, on peut arriver à déterminer les conditions qui génèrent un code refus particulier. Par exemple, le code refus 10 se produit exclusivement aux stations Cartier, de la Concorde et Montmorency. La transaction suivant une transaction refusée avec le code 10 est toujours acceptée dans une station située sur l'île de Montréal. Ainsi, il est possible de conclure que le code refus 10 correspond à l'erreur générée lorsqu'un usager qui possède un titre mensuel acheté sur l'île se rend à l'extérieur de l'île, et tente de revenir à Montréal avec le même titre. Un abonnement mensuel provenant de l'extérieur de l'île ne coûte pas le même prix qu'un abonnement mensuel acheté sur l'île. Les utilisateurs avec un titre mensuel acheté sur l'île se voient donc obligés d'acheter un ticket pour revenir à Montréal lorsqu'ils se rendent à Laval. D'ailleurs, puisque cette transaction est habituellement complétée avec un ticket, le numéro de support est différent et le déplacement n'apparaît pas dans la chaîne de déplacement d'un numéro de support.

Un autre aspect à vérifier est, entre autres, le nombre de titres de transport contenus sur un même support. Un usager peut accumuler jusqu'à 4 titres de transport différents sur sa carte (STM, 2011b). Il est donc approprié de vérifier si cette condition est respectée dans tous les cas.

La connaissance des ensembles de données permet d'apprécier la qualité des analyses qui pourront ensuite être conduites. Autrement dit, il s'agit d'un moyen d'évaluer la rentabilité du système de perception par cartes à puce.

4.2 Procédures de traitement et d'enrichissement

Les bases de données de transactions générées quotidiennement dans le métro de Montréal contiennent environ 900 000 enregistrements. Ainsi, les outils de traitement et d'enrichissement

doivent permettre la manipulation de tels ensembles de données. L'efficacité d'Excel est limitée pour les très gros ensembles de données. C'est pourquoi les traitements seront appliqués sur les données avec le logiciel Stata. Stata est un outil de statistique qui possède une interface graphique et une ligne de commande, et qui permet d'automatiser les traitements à appliquer à de grands ensembles de données. Excel n'est cependant pas écarté pour les analyses : en effet, le logiciel pourra être utilisé pour traiter et visualiser des sous-ensembles de données.

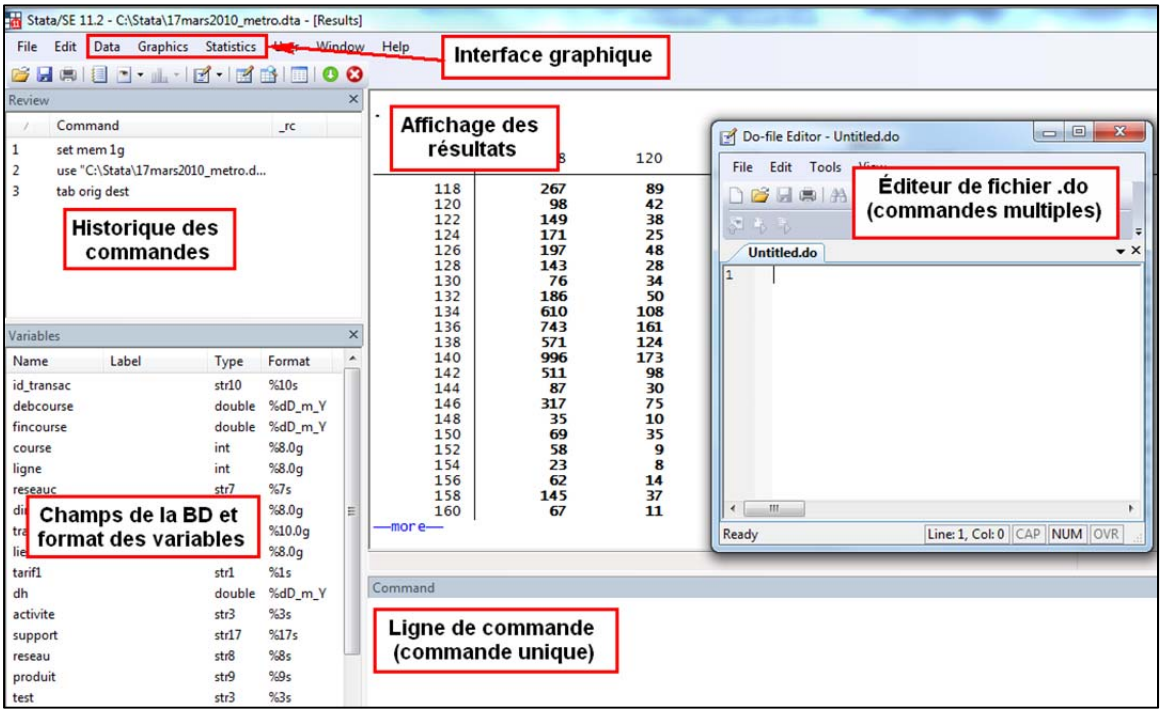


Figure 4.5 – Interface du logiciel Stata

4.2.1 Procédures de traitement

Les procédures de traitements permettent de corriger et d'uniformiser les données. En effet, on peut retrouver des enregistrements doubles, des formats de valeurs particuliers, ou des valeurs qui ne seront pas utiles aux analyses. Il faut donc que les problèmes soient identifiés, et que les données passent à travers une série de procédures de traitements programmée d'avance pour la correction des données.

D'abord, les fichiers de base de données, en format .dbf à l'origine, doivent être convertis en format .dta pour être traités dans Stata. Une fois la conversion complétée avec le logiciel Stat/Transfer (environ 3 minutes avec un processeur à quatre cœurs, 2.53 GHz, 6 Go Ram, sous Windows 7 64 bits, pour un fichier d'environ 238 Mo), la base de données peut être importée dans Stata, et les manipulations peuvent être appliquées sur les données.

Stata permet l'automatisation de procédures grâce à son interface de programmation. Le langage utilisé est propre à Stata (extension .do), et relativement simple à implémenter. Ce qui caractérise la programmation .do est le fait que toute commande entrée s'applique toujours à l'ensemble des données en entier, à moins de spécifier des exceptions. Les procédures suivantes s'avèrent utiles pour le traitement des données de transactions. La programmation complète de ces procédures se trouve en annexe.

4.2.1.1 Enregistrements doubles (Annexe A1)

La première étape du traitement consiste à vérifier si des enregistrements sont présents en double, triple, etc. On observe rapidement cette anomalie en triant les données selon le # de support, l'heure et le tourniquet auquel est effectué la transaction. Les doubles sont ensuite supprimés pour obtenir un ensemble de transactions uniques.

4.2.1.2 Types d'activités inutilisés (Annexe A1)

Les transactions pour lesquelles les valeurs du champ *activité* sont *BF* ou *BG* correspondent à des « dummy » qui identifient le début d'une nouvelle série de transactions (debcourse/fincourse), et ne sont pas effectuées par des usagers. Celles-ci peuvent donc être supprimées.

4.2.1.3 Correction de champ (Annexe A1)

Le champ *heure28* correspond à l'heure de transaction en format hh**.**mm**.**ss. Toutefois, les valeurs de ce champ prennent le format hh**:**mm**:**ss pour certaines transactions effectuées après minuit, ce

qui peut s'avérer problématique lors du tri des données par heure. Puisque cette valeur est cruciale pour la reconstitution de la séquence de transactions d'un usager, les valeurs doivent nécessairement être corrigées. Cette procédure permettra d'uniformiser le format des valeurs de ce champ.

4.2.1.4 Refus consécutifs (Annexe A2)

Puis, dans le cas où l'on observe deux refus consécutifs ou plus à la même station avec le même numéro de support, les enregistrements répétés peuvent être supprimés. En effet, lorsqu'un usager se voit refuser l'accès lors de la validation au tourniquet, il peut tenter de revalider sa carte à répétition au même tourniquet. Ces transactions refusées sont toutes enregistrées, et font en sorte qu'on pourrait surestimer la quantité de transactions refusées. Ainsi, puisque les refus consécutifs peuvent fausser les estimations, ils doivent être supprimés.

4.2.2 Procédures d'enrichissement

Par procédures d'enrichissements, on entend toutes les procédures de calculs ou de dérivation d'information supplémentaire à partir des données existantes. Contrairement aux procédures de traitement, qui servent à corriger l'information existante, les procédures d'enrichissement permettent la production de nouvelles données. Celles-ci découlent de calculs, de regroupement de champs, et du respect de certaines hypothèses.

4.2.2.1 Nombre de transactions par usager (Annexe A3)

Le nombre de transactions par usager est une statistique clé tant au niveau du traitement que de l'analyse des données. En effet, en se basant sur ce champ, il sera possible de dériver les stations de destinations, mais aussi de grouper les supports selon leur fréquence d'utilisation dans le métro. Le nombre de transactions peut être identifié en triant les données par numéro de support et par heure de transaction.

4.2.2.2 Stations de destination (Annexe A4)

À partir de la station d'origine d'un déplacement, il est possible d'inférer sur la destination selon certaines hypothèses. Ainsi, on suppose un déplacement symétrique pour tous les numéros de support pour lesquels on retrouve 2 transactions, et un retour à l'origine pour les numéros ayant conclu 3 transactions ou plus. Pour un numéro de carte ayant conclu exactement 2 transactions, la station d'origine de la 1re transaction correspond à la station de destination de la 2e transaction, et la station d'origine de la 2e transaction correspond à la station de destination de la 1re transaction. Dans le même ordre d'idées, si un même numéro de support conclut 3 transactions ou plus, alors la station d'origine de la 1re transaction correspond à la station de destination de la dernière transaction, et la station de destination de la transaction « n » correspond à la station d'origine de la transaction « n+1 ». La codification de cette procédure dans Stata se retrouve en annexe.

Cette méthodologie est proposée tenant compte du fait qu'une seule journée de transactions était disponible pour les analyses. En principe, l'information sur la station de destination devrait être dérivée à partir de deux ou plusieurs journées successives de transactions, pour limiter les hypothèses non vérifiées. Des données sur une plus longue période permettraient aussi d'identifier des groupes d'usagers, par exemple les travailleurs de 9 h à 17 h, et des points d'ancrage qui sont en fait les stations habituelles des différents usagers.

Une fois épuré, l'ensemble de données contient 797 975 enregistrements valides, dont 17% sont des transactions uniques, soit des numéros de support qui n'apparaissent qu'une seule fois dans la base de données.

4.2.2.2.1 Destination des supports à transaction unique

La base de données étudiée comporte environ 140 000 transactions uniques, c'est-à-dire des supports pour lesquels seulement 1 validation est faite au courant de la journée. Cela peut se produire lorsqu'une personne utilise plus d'un mode pour se déplacer, lorsqu'un titre mensuel est refusé à une station située à l'extérieur de l'île de Montréal et où l'usager a dû se procurer un ticket, ou bien qu'un usager ne soit pas rentré à son lieu habituel de domicile à la fin de la

journée. Il peut aussi s'agir de transactions effectuées par des utilisateurs occasionnels, ou par des touristes, qui eux, peuvent avoir des patrons de déplacements irréguliers.

Les transactions uniques seront laissées de côté pour les analyses et la modélisation du métro. En effet, puisque la station de destination ne peut être dérivée selon la méthode présentée à la section 4.2.2.2, la modélisation de ces déplacements n'est pas possible dans leur forme actuelle. De plus, puisque le numéro de support de ces transactions n'est présent qu'une seule fois dans la base de données, la durée des activités ne peut pas être calculée.

Ces transactions ne doivent cependant pas être ignorées. Au contraire, elles doivent être caractérisées afin de connaître l'impact potentiel de leur absence dans les analyses et la modélisation du métro.

La distribution horaire de ces transactions uniques est présentée à la Figure 4.6. La distribution des transactions uniques est pratiquement identique à la distribution de l'ensemble des transactions, comme on le verra à la section 6.4. De ce fait, on est assuré que les résultats de la modélisation du métro, qui seront présentés subséquemment, ne sont pas biaisés par l'absence de ces transactions.

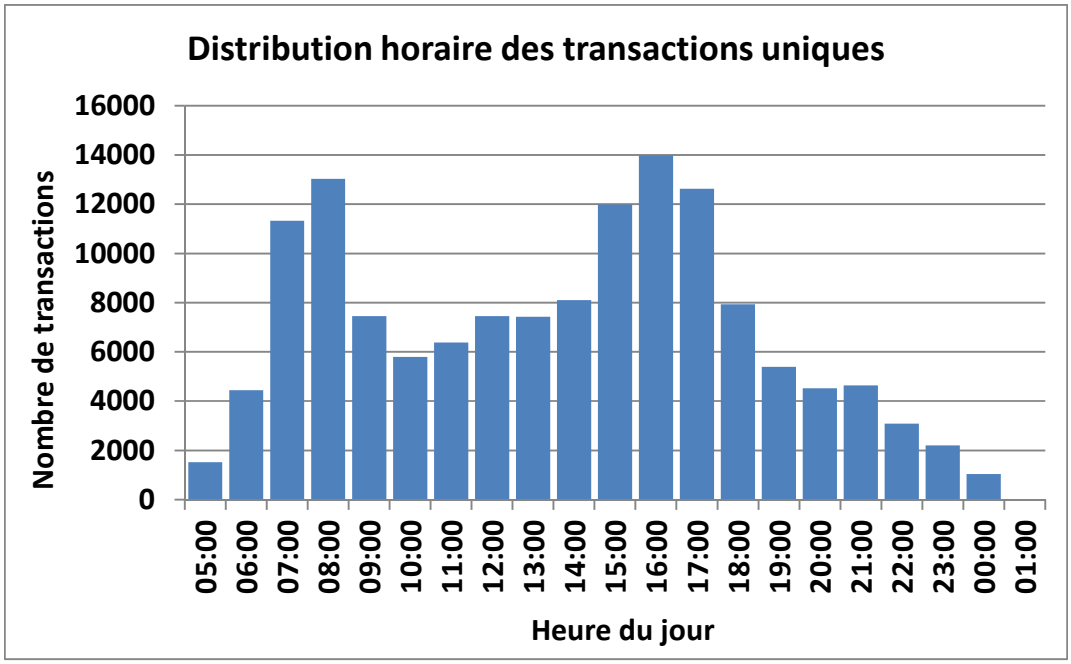


Figure 4.6 - Distribution horaire des transactions uniques

Le Tableau 4.3 présente les 5 types de produit les plus fréquents, issus du réseau de la STM, pour les validations uniques. En examinant le tableau, on réalise que les transactions uniques peuvent provenir de titres mensuels, hebdomadaires ou quotidiens. Cela signifie que ce ne sont pas seulement les utilisateurs occasionnels du métro qui font une seule transaction dans la journée, et que d'autres événements peuvent avoir une influence sur le profil de déplacements des usagers. Encore une fois, une base de données qui s'étend sur plusieurs jours permettrait d'éliminer certains cas d'exception. En effet, il peut survenir une journée pour laquelle un usager emprunte un itinéraire inhabituel pour différentes raisons, et les données de la journée suivante permettraient de confirmer les lieux habituels d'embarquement et de débarquement de celui-ci.

Tableau 4.3 - Types de produits les plus fréquents lors de transactions uniques

Produit	Type	Fréquence	%
857/T10O1	10 passages tarif ordinaire	36 860	26
176/CAMOP	Abonnement mensuel tarif ordinaire	25 006	18
856/660O1	6 passages tarif ordinaire	18 081	13
177/CAMRP	Abonnement mensuel tarif réduit	15 258	11

Tableau 4.3 - Types de produits les plus fréquents lors de transactions uniques (suite)

456/TDXOP	2 passages tarif ordinaire	10 584	8
Autres		34 589	25
TOTAL		**140 378**	**100**

Enfin, le Tableau 4.4 identifie les stations auxquelles ont lieu le plus de validations uniques, et les comparent aux stations auxquelles il y a le plus de transactions, toutes transactions confondues. Cette comparaison permet de voir si certaines stations possèdent un nombre anormalement élevé de validations uniques, comme il semble que ce soit le cas pour la station Montmorency. En effet, cette station occupe le 5e rang des stations avec le plus de validations uniques, et occupe le 25e rang du nombre total de transactions par station. Ce phénomène est dû au fait que la tarification à Laval fait en sorte qu'un usager qui possède un abonnement mensuel à Montréal peut se rendre à Laval, mais ne peut en revenir avec le même titre (tel qu'expliqué à la section 4.1.4.3). L'usager doit se procurer un ticket, et son déplacement ne sera pas associé à son numéro de carte. Il n'est donc pas possible de modéliser ce déplacement, puisqu'une partie de celui-ci est associée à un numéro de support différent.

Tableau 4.4 - Rang des stations selon le nombre de validations uniques et le nombre total de transactions

#	*Station*	*Rang transac. uniques*	*Rang total transac.*
146	**Berri-UQAM**	1ère	1ère
140	**McGill**	2^{e}	2^{e}
136	**Guy-Concordia**	3^{e}	3^{e}
454	**Longueuil-Université Sherbrooke**	4^{e}	4^{e}
288	**Montmorency**	5^{e}	25^{e}
252	**Bonaventure**	6^{e}	5^{e}
222	**Côte-Vertu**	7^{e}	6^{e}
242	**Vendôme**	8^{e}	9^{e}
280	**Henri-Bourassa**	9^{e}	10^{e}
170	**Honoré-Beaugrand**	10^{e}	13^{e}

4.2.3 Données GTFS : précision à la seconde

Les données GTFS, telles que présentées à la section 3.1.5.1, fournissent l'horaire de service planifié à la minute près. Elles permettent de connaître, pour une journée de service, la position de tous les véhicules dans le temps.

Par ailleurs, la description de l'achalandage au niveau du véhicule nécessite un format de données qui permet d'associer des transactions à des départs spécifiques. Autrement dit, on doit pouvoir connaître la rame dans laquelle un usager embarque, en fonction de l'heure à laquelle il a franchi le tourniquet. Ainsi, il serait préférable d'utiliser les données de service planifié dans leur précision la plus fine, soit à la seconde.

Quelques principes simples sont établis pour raffiner la précision des données de service planifié. D'abord, certaines stations dites « stations de contrôle » ont des heures de passage que l'on considère comme étant toujours respectées. Ces stations sont les terminus de départ et d'arrivée, et les stations de correspondance. Le temps entre les passages à deux stations de contrôle successives doit être distribué proportionnellement selon la longueur des interstations contenues entre les deux stations de contrôle. De plus, l'achalandage moyen à une station dicte le temps pendant lequel le véhicule reste arrêté en station. De façon générale, les temps d'attentes aux stations sont tous d'environ 15 à 20 secondes, sauf aux stations de contrôle, où ces temps sont d'environ 45 secondes. Une source de données qui exprime la position réelle des véhicules dans le temps sera utilisée pour tenir compte de ces différentes contraintes.

En effet, pour passer de la minute à la seconde, deux sources de données seront croisées. D'une part, les données GTFS du service planifié pour un jour moyen de semaine de la période janvier 2010 seront utilisées comme structure de base. D'autre part, les résultats de la modélisation spatio-temporelle du métro à partir des données de cantons (Blais, 2007) permettront d'enrichir les données GTFS. Les données de cantons expriment l'activation et la désactivation dans le

temps des cantons, ou bornes électrifiées, au moment du passage d'un véhicule sur celles-ci. Dans la recherche de Blais (2007), la position des véhicules est connue sur l'ensemble du réseau (sauf la ligne jaune et les stations de la ligne orange situées à Laval) à la seconde près, pour la période de pointe du matin. Il est donc possible, en utilisant un temps de départ précis, de modéliser le déplacement d'un véhicule, et de déterminer ses temps de passage aux stations à la seconde près.

La Figure 4.7 présente les relations entre les mesures et les données en présence lors du raffinement des données GTFS.

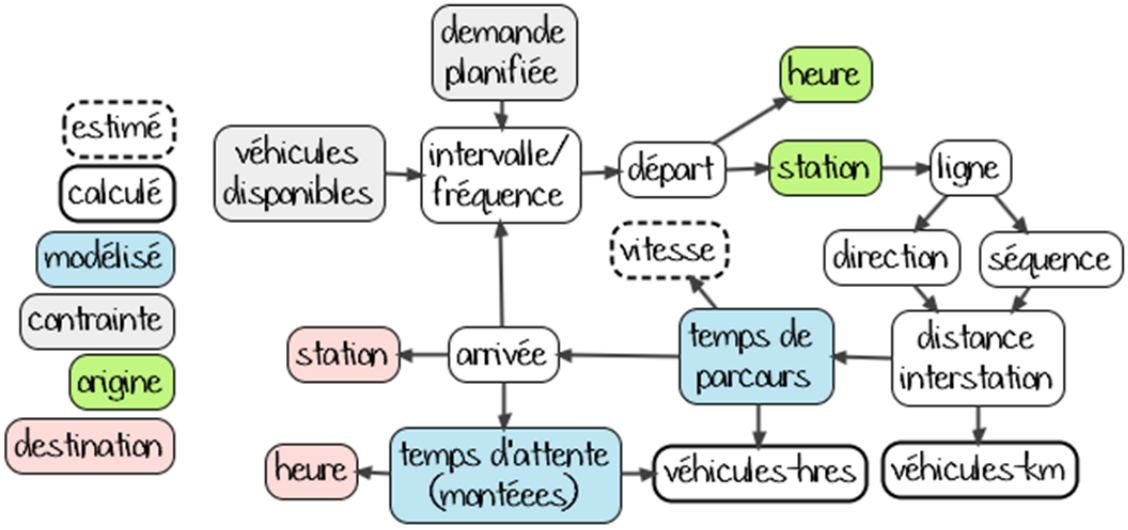


Figure 4.7 Structure des données GTFS

4.2.3.1 Méthode

Le raffinement des données GTFS passe par :

1. La détermination des temps de parcours interstations et des temps passés en station moyens;
2. La segmentation des lignes de métro en points de contrôle définis par les terminus et stations de correspondances;

3. Le calcul des coefficients de distribution du temps de parcours par interstation et du temps d'attente par station;
4. L'application des coefficients aux temps planifiés GTFS d'arrivée et de départ.

Les détails de la modélisation espace-temps du métro sont explicités dans le mémoire de Blais, 2007, et ne seront pas repris ici. Cependant, les résultats de la modélisation seront utilisés comme référence en ce qui concerne les temps de parcours et les temps d'attente en station moyens.

Au moment où la recherche sur la modélisation espace-temps du métro a été entreprise, la portion de la ligne 2 située à Laval était inexistante. C'est pourquoi les temps de parcours pour cette portion ont été estimés avec les données d'accélération et de vitesse moyenne des rames, aussi disponibles dans le mémoire de Blais (2007). La même procédure a été appliquée aux stations de la ligne 4, pour laquelle les données n'étaient pas disponibles.

Par ailleurs, il faut noter que les données de modélisation spatio-temporelle du projet de recherche de 2007 correspondent à la période de pointe du matin, entre 7 h 30 et 9 h. Les temps de parcours calculés avec les données GTFS présentent une différence moyenne d'environ 2 minutes entre la période de pointe et la période hors pointe, tel que présenté au Tableau 4.5. On posera donc l'hypothèse que ces différences de temps sont distribuées en respectant les proportions des coefficients calculés.

Tableau 4.5 - Temps de parcours par ligne pour les périodes de pointe et hors pointe selon les données GTFS

Ligne	Hors pointe (minutes)	Pointe (minutes)
L1	36	39
L2	45	49
L4	5	6
L5	14	15

4.2.3.1.1 Détermination des temps de parcours moyens entre chaque station

D'abord, la démarche consiste à déterminer les temps en station et interstations moyens. Ceux-ci sont pris directement des valeurs obtenues à la suite de la modélisation espace-temps du métro de Montréal (Blais, 2007). Le Tableau 4.6 présente les données GTFS comparées à celles obtenues par la modélisation espace-temps pour la ligne 5, en direction Snowdon.

Tableau 4.6 - GTFS vs Modélisation pour la ligne 5, direction Snowdon

	Distance (km)	GTFS		Modélisation			
		Temps	Vitesse (km/h)	Interstation	En station	Total	Vitesse (km/h)
St-Michel	-	**00:00**	-	00:00	00:00	**00:00**	**-**
Iberville	1.2	**01:00**	**72**	00:49	00:25	**01:14**	**88**
Fabre	0.65	**01:01**	**38**	00:52	00:26	**01:18**	**45**
Jean-Talon	0.82	**02:00**	**24.6**	00:58	00:48	**01:46**	**51**
Castelnau	0.7	**00:59**	**43**	00:47	00:23	**01:10**	**54**
Parc	0.6	**01:01**	**35**	00:42	00:24	**01:06**	**51**
Acadie	0.52	**02:00**	**15.6**	00:50	00:23	**01:13**	**37**
Outremont	0.93	**01:00**	**56**	01:00	00:25	**01:25**	**56**
Édouard-M.	0.88	**02:00**	**26**	01:18	00:24	**01:42**	**41**
U-de-M	0.78	**00:59**	**48**	00:52	00:26	**01:18**	**54**
CDN	1.2	**02:00**	**36**	00:57	00:24	**01:21**	**76**
Snowdon	0.7	**01:01**	**41**	01:02	00:00	**01:02**	**41**
TOTAL	**9**	**15:01**	-	10:07	04:28	**14:35**	**-**

La différence entre le temps de parcours pour la ligne 5 en direction Snowdon est en moyenne de 25 secondes entre les données GTFS et la modélisation spatio-temporelle à partir des données de cantons.

4.2.3.1.2 La segmentation des lignes de métro en points de contrôle

Ensuite, les coefficients de temps de parcours et de temps en station doivent être calculés entre les points de contrôle. Les points de contrôle ont été fixés aux terminus et aux stations de correspondance, tel qu'illustré à la Figure 4.8.

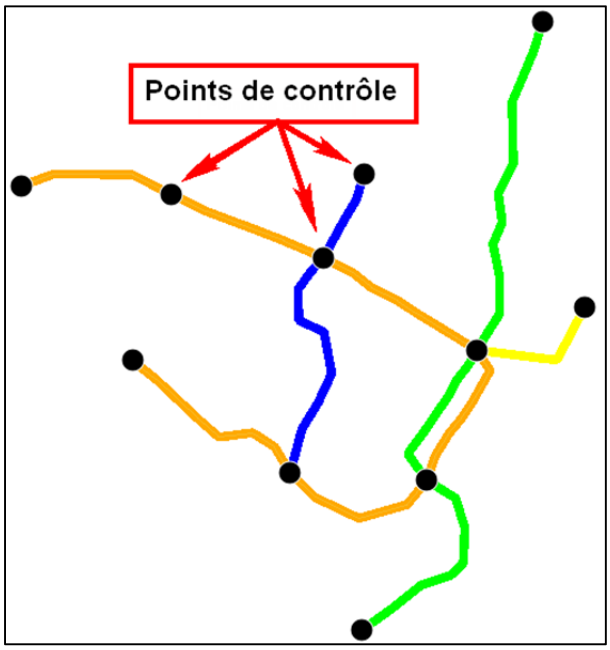


Figure 4.8 - Positionnement des points de contrôle

4.2.3.1.3 Le calcul des coefficients de distribution du temps de parcours par interstation

Un exemple de calcul des coefficients entre le départ et l'arrivée à 2 points de contrôles successifs est présenté à la Figure 4.9. En jaune, on voit les points de contrôle sélectionnés. Les formules utilisées pour calculer les coefficients de temps de parcours « P » et de temps d'attente « A » sont respectivement :

$$P_j = \frac{Temps\ interstation}{(Temps\ Cumulé_{Cj} - Temps\ Cumulé_{Cj-1})}, \text{ et}$$

$$A_j = \frac{Temps\ en\ station}{(Temps\ Cumulé_{Cj} - Temps\ Cumulé_{Cj-1})},$$

où C est un point de contrôle, Temps Cumulé est le temps cumulé d'un voyage lorsqu'il est rendu à une station spécifique, et j est un indice qui désigne le point de contrôle à utiliser.

	GTFS	Modélisation					
		Interstation	En station	Total	Total Cumulé	Coeff. P	Coeff. A
St-Michel	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00	0	0
Iberville	00:01:00	00:00:49	00:00:25	00:01:14	00:03:49	0.233	0.119
Fabre	00:01:01	00:00:52	00:00:26	00:01:18	00:02:06	0.248	0.124
Jean-Talon	00:02:00	00:00:58	00:00:48	00:01:46	00:03:30	0.276	
Castelnau	00:00:50	00:00:47	00:00:23	00:01:10	00:05:05		

Figure 4.9 - Exemple de calcul de coefficient

Entre le départ à Saint-Michel et l'arrivée à Jean-Talon, la somme des coefficients P et A donne toujours 1, car ce sont deux points de contrôle.

4.2.3.1.4 L'application des coefficients aux temps planifiés GTFS d'arrivée et de départ

Une fois que tous les coefficients ont été calculés pour chaque station et chaque direction, ceux-ci sont appliqués aux heures de passages GTFS. L'heure d'arrivée aux stations est prise directement des données GTFS, si la station est un point de contrôle : dans le cas contraire, l'heure d'arrivée est estimée. L'heure de départ, elle, est calculée partout sauf aux terminus, où elle est fournie par les données GTFS.

Ainsi, pour chaque heure de passage des 1581 voyages décrits dans les données GTFS, on doit appliquer une formule pour l'heure d'arrivée, et une autre pour l'heure de départ. Les formules sont :

$Ha_i = Hd_{(i-1)} + [P_i * (C_j - C_{j-1})]$ si « i » n'est pas un point de contrôle, et

$Hd_i = Ha_{(i)} + [A_i * (C_j - C_{j-1})]$ si « i » n'est pas un terminus,

où Ha = heure d'arrivée, Hd = heure de départ, P = coefficient temps de parcours, A = coefficient temps d'attente, C = heure GTFS au point de contrôle, i = station et j = point de contrôle.

En appliquant ces formules aux 30 254 enregistrements du fichier GTFS utilisé, on obtient les heures de passages estimées à la seconde près. Un échantillon des données modifiées est présenté au Tableau 4.7, où les champs « arrivée » et « départ » sont les résultats de l'application des formules précédentes sur les données GTFS originales.

Tableau 4.7 - Échantillon de données GTFS modifiées

Ligne	Véhicule	DepTerminus	arrival_time	departure_time	stop_id	stop_sequence	arrivée	départ
5	9	05:30:00	05:30:00	05:30:00	64	1	05:30:00	05:30:00
5	9	05:30:00	05:30:59	05:30:59	63	2	05:30:56	05:31:24
5	9	05:30:00	05:32:00	05:32:00	62	3	05:32:23	05:32:53
5	9	05:30:00	05:33:59	05:33:59	5	4	05:33:59	05:34:47

Tableau 4.8 - Provenance des nouvelles heures de départ et d'arrivée selon le type de station

	Terminus	**Point de contrôle**	**Station intermédiaire**	**Heure d'arrivée**	**Heure de départ**
Type de station	x	x		GTFS	GTFS
		x	x	GTFS	Estimée
			x	Estimée	Estimée

L'application de cette méthode respecte donc les heures planifiées de départ et d'arrivée aux terminus, et fournit une représentation plus réaliste de la position des véhicules dans le temps. Tout en modélisant les temps d'attente en station, les nouvelles données règlent les incohérences rencontrées dans les données GTFS originales. En effet, on retrouve différents cas où les rames

sont à 2 stations à la même heure dans les données GTFS. Un exemple de cette situation est fourni à la Figure 4.10.

C	D	E	F	G	H	I	J	K
Ligne	Véhicule	DepTerminus	arrival_time	departure_time	stop_id	stop_sequence	arrivée	départ
1	1	05:30:00	05:50:00	05:50:00	11	15	05:50:00	05:50:22
1	1	05:30:00	05:50:00	05:50:00	29	16	05:51:05	05:51:24

Figure 4.10 - Exemple de positionnement incohérent d'une rame de métro tiré des données GTFS originales et la correction apportée

Le résultat de l'application de la méthode ne donne pas les heures de passage telles que planifiées à la seconde près, mais s'en approche davantage que les données GTFS pures. En effet, en comparant la Figure 4.11, dressé à partir des données modifiées, à la Figure 3.6 qui contient les données GTFS originales, on observe une grande différence au niveau de l'écart entre les vitesses d'une même interstation. Cela est dû au fait que dans les données GTFS originales, une augmentation du temps de parcours d'une minute est nécessairement appliquée à 1 seule interstation. Pour les données GTFS modifiées, la même augmentation est appliquée à l'ensemble des stations contenues entre les deux points de contrôles qui bornent l'interstation où l'augmentation a eu lieu dans les données GTFS originales. De cette façon, le service planifié est beaucoup plus réaliste qu'en prenant les données GTFS directement, sans toutefois correspondre au service planifié réel.

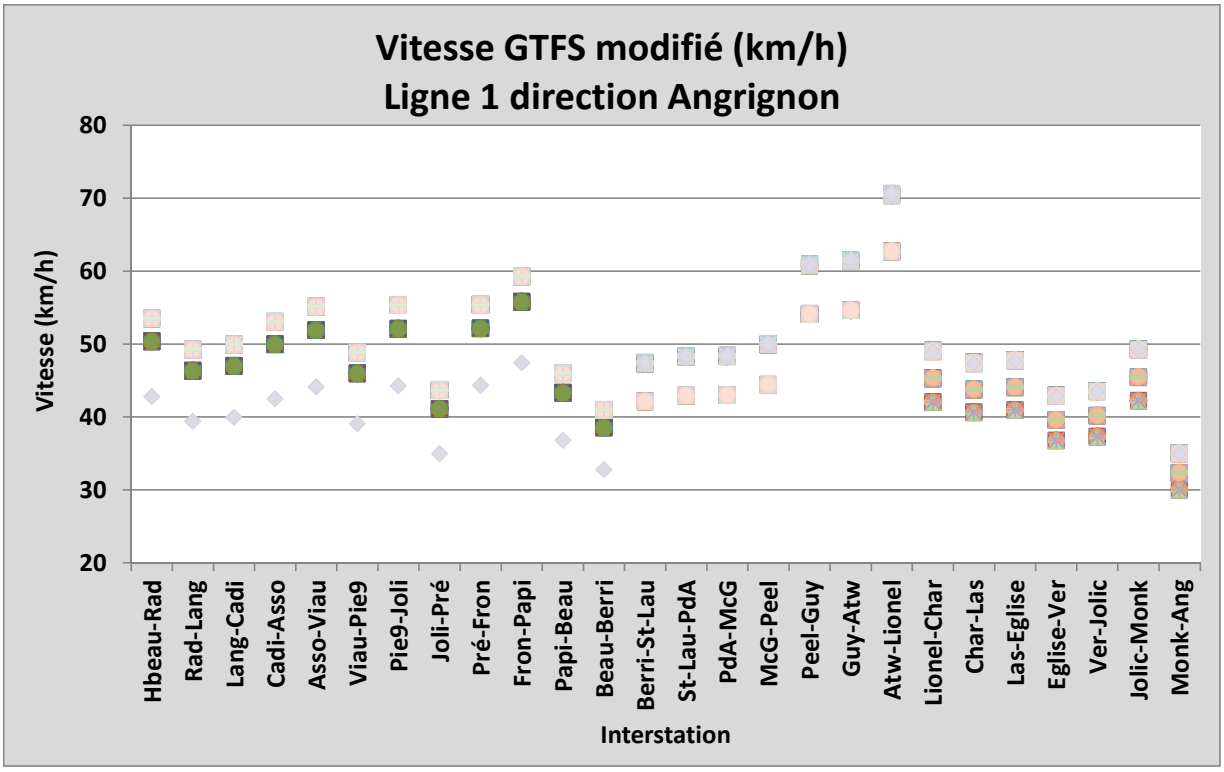


Figure 4.11 Vitesse GTFS modifié ligne verte direction Angrignon

Les heures de passages maintenant connues à la seconde près, il devient possible de connaître le profil de charge des rames de métro dans le temps, en leur associant les données de transaction de cartes à puce correspondantes.

4.3 Applications

4.3.1 Profils de charge : transactions cartes à puce et données GTFS enrichies

La charge à bord des rames est un aspect fondamental du processus de planification du service de métro. C'est en fonction de celle-ci que l'on détermine la fréquence de passage des rames aux stations pour les différentes périodes de la journée. De plus, la connaissance de la charge à bord des véhicules permet de connaître le nombre d'usagers touchés dans le cas d'une interruption de service causé par une panne. Les données de cartes à puce offrent la possibilité de calculer ces attributs, qui devaient auparavant être estimés.

Les données GTFS, dont la précision a été portée à la seconde près, décrivent maintenant la position des rames de manière plus stricte. En effet, les temps passés en station sont maintenant pris en considération. De cette façon, il est possible d'associer chaque transaction carte à puce à une rame, en fonction de l'heure à laquelle la transaction est réalisée.

La génération des profils de charge est faite en respectant l'hypothèse qu'un usager embarque dans la première rame qui passe après l'heure de transaction au tourniquet, additionnée au temps de marche, entre les tourniquets et les quais. Ces temps de marche sont dérivés des données de profondeur des quais.

4.3.1.1 Méthode

La méthode employée est la suivante :

1. Choix de l'ensemble de transactions à associer;
2. Définition d'une direction pour chaque transaction;
3. Énumération des heures de passage pour l'échantillon de données de transactions choisi;
4. Association des transactions aux passages des véhicules.

4.3.1.1.1 Choix de l'ensemble de transactions

D'abord, l'ensemble de transactions à associer à des véhicules doit être défini. Les aspects dont il faut tenir compte pour le choix de l'échantillon sont, notamment : les lignes sur lesquelles sont situées les stations d'origine et de destination du déplacement, et la présence d'une ou plusieurs correspondances. Ces propriétés définissent la plage de l'horaire planifié à extraire des données GTFS, qui sera comparée aux heures de transactions pour permettre l'association entre les deux groupes de données. Étant donné que l'itinéraire d'un usager dépend de facteurs plus ou moins prévisibles tels que le fait d'être assis dans un véhicule ou bien le premier véhicule qui passe comme condition d'embarquement, l'analyse se fera sur un échantillon de transaction dont les

déplacements excluent toute correspondance : l'origine et la destination du déplacement se situent sur la ligne bleue. Cela permettra de simplifier les manipulations, et de vérifier les résultats avant de les répéter sur des ensembles de données complets.

4.3.1.1.2 Définition d'une direction pour chaque transaction

Chaque déplacement implique une origine et une destination. La direction correspondante est numérotée 1 ou 2, selon une séquence fixée au départ. Par exemple, pour la ligne bleue, on définit la direction Saint-Michel comme étant 1, et la direction Snowdon comme étant 2. Cette numérotation permet d'appliquer une condition de recherche programmable, et ainsi, de faire le lien entre les données de transaction et les données de service planifié.

4.3.1.1.3 Heures de passages pour l'ensemble choisi

Les temps de passages pour l'ensemble de données de transactions choisi doivent inclure les passages à toutes les stations des lignes sur lesquelles sont effectués les déplacements. On listera les heures de passages dans un format spécifique, illustré à la Figure 4.12. Cette structure présente les données d'une ligne seulement, les voyages y sont triés par heures de départ de façon décroissante, et les directions ont été fixées à 1 (Saint-Michel) et 2 (Snowdon). On retrouve donc 374 départs au total sur la ligne bleue, et le dernier départ se fait à 0 h 15 dans les 2 directions.

					00:15:00	00:15:00	00:06:00	00:03:00
					374	373	372	371
origine	destination	direction	stop_id	Rangée	5	11	9	4
49	64	1	5	4	00:26:00			00:15:00
49	64	1	49	5	00:15:00			00:03:00
49	64	1	55	6	00:16:12			00:04:18
49	64	1	56	7	00:17:41			00:05:55
49	64	1	57	8	00:19:00			00:07:21
49	64	1	58	9	00:20:46			00:09:17
49	64	1	59	10	00:22:14			00:10:53
49	64	1	60	11	00:23:38			00:12:25
49	64	1	61	12	00:24:53			00:13:47
49	64	1	62	13	00:27:17			00:16:17
49	64	1	63	14	00:28:26			00:17:26
49	64	1	64	15	00:29:59			00:18:59
64	49	2	5	16		00:18:59	00:09:59	
64	49	2	49	17		00:29:59	00:20:59	
64	49	2	55	18		00:28:34	00:19:34	
64	49	2	56	19		00:27:11	00:18:11	
64	49	2	57	20		00:25:56	00:16:56	
64	49	2	58	21		00:24:14	00:15:14	
64	49	2	59	22		00:22:51	00:13:51	
64	49	2	60	23		00:21:38	00:12:38	
64	49	2	61	24		00:20:33	00:11:33	
64	49	2	62	25		00:17:23	00:08:23	
64	49	2	63	26		00:15:56	00:06:56	
64	49	2	64	27		00:15:00	00:06:00	

Figure 4.12 – Structure des données pour l'énumération des heures de passage

4.3.1.1.4 Association des transactions aux rames

L'association entre les transactions et les rames se fait par une fonction « match » du logiciel Excel. Celle-ci permet, pour 2 ensembles de données, d'associer 2 valeurs selon l'écart numérique qui les sépare. Donc, si une transaction se fait à 13:03:00 à la station Université-de-Montréal, et que le temps de marche est de 15 secondes, la fonction permet d'associer cette heure à l'heure de passage de la première rame qui passe après 13:03:15, soit la rame 6, à 13:05:00.

4.3.1.2 Résultat

En appliquant cette méthode aux données de transactions dont l'origine et la destination sont sur la ligne 5, on obtient le résultat illustré à la Figure 4.13. L'échantillon comporte 35 503 transactions effectuées entre des stations de la ligne 5 uniquement, et pour lesquels la station de

destination a pu être dérivée. Les correspondances à partir d'autres lignes ne sont pas prises en compte.

Les temps de marche entre les tourniquets et les quais ont été calculés en fonction d'une vitesse de marche de 0.35 m de profondeur/seconde lorsqu'il y a un escalier mécanique, et 0.25 m de profondeur/seconde s'il n'y a qu'un escalier fixe. L'information concernant la profondeur des stations provient du site Metro de Montréal (Métro de Montréal, 2012). Seule la portion de profondeur située après les tourniquets est utilisée pour les calculs de temps de marche. Celle-ci a été estimée en fonction de la position des tourniquets observée dans chaque station.

En direction Saint-Michel, on observe que les rames sont les plus remplies en période de pointe du soir. De plus, les résultats permettent d'affirmer que les usagers dont l'origine et la destination sont sur la ligne bleue comptent, au maximum, pour 21% de la charge maximale des rames (250/1200). Enfin, l'écart d'achalandage par véhicule entre l'affectation des données avec une précision à la seconde et une précision à la minute est d'environ 15%.

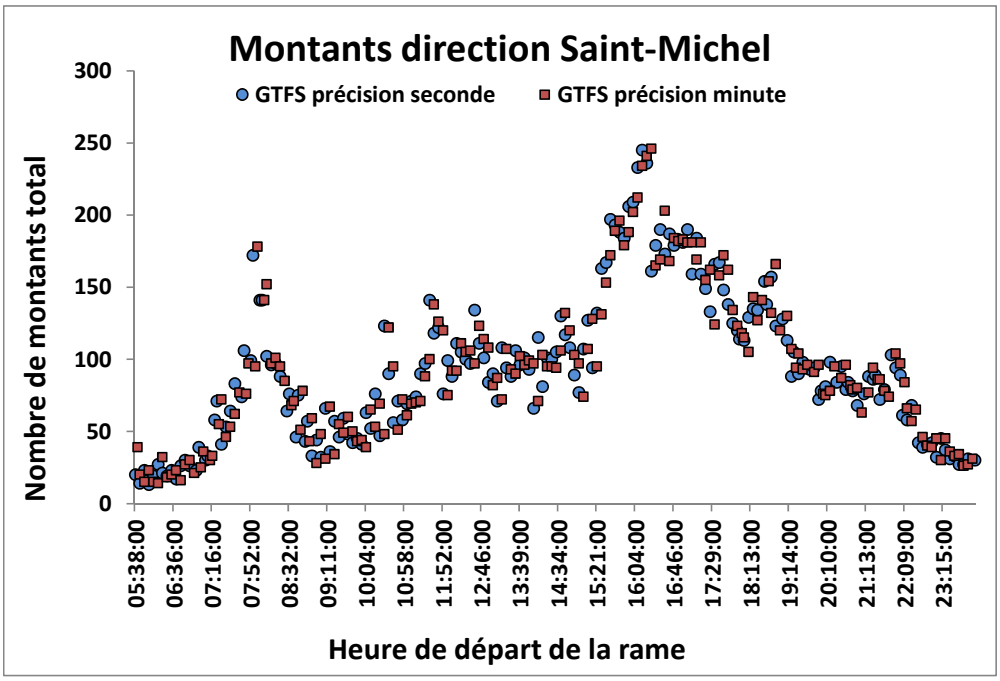


Figure 4.13 - Montants par départ sur la ligne bleue en direction Saint-Michel

En dessinant l'évolution du profil de charge du voyage de 16:11:00, dont le nombre de montants est maximal à la figure précédente, on obtient le résultat présenté à la Figure 4.14.

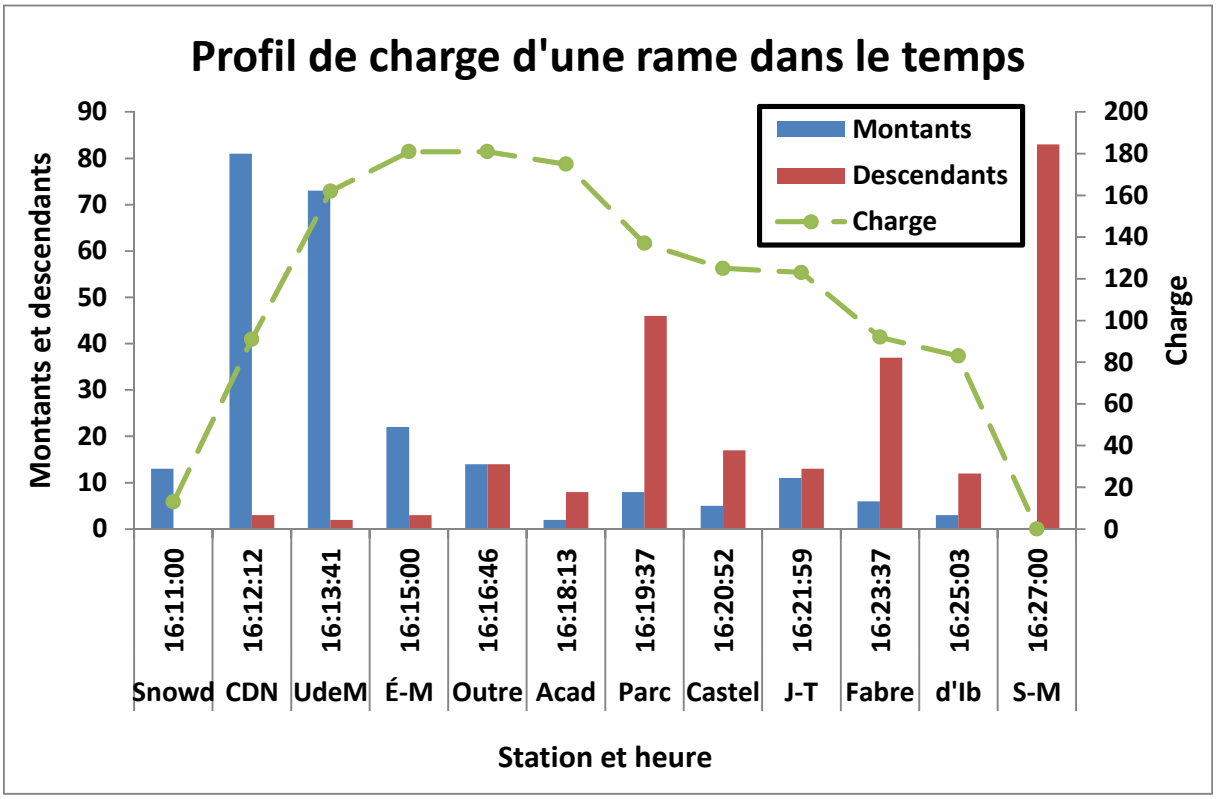


Figure 4.14 - Profil de charge du départ de 16:11:00 en direction Saint-Michel

L'exemple précédent ne tient compte que des transactions dont l'origine et la destination sont situées sur la ligne 5. La situation devient plus lourde à traiter lorsqu'on en vient à traiter tous les mouvements possibles dans le métro de Montréal. En effet, il est impossible de connaître, seulement à partir des données de cartes à puce, l'itinéraire emprunté par l'usager pour se rendre à sa destination, dans le cas d'un déplacement pour lequel une correspondance est requise. Les choix des individus dépendent de conditions propres à chacun, comme le nombre de places assises, le temps de déplacement, ou le nombre de correspondances à effectuer dans un déplacement. Les déplacements plus complexes seront donc simulés à l'aide d'un logiciel de modélisation. Les résultats des simulations pourront être validés grâce à la méthode d'association présentée ci-dessus. En intégrant la même demande au logiciel que celle qui a été utilisée ici, on

devrait retrouver le même profil d'utilisation des rames. C'est ce qui nous permettra d'assurer la correspondance des simulations. Cette méthode permettrait, ultimement, de connaître le profil de charge de chaque rame, pour chacun des 1581 départs de la journée.

4.3.2 Usagers touchés par une panne de service

Maintenant que la position et l'occupation des rames sont connues, différentes analyses peuvent être réalisées. L'une d'entre elles consiste à calculer le nombre d'usagers touchés par une panne de service à une heure donnée. Cela implique la connaissance du nombre de véhicules en service, le nombre d'usagers à bord des véhicules, et le nombre d'usagers qui s'accumulent aux différentes stations.

4.3.2.1 Mise en situation

Le sous-ensemble de données utilisé pour la ligne 5 seulement sera réutilisé pour illustrer la problématique. Il faut noter que les transactions uniques auraient pu être intégrées, mais elles ont été ignorées pour conserver l'ordre de grandeur de l'ensemble utilisé. Donc, en supposant une panne de service sur la ligne 5 en direction Snowdon à 7 h 30, combien d'usagers seraient touchés si la panne durait 15 minutes?

Si la panne débute à 7 h 30, alors on suppose que les véhicules s'immobilisent à la station à laquelle ils sont à partir de 7 h 30 s'ils étaient arrêtés, ou à la station à laquelle ils arriveront tout juste après 7 h 30 s'ils étaient en mouvement.

À partir des données de service planifié, on observe qu'à 7 h 30, il y a exactement 4 rames en service sur la ligne 5 en direction Snowdon selon les données GTFS. Ces rames se trouvent respectivement aux stations Saint-Michel, Jean-Talon, Outremont et Côte-des-Neiges, tel qu'illustré aux Figure 4.15 et Figure 4.16.

Une fois la position des véhicules connue, on doit déterminer combien d'usagers sont à bord des véhicules à 7 h 30. Pour ce faire, on calcule la somme des montants, pour les 4 rames en service entre 5 h 30 et 7 h 30, à laquelle on soustrait le nombre de descendants pour la même période. Pour l'ensemble de données utilisé, le nombre de passagers à bord des véhicules au moment de la panne est de 253.

Count of Ligne		origin	stop_	stop_i									
		64											
		1	2	3	4	5	6	7	8	9	10	11	12
arrivée	Véhicule	64	63	62	5	61	60	59	58	57	56	55	49
07:30:00	2				1								
	3	1											
07:30:15	1								1				
07:30:35	12											1	
07:30:56	3		1										
07:31:34	2					1							
07:31:57	1									1			
07:32:00	12												1

Figure 4.15 - Données de service planifié : nombre de véhicules en service

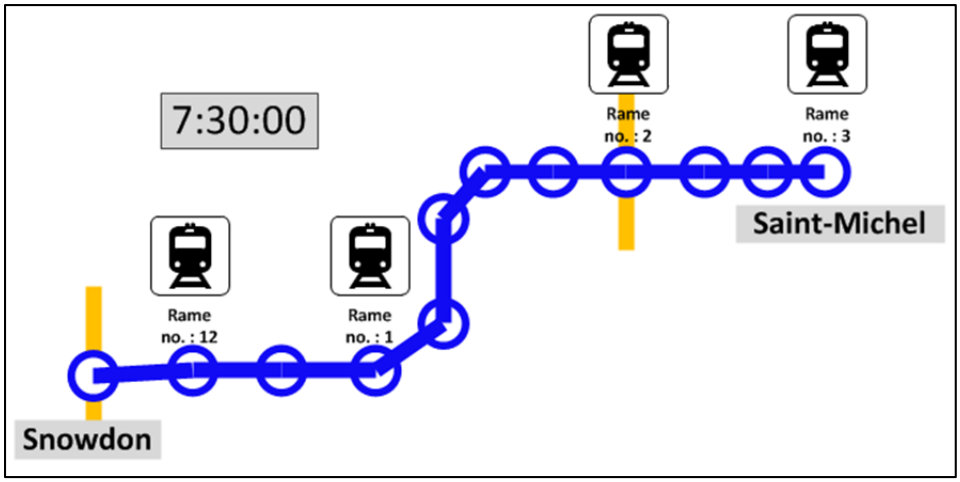


Figure 4.16 - Position des rames à 7 h 30 min sur la ligne 5 en direction Snowdon

Ensuite, les passagers qui ont franchi les tourniquets sans pouvoir attraper le dernier passage à leur station restent donc coincés sur leur quai. Ce sont donc les transactions effectuées entre l'heure du dernier passage avant la panne, et l'heure de la panne elle-même qui sont visées. Pour s'assurer de la précision, il convient de tenir compte du temps de marche. En effet, un usager qui franchit le tourniquet à 7 h 25 ne pourra pas attraper le départ de 7 h 25 min 30 s si son temps de

marche est supérieur à 30 secondes. Le nombre d'usagers qui franchissent les tourniquets sans être capables d'embarquer dans le dernier départ avant la panne est de 50.

Enfin, les usagers s'accumuleront sur les quais entre 7 h 30 et 7 h 45, puisque la panne dure 15 minutes. Donc, il suffit de comptabiliser toutes les transactions effectuées dans cet intervalle de temps, pour obtenir un total de 760 usagers, dont environ la moitié provient de la station Saint-Michel. Ce qui fait qu'au total, ce sont 1063 déplacements qui seraient affectés par la panne.

CHAPITRE 5 MODÉLISATION DU RÉSEAU DE MÉTRO DE MONTRÉAL

La disponibilité de données fines de service planifié et des données de transactions de cartes à puce permet la modélisation du métro à un très fin niveau de résolution. En effet, en intégrant les données à un logiciel de simulation, il devient possible de suivre dynamiquement la distribution de la charge dans chacune des rames de métro pendant la journée. La définition des objets et des paramètres du modèle sont la clé d'une représentation fidèle à ce que l'on observe dans une situation réelle. Ainsi, ce chapitre sera consacré aux détails qui concernent la codification du réseau, l'affectation de déplacements, la simulation des déplacements, et la visualisation et l'analyse des résultats.

La variété d'outils « open source », ou logiciels ouverts, de qualité se fait de plus en plus grande. Les bénéfices de tels logiciels peuvent être immenses, à condition que leur utilisation soit conviviale, et qu'ils permettent des expérimentations reproductibles. C'est le cas des deux logiciels qui seront utilisés tout au long de ce chapitre, soient Quantum GIS et TRANSIMS.

5.1 Présentation des logiciels

5.1.1 Quantum GIS

Publié sous la GNU Public License, Quantum GIS est défini comme étant « un système d'information géographique (SIG) aisé à prendre en mains et s'utilisant sous Linux, Unix, Mac OS X, et Windows, et qui supporte les formats vectoriels, rasters et de bases de données » (QGIS, 2011a). Le logiciel présente plusieurs fonctionnalités de création de carte et d'exploration interactive des données spatiales dont : l'édition, la visualisation et la recherche d'attributs, l'étiquetage d'entités, et la modification de la symbologie (QGIS, 2011b). Ces fonctions seront des plus utiles dans les prochaines sections, car elles permettront de distinguer rapidement les nombreux objets desquels il sera question. De plus, QGIS admet les chemins relatifs lors de la sauvegarde d'un projet, ce qui facilite le transport des données dans différents répertoires.

5.1.2 TRANSIMS

5.1.2.1 Introduction

TRANSIMS, ou *Transportation Analysis Simulation System*, est un ensemble de procédures visant la simulation à la fois macro et microscopique du transport multimodal, basé sur un système de calcul par automate cellulaire, et publié sous la NASA Open Source Agreement Version 1.3 (TRANSIMS Google Code, 2011a). Le projet, soutenu par le U.S. Department of Transport, et par l'Environment Protection Agency, a été confié au Los Alamos National Laboratory suite à l'adoption, aux États-Unis, de différents accords visant à réduire la pollution de l'air, et à rendre les transports plus efficaces (National Transportation Library, 1995).

Pour faciliter l'utilisation de la suite de programmes de TRANSIMS, à l'origine exécutée au moyen de fichiers « batch », l'outil TRANSIMS studio a été développé, et rendu disponible au printemps 2010. Cet outil combine une interface graphique et un Run Time Environment (RTE) propre à TRANSIMS. De plus, il permet l'exécution des programmes à partir de procédures python (Argonne National Laboratory, 2010).

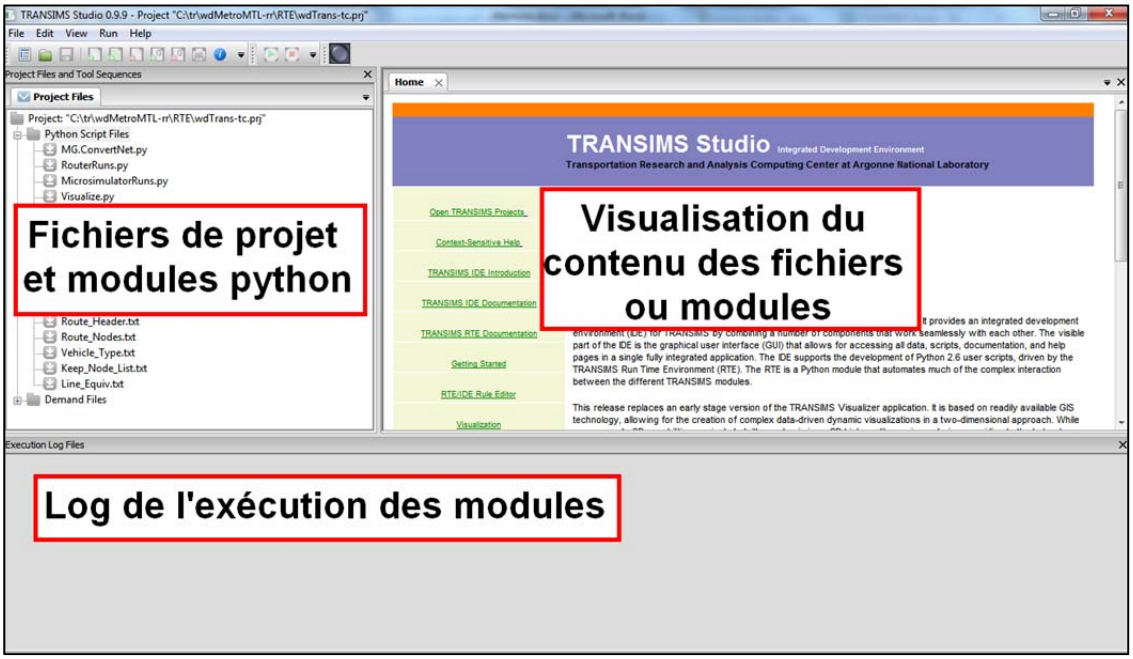


Figure 5.1 - TRANSIMS Studio

5.1.2.2 Fonctionnement

TRANSIMS a été développé pour permettre la simulation d'un réseau de transport codifié, en passant successivement par les étapes : de génération d'une population synthétique, de génération de la demande, d'affectation des déplacements et de microsimulation. Toutefois, les expérimentations réalisées dans le cadre de ce projet touchent seulement la codification, l'affectation des déplacements et la microsimulation, puisque la demande sera basée sur les données de transactions de cartes à puce. Le fonctionnement de TRANSIMS tel qu'utilisé pour la réalisation des expérimentations qui suivront est illustré à la Figure 5.2. Par ailleurs, la réalisation d'un projet de modélisation passe par une organisation systématique des répertoires de projet, tel qu'illustré au Tableau 5.1.

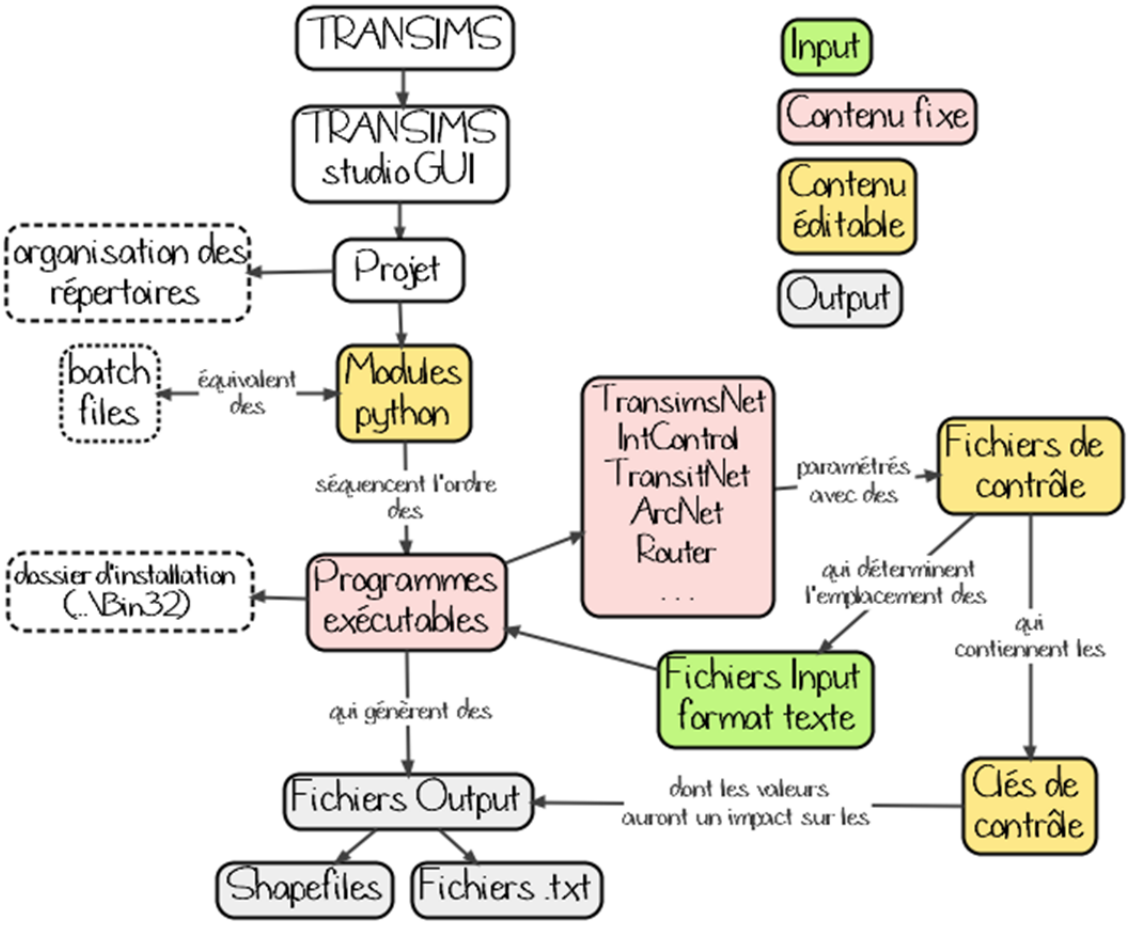
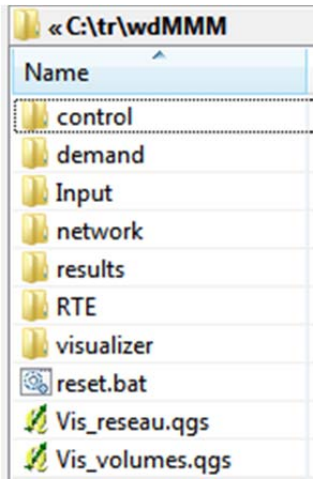


Figure 5.2 - Schéma organisationnel de l'utilisation de TRANSIMS

TRANSIMS studio, les modules pythons et les fichiers de contrôle de la version 4 de TRANSIMS peuvent être téléchargés sur le site Source Forge (Source Forge, 2010a et Source Forge, 2010b).

Tableau 5.1 - Répertoire de projet TRANSIMS et son contenu

Répertoire de projet	Contenu
	**/control** : Fichiers de contrôle
	/demand : Fichiers de déplacements, de possession véhiculaire, et de type de véhicules
	/Input : Fichiers nœuds, liens, arrêts TC, et autres fichiers requis pour la codification du réseau au niveau de précision désiré
	/network : Fichiers résultats de la codification exécutée par TRANSIMS (fichiers textes et shapefiles)
	/results : Fichiers résultats de l'affectation des déplacements et de la microsimulation exécutées par TRANSIMS (fichiers textes et shapefiles)
	/RTE : Modules pythons (.py), fichier de projet (.prj), et rapports générés par les différents modules
	/visualizer : Fichier de contrôle du TRANSIMS Visualizer et fichiers vidéo
	reset.bat : Fichier batch permettant de vider les dossiers/network et/results (N.B. : ce fichier n'est pas associé à TRANSIMS)
	Vis_reseau.qgs et **Vis_volumes.qgs :** Projet QGIS permettant de visualiser les shapefiles générés par TRANSIMS

TRANSIMS fonctionne sur un principe de lieux d'activités (activity location), comme l'illustre la Figure 5.3. En effet, un déplacement est effectué entre 2 lieux d'activités, qui peuvent être un domicile, un arrêt d'autobus, un lieu de travail, etc. De plus, tout déplacement comportera d'abord une portion « accès au mode de transport », faite à pied. La portion suivante constitue le déplacement dans le mode choisi, du point de départ au point d'arrivée. La dernière portion du

déplacement est « l'accès à la destination », qui modélise la marche du piéton de son mode de transport à sa destination.

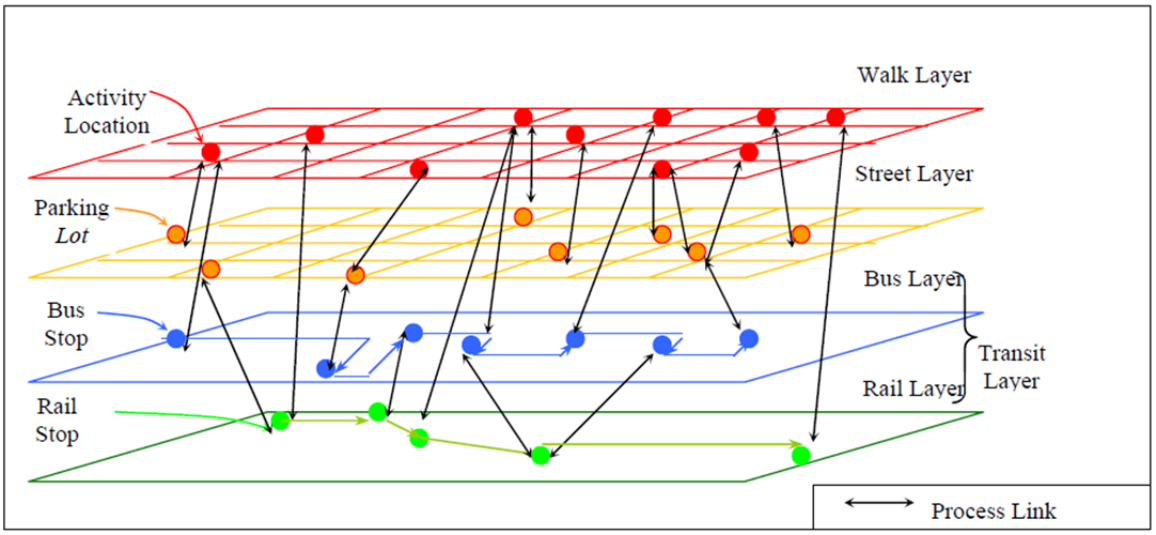


Figure 5.3 - Lieux d'activité et couches telles qu'interprétées par TRANSIMS (FHWA TMIP, 2005)

5.1.2.3 Études de cas

5.1.2.3.1 Dallas-Ft. Worth, TX

L'objectif de cette étude de cas était de définir un modèle qui démontrait la robustesse des programmes de TRANSIMS, et qui restait assez simple pour pouvoir être reproduit dans les sociétés de transports (Beckman, 1997). L'expérimentation consistait à simuler près de 3.5 millions de déplacements générés par TRANSIMS à partir de données d'enquête de 1990, et de comparer les résultats de la simulation originale à deux scénarios où la géométrie avait été modifiée. Différentes analyses ont été conduites, dont la variabilité du temps de transport des usagers, et la validation de la génération automatique d'éléments aléatoires.

5.1.2.3.2 Portland, OR

L'étude de cas de la ville de Portland, Oregon, consistait à utiliser TRANSIMS pour modéliser le transport dans toute la région métropolitaine. De plus, on voulait savoir si l'intégration de données d'enquête et de réseaux codifiés pour d'autres modèles était possible et efficace dans TRANSIMS (AECOM, 2009). Ce projet est à beaucoup plus grande échelle que celui de Dallas, car on procède à la codification et à la simulation du réseau en entier plutôt qu'à la seule étude d'un scénario particulier, pour une demande d'environ 5 millions de déplacements. Par ailleurs, on y étudie en détail le modèle de « générateurs d'activités ». Les paramètres de simulation y sont fournis, de même que les résultats de celles-ci.

5.1.2.3.3 Alexandria, VA

Le projet le plus riche au point de vue expérimental est certainement celui d'Alexandria. En effet, sans être rigoureusement documenté, le projet est entièrement téléchargeable et modifiable. On peut tester le fonctionnement des modules, analyser le format des fichiers d'entrée et de sortie, et modifier différents paramètres pour observer l'impact de ceux-ci sur les simulations. De cette façon, il est possible de comprendre la logique du logiciel, et de s'en servir pour la modélisation de cas spécifiques. Le projet peut être téléchargé à l'adresse suivante : http://sourceforge.net/projects/transims/files/test%20data/Alexandria%202008-06-10/.

5.2 Sommaire de l'expérimentation

Le but ultime de l'expérimentation est de comprendre le fonctionnement de TRANSIMS pour la création d'un modèle multimodal qui comprend le métro, les bus, et l'automobile, en intégrant les données de service planifié et de transactions de cartes à puce. Les expérimentations suivantes ne touchent cependant qu'à la modélisation du métro. Les principales préoccupations desquelles il sera question sont l'organisation d'un projet (Tableau 5.1), les formats de fichiers, les principes de codification, le contrôle des paramètres, de même que le format, la cohérence et la validité des fichiers résultat générés par TRANSIMS.

La modélisation inclut les 68 stations du métro de Montréal, et les temps de passages des 1581 départs planifiés lors d'une journée moyenne de semaine selon les données GTFS. Au total, on compte 11 lignes-direction : 2 par ligne sauf pour la ligne 2, qui elle, possède 5 tracés différents. La vitesse maximale des rames a été fixée à 72 km/h (20 m/s dans TRANSIMS), et leur capacité à 1200 passagers. L'étape de génération des déplacements à partir d'une population sera remplacée par une demande existante, soit les transactions de données de cartes à puce du métro pour une journée. Seules les 645 880 transactions pour lesquelles une destination a pu être dérivée seront utilisées sur les 797 975 transactions valides enregistrées.

5.3 ConvertNet : codification du réseau de métro

La première étape de la modélisation consiste à codifier le réseau de métro selon les spécifications requises par TRANSIMS. Le module python ConvertNet.py permet d'accomplir cette tâche en exécutant, dans l'ordre, les programmes TransimsNet, TransitNet et ArcNet.

Les coordonnées des objets codifiées doivent être de type UTM afin de permettre le calcul et l'interprétation de certains attributs par TRANSIMS. Le fonctionnement du module ConvertNet est illustré à la Figure 5.4. Tous les fichiers générés par le module ConvertNet se retrouvent dans le dossier « ../network » de l'organisation présentée au Tableau 5.1.

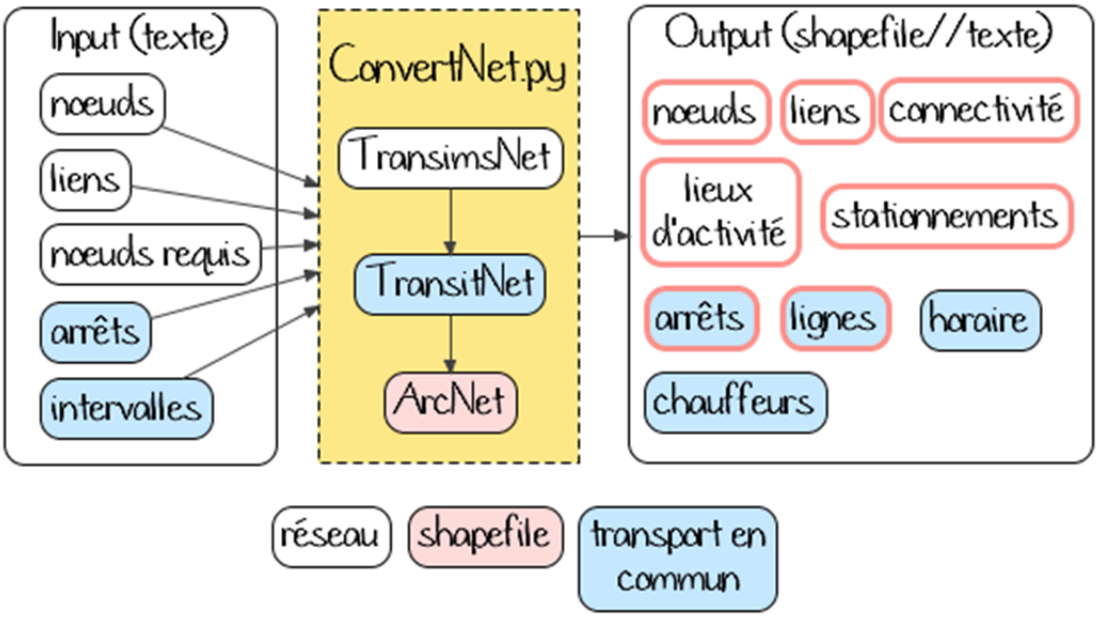


Figure 5.4 - Module ConvertNet.py

5.3.1 TransimsNet

Chaque programme permet la codification de différents objets. D'abord, le programme TransimsNet est celui qui prend en charge tous les objets relatifs aux « walk layer » et « street layer » de la Figure 5.3. On y codifie, à travers des fichiers inputs textes formatés pour TRANSIMS, les nœuds, les liens et les zones. À l'exécution du programme, de nouveaux objets sont créés automatiquement, dont : les « process link », qui font les connections entre les liens aux intersections, les lieux d'activités, d'où proviennent et se destinent tous les déplacements, et les stationnements pour automobiles.

Les paramètres d'exécution, qui peuvent être modifiés dans le fichier de contrôle du programme, déterminent les propriétés des objets générés. On peut, par exemple, modifier la valeur de la clé de contrôle liée à l'espacement maximal des lieux d'activités générés automatiquement sur un lien. Le rôle de chacune des clés est défini dans la documentation (TRANSIMS Google Code, 2011b). Tous les paramètres des fichiers de contrôle, de même que les fichiers input utilisés dans leur format requis se trouvent aux annexes C et D respectivement.

Dans le cas du métro de Montréal, la codification a été faite en spécifiant des liens de type « HeavyRail » sauf un, de type local, qui doit être codifié pour permettre d'exécuter le module TransimsNet sans rencontrer d'erreur. TRANSIMS n'autorise pas l'exécution des différents modules tant qu'il ne reconnaît pas au moins 1 lien routier assez long pour y générer des lieux d'activité.

TRANSIMS calcule automatiquement la longueur des liens à partir des nœuds, et interprète les types associés aux liens pour générer des attributs par défaut. TRANSIMS supprime les nœuds considérés inutiles, c'est pourquoi il faut spécifier une liste qui contient les nœuds à ne pas supprimer, appelée la « keep_node_list ».

5.3.2 TransitNet

Le second programme exécuté dans le module ConvertNet.py est TransitNet. Ce programme a trait aux objets liés au transport en commun : stations, lignes, intervalles et chauffeurs. Les fichiers input sont : les stations séquencées dans l'ordre de passage des véhicules et les temps moyen interstations pour chaque ligne-direction (../input/Route_Nodes.txt), et les intervalles de départ des rames (../input/Route_Header.txt). Le fichier de contrôle TransitNet.ctl permet de diviser la journée en périodes de service, et d'adapter l'horaire en conséquence.

Le fichier Route_Nodes.txt peut aussi prendre en input les « dwell times », c'est-à-dire les temps d'attente en station, ainsi que la vitesse interstation. Si ces paramètres ne sont pas fournis, TRANSIMS les calculera automatiquement selon les paramètres fournis dans la spécification des liens et des types de véhicules. Par ailleurs, le fichier Route_Header.txt demande de spécifier les « offset ». Ceux-ci correspondent au temps écoulé entre le début de la période de service spécifiée dans le fichier de contrôle, et le premier départ d'un véhicule. Si la valeur de « offset » est laissée à 0, le premier départ sera fait en même temps que l'heure spécifiée pour la division des périodes de service; si la valeur est négative, une valeur sera déterminée au hasard par TRANSIMS.

À partir des paramètres fournis en input, le programme génère un horaire de passage aux stations, assigne un chauffeur à chaque départ, et associe des lieux d'activités aux stations. Afin de pouvoir réutiliser ces lieux d'activités à l'étape d'affectation des déplacements, il faut leur apporter une modification. En effet, toutes les valeurs du champ « Zone » contenues dans le fichier output « ../network/Activity_Location_2 » doivent être remplacée par la valeur « 1 ».

La validation exhaustive des fichiers générés par TRANSIMS n'a pas été complétée. Cependant, si l'utilisation de TRANSIMS s'avère pertinente, cette tâche devra être accomplie. Pour l'instant, les résultats obtenus par défaut sont sommairement vérifiés, et considérés comme valides si aucun problème majeur n'est observé.

5.3.2.1 Intégration de l'horaire GTFS

Il est possible d'intégrer les données d'horaire planifié GTFS dans TRANSIMS en les formatant de la même façon que les données du fichier output « ../network/Transit_Schedule » généré par le programme TransitNet. Pour ce faire, il faut respecter la numérotation des lignes suivie dans le fichier Route_Nodes.txt, utiliser les numéros d'arrêts du fichier Transit_Schedule, et formater les heures de passages en secondes après minuit. De cette façon, un nouveau fichier Transit_Schedule peut être produit en calquant celui généré par TRANSIMS avec, comme seule différence, les temps de passages officiels fournis dans les données GTFS. En remplaçant le fichier original par le fichier créé, TRANSIMS pourra utiliser l'horaire planifié réel, plutôt qu'un horaire basé sur des intervalles de service à différentes périodes de la journée. Toutefois, il faut noter que TRANSIMS respecte seulement les heures de départ aux terminus. Les heures de passage aux stations intermédiaires sont calculées en se basant sur la vitesse des véhicules, et sur la demande aux différentes stations.

Il est possible de vérifier le résultat de la simulation en comparant les heures de passages avec les heures fournies dans le fichier GTFS.

5.3.3 ArcNet

Le dernier programme exécuté dans le module ConvertNet.py, ArcNet, permet de générer les shapefiles du réseau codifié. Les fichiers shapefiles créés sont identifiés par un contour rouge à la Figure 5.4. Une fois importés dans QGIS, il est possible de modifier leur aspect afin de distinguer chacun des objets.

La Figure 5.6 présente 4 différents objets du réseau de métro : la station, représentée par le point rouge au centre, les lignes, verte, orange et jaune, les quais, représentés par les symboles numérotés, et les liens interstations de chaque ligne-direction représentés par des lignes noires pointillées. En gris, on voit la connectivité entre les quais à la station Berri-UQAM, qui est générée automatiquement par TRANSIMS.

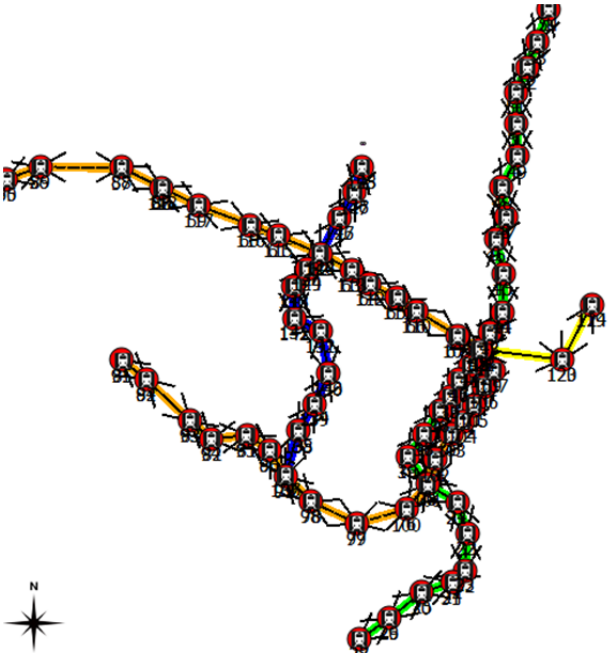


Figure 5.5 - Réseau de métro codifié dans TRANSIMS et illustré avec QGIS

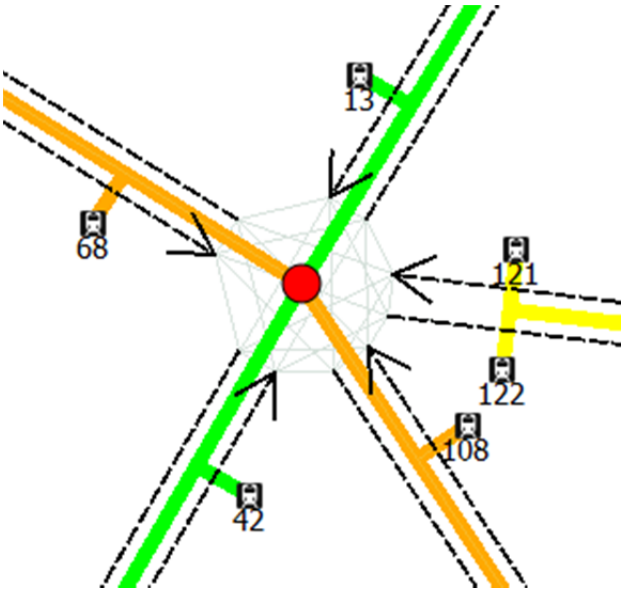


Figure 5.6 - Station Berri-UQAM codifiée dans TRANSIMS et illustrée avec QGIS

Le nombre de quais total du réseau de métro se calcule à l'aide de la formule suivante :

$$Q = 2 * \sum_{L=1}^{n} s_L$$

où Q correspond au nombre total de quais, et s_L au nombre de stations de la ligne L (pour les lignes 1 à n). Pour le métro de Montréal, on obtient [2*(27 + 31 + 12 + 3)] = 146 quais. Le nombre de liens interstation, lui, se calcule de la façon suivante :

$$li = \sum_{L=1}^{n} (s_L - 1)$$

où li correspond au nombre de liens interstation, et s_L au nombre de ligne-stations de la ligne L (pour les lignes 1 à n). On obtient donc [26 + 30 + 11 + 2] = 69 liens interstation.

5.4 RouterRuns : affectation des déplacements

La deuxième étape de la modélisation consiste à affecter les déplacements sur le réseau codifié à l'aide du module ConvertNet. L'affectation peut être faite avec un nombre déterminé d'itérations. Dans le cadre de cette expérimentation, il n'y aura qu'une seule itération. En effet, une itération suffit, car on tente de valider le fonctionnement général de TRANSIMS, et non d'analyser les procédures d'optimisation.

Le module RouterRuns prend d'abord, comme fichiers input, les fichiers constituants le réseau codifié précédemment et contenus dans le dossier « ../network ». Il prend ensuite le fichier qui caractérise la demande. Celui-ci doit indiquer le ménage, la personne, le mode, et le numéro de véhicule s'il s'agit d'un déplacement auto. Il doit aussi spécifier le numéro de lieu d'activité d'origine et de destination, ainsi que l'heure de départ. L'heure d'arrivée, pour sa part, sera calculée par TRANSIMS. Un extrait du fichier de la demande, de même que le fichier de types de véhicules sont fournis en annexe. Le schéma du fonctionnement du module RouterRuns est présenté à la Figure 5.7.

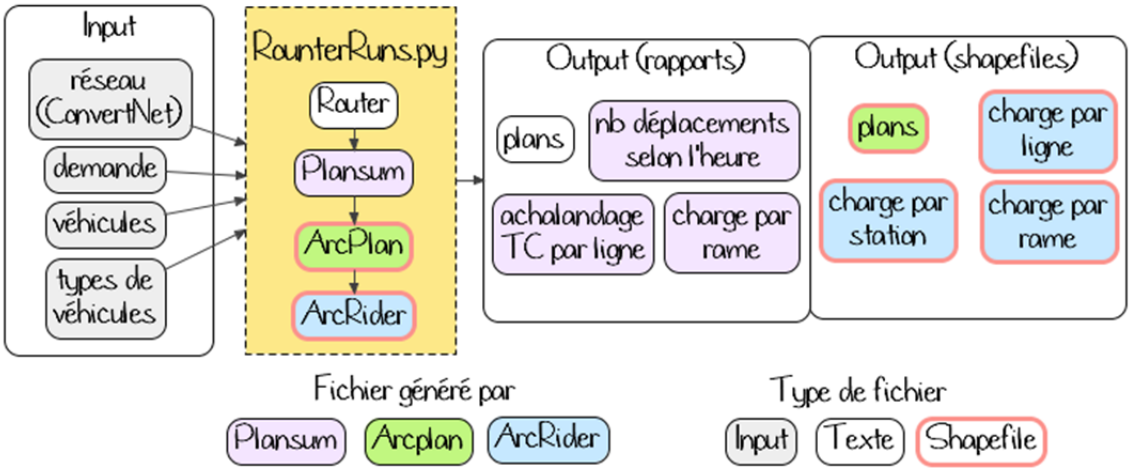


Figure 5.7 - Module RouterRuns.py

5.4.1 Router

Le premier programme exécuté dans le module RouterRuns est le Router. Celui-ci détermine l'itinéraire emprunté pour chaque déplacement du fichier de demande « trips.txt » fourni en annexe. Le fichier « Plans », résultat de l'exécution du programme Router, contient tous les itinéraires, et sera le fichier input principal des programmes suivants.

Les déplacements effectués dans cette expérimentation sont tous de type transport en commun. Dans le fichier « trips.txt », le mode correspondant est « 3 ». Pour ce mode, le numéro de véhicule est toujours « 0 », étant donné que le véhicule dans lequel l'usager embarquera est attribué automatiquement par TRANSIMS, en se basant sur l'heure à laquelle le déplacement est réalisé. L'heure de départ doit être spécifiée, tandis que l'heure d'arrivée sera modélisée par TRANSIMS. Il faut spécifier qu'un départ réalisé après minuit doit être codifié 24:05:00 plutôt que 00:05:00.

Le type de véhicule auquel on s'intéresse dans le cadre de ce projet est le type 9 (RapidRail). Ses caractéristiques sont présentées dans le fichier « vehicle_type.txt » fourni en annexe. Seules les

propriétés de vitesse maximale et de capacité ont été modifiées, toutes les autres propriétés sont celles fournies par défaut.

5.4.2 PlanSum

Une fois les déplacements affectés sur le réseau, le programme PlanSum permet de résumer l'utilisation du réseau par mode, par lien, et par division de temps. En effet, ce programme permet la génération de plusieurs rapports caractérisant les débits, l'achalandage en transport en commun et la répartition temporelle des déplacements. Il permet aussi de connaître toutes les statistiques de bases sur les déplacements : temps total de déplacement, passagers-miles, temps moyen de déplacement, distance moyenne parcourue, etc.

Échantillon : 645 125 déplacements

Passagers-Heures : 129 107

Passagers-Miles : 3 125 432 (5 029 895 passagers-km)

Nombre de correspondances : 267 077

Déplacements non simulés : 534

5.4.3 ArcPlan

Par la suite, l'exécution du programme ArcPlan permet de générer un shapefile qui contient les débits automobiles sur chaque lien, et un autre qui contient les chemins de chaque déplacement. Dans ce projet, il n'y a que le second shapefile qui est pertinent pour valider l'itinéraire établi par TRANSIMS pour chaque déplacement. Sans être explicitement défini dans le shapefile, l'itinéraire emprunté par l'usager peut être déduit des temps passés en véhicule. En effet, les temps passés à bord des rames correspondent aux temps interstations et temps d'attente fournis dans les données GTFS. Il est donc possible de retracer l'itinéraire des usagers en comparant les temps passés à bord des véhicules, car ceux-ci sont segmentés par des temps de marche qui simulent l'embarquement et le débarquement des usagers.

5.4.4 ArcRider

Finalement, le programme ArcRider permet de générer les shapefiles qui concernent l'achalandage du réseau de transport en commun. Les profils de charge par tronçon et par station sont créés automatiquement par le programme ArcRider. Il est d'ailleurs possible de changer l'échelle de la charge par tronçon qui est spécifiée dans le fichier de contrôle en passagers/mètre de largeur. La charge peut ensuite être illustrée dans QGIS, comme à la Figure 5.8, qui reprend l'échantillon présenté à la section 5.4.2. On y présente le profil de charge par tronçon, coloré en vert pour une faible charge et en rouge pour une charge élevée, et le profil de charge par station, par les cercles bleus, dont la superficie croît selon le nombre d'embarquements.

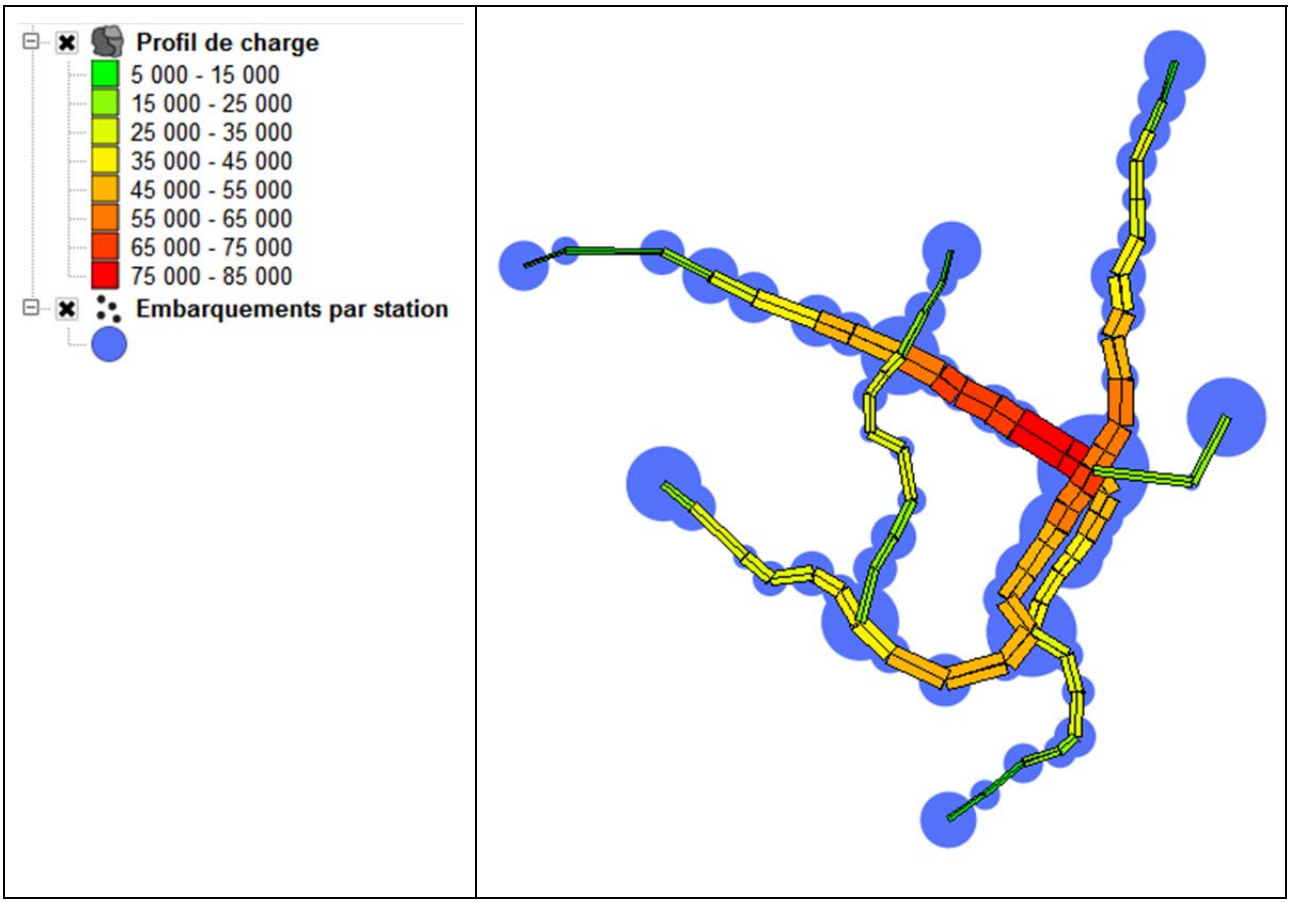


Figure 5.8 - Profil de charge du métro de Montréal pour la journée du 17 mars 2010

5.4.5 Validation

La validation des résultats implique l'analyse de 2 composantes :

- Les déplacements effectués (nombre, heures de départ et d'arrivée, temps de correspondances, itinéraires empruntés);
- Caractéristiques opérationnelles telles que le nombre de départs, la capacité, les heures de passages et les temps de parcours.

Le module d'affectation des déplacements contient 4 programmes principaux. Chacun de ceux-ci génère 2 types de fichiers : les « outputs », soit les fichiers de résultats, et les rapports, qui eux, contiennent les paramètres et le résultat global de l'exécution.

En examinant d'abord les rapports, on s'assure de savoir si des problèmes ont été rencontrés lors de l'exécution d'un programme. On y vérifie entre autres le nombre de déplacements effectués, et le nombre de déplacements « problématiques ». Par exemple, un déplacement peut être classé comme problématique si son heure de départ est après le dernier passage d'un bus ou d'une rame de métro à son arrêt. Une liste complète des problèmes pouvant être rencontrés se trouve dans la documentation de TRANSIMS (fichier Router QR).

Puis, viennent les fichiers « output » générés par TRANSIMS. Ceux-ci sont spécifiques au programme exécuté. Dans le module d'affectation, le programme intitulé « PlanSum » génère le fichier d'utilisation des rames, à la seconde près, de même que des heures de passages des rames aux stations. En groupant les données selon la charge des rames dans le temps, on observe que l'affectation ne tient pas compte des contraintes de capacité et de temps d'attente contenus dans le fichier « vehicle_type ». En effet, des valeurs par défaut sont attribuées à ces paramètres à l'étape d'affectation. Ces valeurs par défaut ne posent pas problème étant donné que les profils de charge sont recalculés à l'étape de microsimulation, cette fois, utilisant les paramètres spécifiés. Ils permettent de se faire une idée de l'ordre de grandeur de l'achalandage du réseau modélisé.

5.4.6 Le cas de la ligne 5

À la section 4.3.1, une méthode d'association des données de transaction aux données de service planifié a été présentée. Celle-ci a permis d'obtenir le profil de charge des rames de métro pour les usagers dont les stations d'origine et de destination étaient sur la ligne bleue uniquement. Maintenant, en utilisant la même demande, il est possible de procéder à une simulation de ces déplacements avec TRANSIMS, et de comparer les résultats obtenus.

La simulation doit respecter toutes les étapes décrites dans les sections 5.3 et 5.4. L'ensemble de données de transactions à simuler doit être extrait de l'ensemble de données original, et doit contenir seulement les données de transactions dont l'origine et la destination sont sur la ligne 5. Au total, cela correspond à 35 503 déplacements à simuler. Les différences observées proviennent du fait que des temps de marche précis ont été appliqués avec la méthode « match ». Aussi, TRANSIMS semble générer des valeurs aléatoires, ce qui doit être pris en considération lors de la création de scénarios comparatifs.

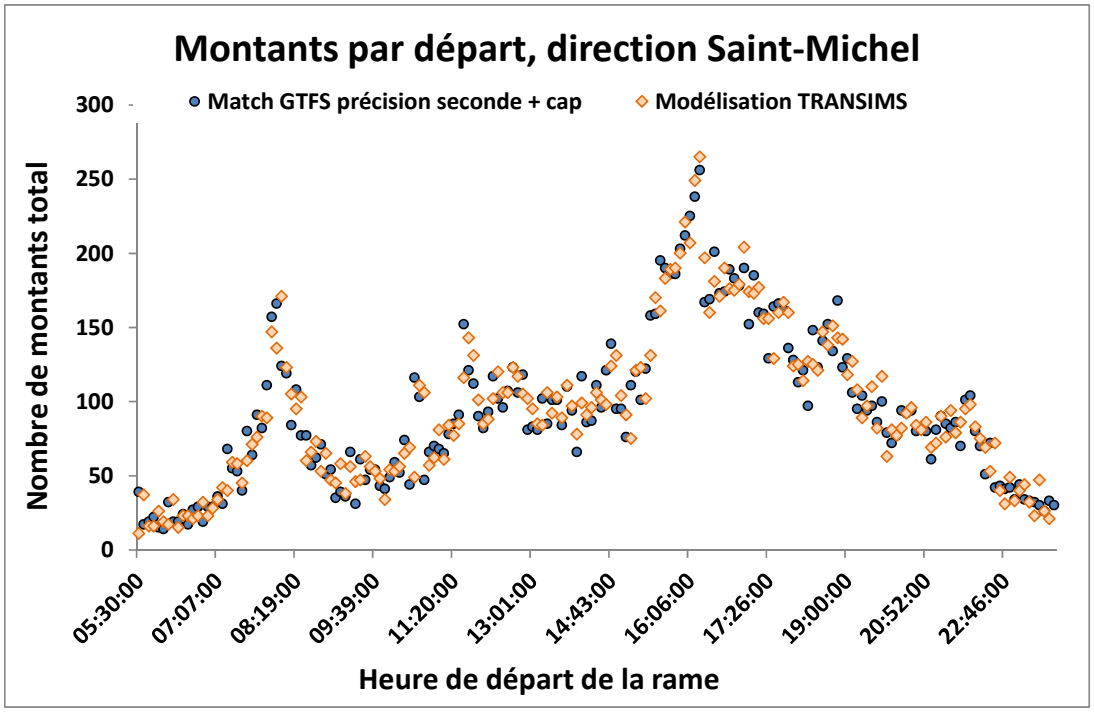


Figure 5.9 - Comparaison des résultats obtenus par simulation à ceux obtenus manuellement

La figure ci-dessus illustre que les résultats de la simulation réalisée avec TRANSIMS respectent avec justesse les résultats obtenus manuellement. Cela nous indique que la simulation respecte la logique décrite dans la méthode manuelle. En effet, un usager qui effectue une transaction à une certaine heure dans le réseau de métro, embarque dans la première rame qui passe à la station empruntée après un temps de marche donné.

De ce sous-ensemble de données, il est possible de retirer le nombre de passagers-heures et le nombre de passagers-km, toujours en utilisant TRANSIMS. D'une part, en simulant 1 déplacement entre chaque origine et chaque destination, TRANSIMS générera un fichier contenant les temps de parcours de chaque déplacement. Le fichier distingue le temps passé en véhicule et le temps d'attente. Les temps en véhicules pourront être multipliés au nombre de déplacements entre chaque paire OD pour obtenir le nombre de passagers-heures. La matrice des temps de parcours pour des déplacements sur la ligne bleue uniquement est présentée à la Figure 5.10. La matrice n'est pas symétrique car le temps entre les stations varie selon le nombre de montants et de descendants à moment de la journée et à chaque station.

		236	534	536	538	540	542	544	546	272	552	554	556
Snowdon	236		01:12	02:40	03:59	05:46	07:13	08:37	09:53	10:59	12:37	14:04	16:00
C-D-N	534	01:25		01:28	02:48	04:34	06:02	07:26	08:41	09:48	11:25	12:52	14:48
U-de-M	536	02:48	01:23		01:19	03:05	04:33	05:57	07:12	08:18	09:56	11:23	13:19
É-M	538	04:03	02:38	01:15		01:46	03:14	04:38	05:53	07:00	08:38	10:03	12:00
Outremont	540	05:45	04:20	02:57	01:42		01:27	02:52	04:07	05:14	06:52	08:17	10:13
Acadie	542	07:08	05:43	04:20	03:05	01:23		01:24	02:39	03:46	05:23	06:50	08:45
Parc	544	08:22	06:57	05:33	04:18	02:36	01:13		01:15	02:22	03:59	05:26	07:21
de Castelnau	546	09:26	08:01	06:38	05:22	03:41	02:18	01:05		01:07	02:44	04:11	06:06
Jean-Talon	272	11:00	09:35	08:12	06:57	05:14	03:52	02:39	01:34		01:37	03:04	05:00
Fabre	552	13:00	11:10	09:48	08:33	06:51	05:52	04:39	03:10	01:36		01:27	03:23
d'Iberville	554	14:04	12:39	11:16	10:47	09:05	06:55	05:43	04:38	03:03	01:28		01:56
Saint-Michel	556	15:00	14:34	13:12	10:57	09:14	07:52	06:38	05:33	04:00	02:23	01:10	

Figure 5.10 - Matrice des temps de parcours pour un déplacement sur la ligne bleue

En effectuant le produit entre ces temps de parcours et la matrice des déplacements dont l'origine et la destination sont sur la ligne 5, il est possible de calculer le total de passagers-heures. Celui-ci peut ensuite être comparé au total généré automatiquement par TRANSIMS, dans le fichier résultat « PlanSum ».

D'autre part, une fois le réseau codifié, il est possible d'en ressortir les distances entre chaque paire origine-destination, comme illustré à la Figure 5.11. On peut ensuite multiplier celles-ci au nombre de déplacements entre les paires OD pour obtenir le nombre total de passagers-km.

		236	534	536	538	540	542	544	546	272	552	554	556
Snowdon	236		0.7	1.9	2.68	3.56	4.49	5.01	5.61	6.31	7.13	7.78	8.98
C-D-N	534	0.7		1.2	1.98	2.86	3.79	4.31	4.91	5.61	6.43	7.08	8.28
U-de-M	536	1.9	1.2		0.78	1.66	2.59	3.11	3.71	4.41	5.23	5.88	7.08
É-M	538	2.68	1.98	0.78		0.88	1.81	2.33	2.93	3.63	4.45	5.1	6.3
Outremont	540	3.56	2.86	1.66	0.88		0.93	1.45	2.05	2.75	3.57	4.22	5.42
Acadie	542	4.49	3.79	2.59	1.81	0.93		0.52	1.12	1.82	2.64	3.29	4.49
Parc	544	5.01	4.31	3.11	2.33	1.45	0.52		0.6	1.3	2.12	2.77	3.97
De Castelnau	546	5.61	4.91	3.71	2.93	2.05	1.12	0.6		0.7	1.52	2.17	3.37
Jean-Talon	272	6.31	5.61	4.41	3.63	2.75	1.82	1.3	0.7		0.82	1.47	2.67
Fabre	552	7.13	6.43	5.23	4.45	3.57	2.64	2.12	1.52	0.82		0.65	1.85
d'Iberville	554	7.78	7.08	5.88	5.1	4.22	3.29	2.77	2.17	1.47	0.65		1.2
Saint-Michel	556	8.98	8.28	7.08	6.3	5.42	4.49	3.97	3.37	2.67	1.85	1.2	

Figure 5.11 - Matrices OD des distances interstations

5.5 Microsimulator : simulation des déplacements

Le troisième et avant-dernier module qui doit être exécuté pour la modélisation est Microsimulator.py. Celui-ci permet de générer des fichiers, appelés « *snapshots* », qui contiennent les mouvements des véhicules à chaque pas de temps déterminé dans le fichier de contrôle, pour la durée de la simulation choisie. L'exécution de ce module permet aussi de générer un fichier qui contient les déplacements pour lesquels un problème est survenu, et d'en identifier la source. La Figure 5.12 illustre le fonctionnement du module.

La différence entre la simulation des déplacements et l'affectation du module précédent est la prise en charge des contraintes de capacité et de temps d'attentes des véhicules aux différentes stations. Le Microsimulator applique les paramètres spécifiés dans le fichier « types de véhicules » à l'affectation du module précédent pour générer des résultats plus fidèles à la réalité.

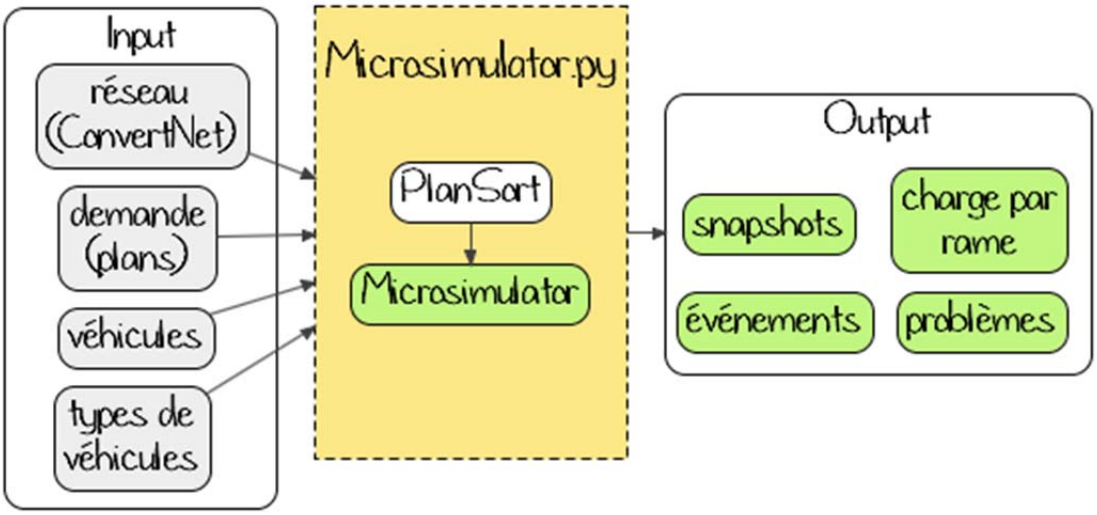


Figure 5.12 - Module Microsimulator.py

5.5.1 PlanSort

D'abord, le programme PlanSort est exécuté pour trier le fichier « plans », qui contient l'itinéraire de chaque déplacement, en ordre croissant d'heure de départ.

5.5.2 Microsimulator

Puis vient le programme Microsimulator, qui lui, génère les fichiers snapshots qui seront utilisés pour visualiser la simulation.

5.6 Visualize : visualisation de la simulation

Le dernier module à exécuter pour compléter la modélisation est Visualize.py. En prenant comme input le fichier *snapshots* créé au module précédent, ce module génère un fichier de snapshots compressé pour la visualisation. Le fonctionnement du module est illustré à la Figure 5.13.

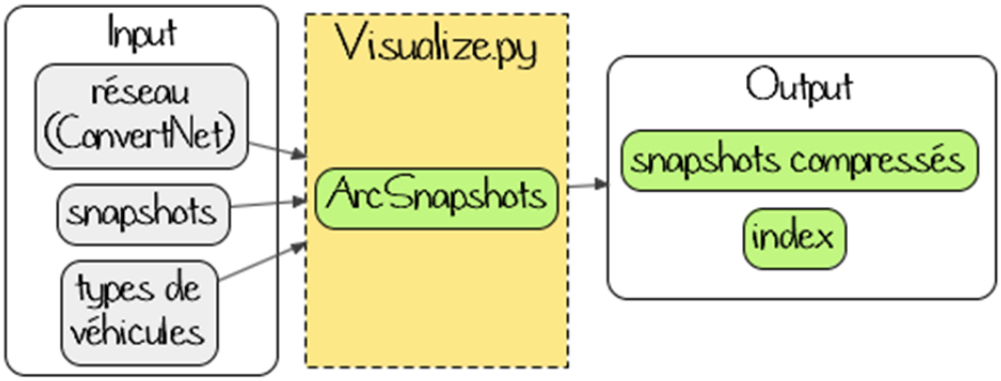


Figure 5.13 - Module Visualize.py

5.6.1 TRANSIMS Visualizer

La suite TRANSIMS studio inclue le TRANSIMS Visualizer. Ce programme permet de visualiser la microsimulation en chargeant, entre autres, le fichier de *snapshots* compressé. Le Visualizer prend en input un fichier de contrôle au format spécifique (fourni en annexe), qui fait appel aux différents fichiers nécessaires à la visualisation. Le Visualizer est présenté à la Figure 5.14.

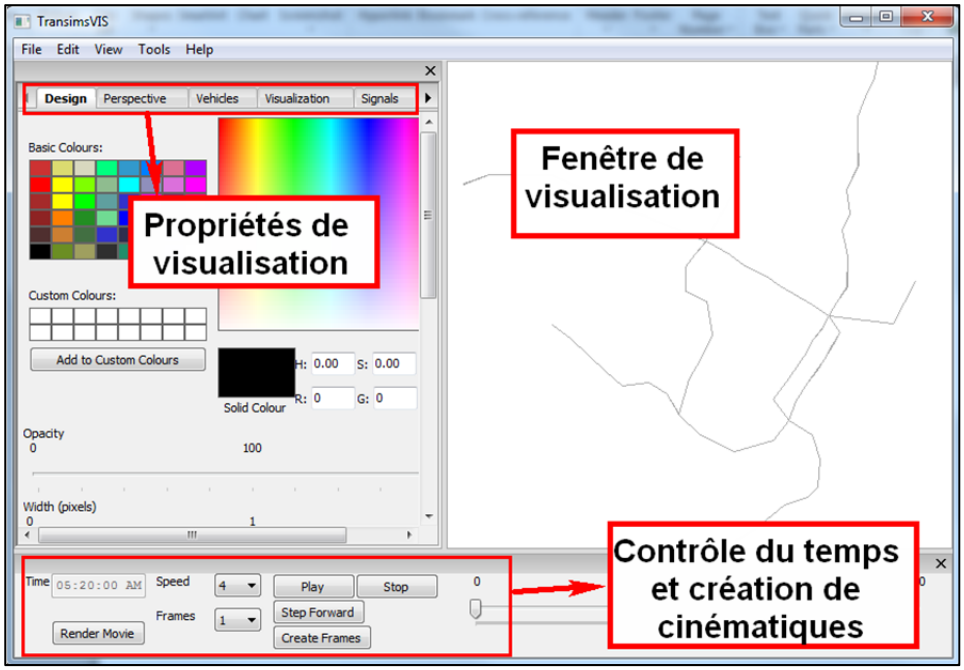


Figure 5.14 TRANSIMS Visualizer

Une des propriétés de visualisation du TRANSIMS Visualizer est la coloration des véhicules selon la charge à bord de ceux-ci. Grâce à cette fonction, il est possible d'observer l'évolution de l'achalandage dans le temps et l'espace, et ce, de façon dynamique, tel qu'illustré à la Figure 5.15. Le Visualizer possède aussi des fonctions de zoom, et de rotation tridimensionnelle, ce qui permet la capture d'images de différentes portions du réseau selon différentes perspectives. Enfin, des cinématiques peuvent être exportées pour que les simulations puissent être visualisées sans nécessairement avoir à passer par la paramétrisation du Visualizer.

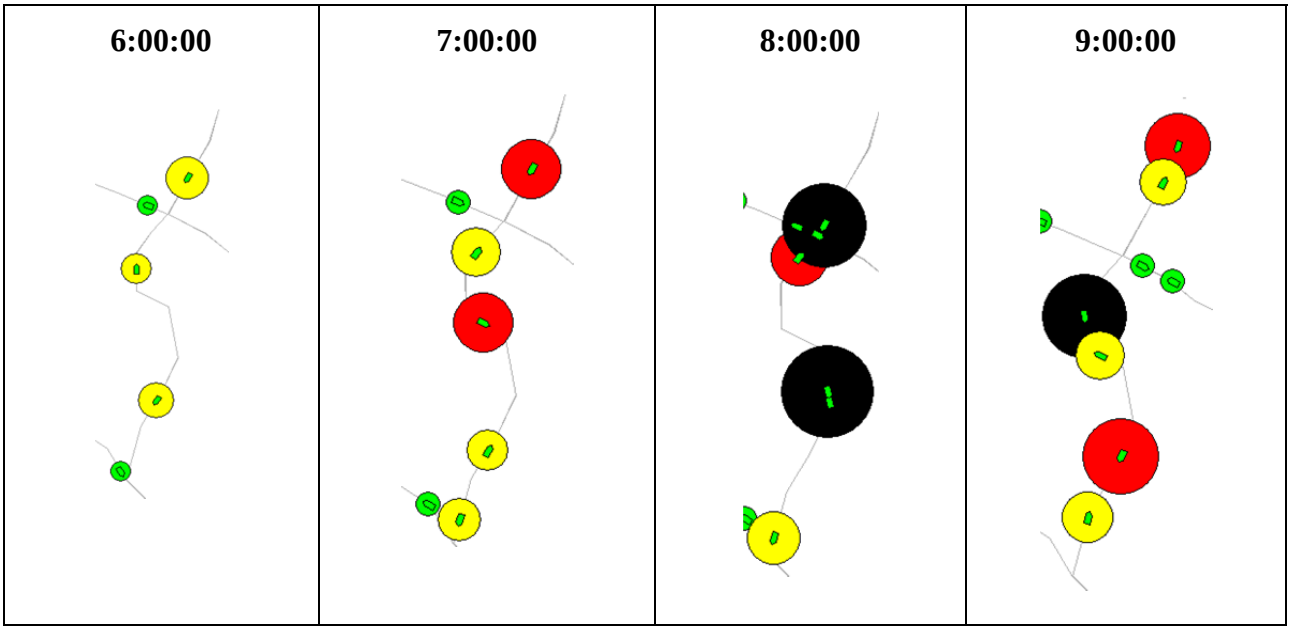


Figure 5.15 - Illustration de l'évolution du profil de charge dans le temps avec TRANSIMS

5.7 Discussion

Les expérimentations menées à l'aide de TRANSIMS constituent une base exploratoire du potentiel du logiciel à intégrer et modéliser les données de transport en commun. La démarche entreprise a d'abord permis de comprendre le fonctionnement général du logiciel, puis de déterminer la structure nécessaire des différents fichiers, et finalement d'exécuter les modules et d'observer les résultats générés.

Une des nombreuses particularités de ce logiciel est la personnalisation des procédures. En effet, il est possible de choisir quelles procédures exécuter, de modifier des procédures existantes, ou encore d'écrire ses propres procédures. Ainsi, il est plus simple de traiter d'une situation de transport en commun en n'exécutant que les procédures qui y sont liées, selon des paramètres adaptés à la situation à modéliser. Il est aussi possible de répliquer le format de fichiers requis, mais avec des données personnelles : d'où l'intégration des données de service planifié GTFS et de transactions par cartes à puce.

La flexibilité de TRANSIMS est évidente, mais la validation des résultats doit être réalisée de façon stricte vu la quantité de fichiers, de paramètres et de résultats impliqués dans une simulation, et le peu de documentation disponible. Une fois complété, le dossier de projet décrit dans ce chapitre contenait 244 fichiers répartis dans 10 dossiers, pour un total de 509 Mo. Dans le cadre de ce projet, la validation des résultats a été faite de façon sommaire, mais les résultats devront être soumis à des vérifications exhaustives si on désire implémenter TRANSIMS comme logiciel de modélisation.

CHAPITRE 6 ANALYSE DES DONNÉES D'ACHALANDAGE

La richesse des données disponibles entraîne une vaste étendue d'analyses possibles, aussi bien statiques que dynamiques. Une fois leur contenu traité et comptabilisé, les ensembles de données peuvent être utilisés pour décrire la demande spatiale et temporelle du métro. Le niveau de résolution utilisé peut varier de la journée complète, à la seconde près, et peut s'étendre de l'achalandage d'un voyage à celui sur l'ensemble du réseau.

Ainsi, les questions auxquelles ce chapitre tente de répondre sont : qui se déplace? Combien de fois un usager se déplace-t-il pendant la journée? D'où les déplacements proviennent-ils? Quelles sont les destinations principales? À quel moment de la journée les déplacements se font-ils? Combien de temps séparent deux déplacements, ou quelle est la durée des activités? Une illustration qui résume les relations entre ces questions est proposée par Klein (2007).

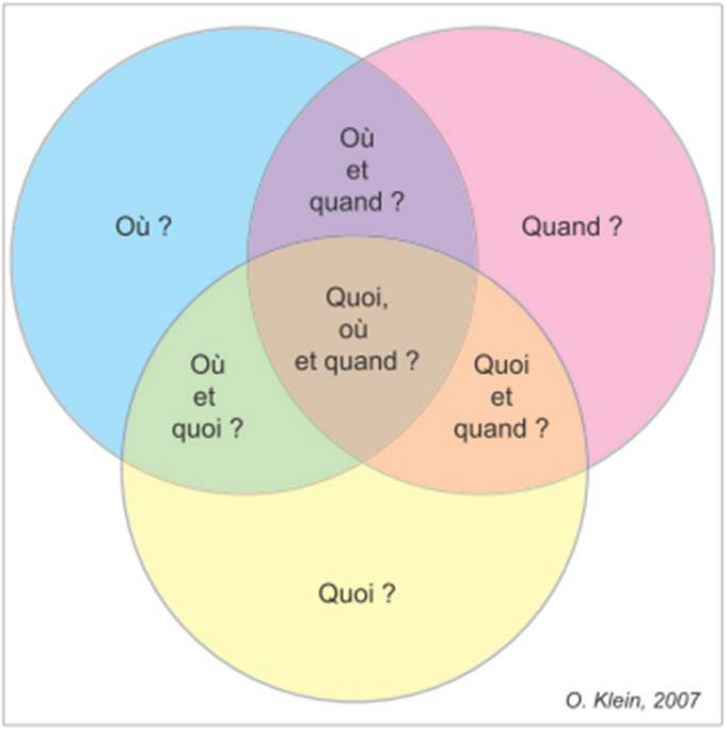


Figure 6.1 - Imbrication des requêtes spatio-temporelles (tiré de Klein, 2007)

Ce chapitre s'attarde, une fois l'échantillon de données décrit, aux analyses spatiales décrivant l'utilisation globale du réseau de métro. Ensuite, des analyses temporelles sont conduites sur des stations en particulier, afin d'observer la distribution de l'achalandage. Enfin, la combinaison des dimensions spatiales et temporelles à différents objets de la base de données sera effectuée pour des analyses variées, telles que les durées d'activités selon le nombre de déplacements, l'origine, la destination et l'heure de départ.

6.1 Description de l'échantillon

D'abord, l'achalandage des lignes et des stations sera décrit afin de connaître l'ordre de grandeur de l'utilisation de chacune d'entre elles. L'ensemble de données de transactions utilisé pour les analyses qui suivront contient 797 975 enregistrements qui contiennent : la station d'origine et l'heure de transaction, le numéro de support, et le type de tarif utilisé. La station de destination a été dérivée en suivant la méthode présentée à la section 4.2.2.2. Au total, ce sont 350 963 cartes à puce, et 68 994 tickets qui composent l'ensemble de données. Le nombre de déplacements moyen par type de support est de 2.07 par carte à puce, et 1 par ticket. Le Tableau 6.1 résume les échantillons utilisés depuis le début du projet.

Tableau 6.1 - Résumé des échantillons utilisés

Section	Description de l'échantillon	Nombre d'observations
	1. Ensemble d'origine	1 647 000
4.2.1.1	2. Ensemble 1 épuré des doublons	815 800
4.2.2.2	3. Ensemble 2 avec *code refus* = 0, donc embarquements et correspondances seulement	797 975
5.2	4. Ensemble 3 avec destinations dérivées et origine ≠ destination	645 880

Tableau 6.1 - Résumé des échantillons utilisés (suite)

5.4.2	5. Nombre de déplacements simulés avec succès dans TRANSIMS à partir de l'ensemble 4.	645 125
6.4.1	6. Ensemble 4 avec support cartes à puce seulement et titres validés de 2-4 fois.	643 080
*Les autres tailles d'échantillons mentionnés sont des approximations.		

6.2 Profil par ligne

La base de données permet une panoplie d'analyses qui peuvent être conduites selon les besoins. Celles-ci peuvent se pencher sur l'étude d'un voyage, d'un véhicule, d'une station, d'une ligne, et peuvent comporter un pas de temps qui passe de la seconde à une journée entière. La Figure 6.2 constitue un repère qui a pour but de pouvoir identifier les différentes possibilités qui s'offrent en termes d'étendue d'analyse, et de niveau d'agrégation des données.

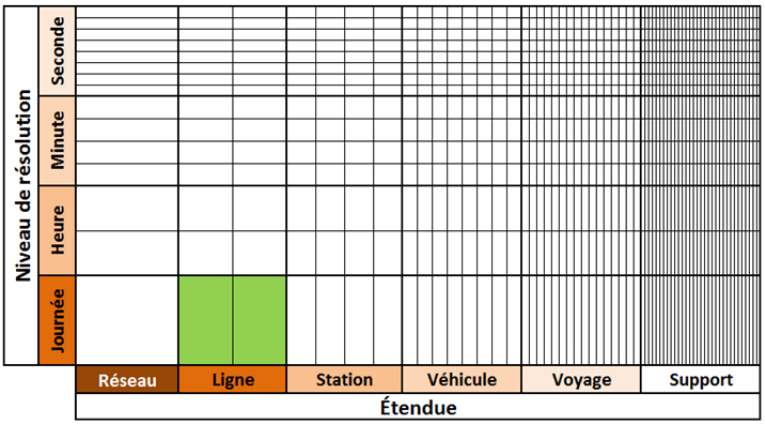


Figure 6.2 - Niveau de résolution et étendue de l'analyse par ligne

La caractérisation de l'utilisation des lignes de métro peut être représentée selon une matrice origine-destination. En effet, cette matrice permet d'identifier les lignes les plus achalandées, de

même que les types de correspondances les plus fréquents. C'est ce qui est illustré au Tableau 6.2.

Tableau 6.2 Matrice origine-destination par ligne

	DESTINATION					
ORIGINE	**1**	**2**	**4**	**5**	**ND**	**Total**
1	176 295	83 989	12 924	10 894	58 882	**342 984**
2	84 655	160 126	8 217	28 891	63 902	**345 791**
4	12 767	8 292	412	1 405	4 578	**27 454**
5	10 506	29 445	1 296	27 483	13 016	**81 746**
Total	**284 223**	**281 852**	**22 849**	**68 673**	**140 378**	**797 975**

Ce tableau nous informe du nombre de transactions par ligne, du nombre de correspondances, et du nombre de transactions dont la destination n'a pu être dérivée. Chaque station est numérotée de façon à correspondre à une ligne, par exemple : la station Berri-UQAM est numérotée 146, donc appartient à la ligne 1. Toutefois, cette numérotation entraîne une surestimation de l'utilisation de certaines lignes au détriment des autres puisqu'une personne peut embarquer à Berri-UQAM et faire un déplacement sur la ligne orange. La méthode utilisée ici est sommaire, et permet de déterminer l'importance relative de chaque ligne sur l'achalandage total du réseau de métro. Évidemment, le nombre de transactions par ligne dépend de plusieurs facteurs, dont le nombre de stations par ligne, et la présence de générateurs de déplacements à proximité des stations.

6.3 Profil par station

En second lieu, une fois les lignes décrites de façon sommaire, le niveau de résolution peut être raffiné au niveau de la station. L'achalandage à une station dépend de la nature de celle-ci, c'est-à-dire du type d'infrastructure qui l'entoure. Une station à proximité d'une zone résidentielle, par exemple, comportera une période de pointe matinale, tandis qu'une station près d'un lieu de travail génèrera une pointe en soirée. Chaque station peut être analysée selon l'heure de la

journée, et selon le type d'usager le plus fréquent à la station. Les 10 stations les plus fréquentées pour la journée du 17 mars 2010 sont présentées au Tableau 6.3.

Tableau 6.3 - Les 10 stations les plus fréquentées le 17 mars 2010

RANG	CODE	STATION	NB TRANSAC	%
1	146	Berri-UQAM	41 840	5.2
2	140	McGill	36 504	4.6
3	136	Guy-Concordia	28 787	3.6
4	454	Longueuil	26 054	3.3
5	252	Bonaventure	25 644	3.2
6	222	Côte-Vertu	23 518	3.0
7	138	Peel	21 888	2.7
8	142	Place-des-Arts	21 519	2.7
9	242	Vendôme	18 999	2.4
10	280	Henri-Bourassa	18 999	2.4
		Total	263 752	33

Cette classification peut être jumelée au réseau de métro lui-même, afin d'apprécier visuellement l'utilisation des stations. C'est ce qui est illustré à la Figure 6.3.

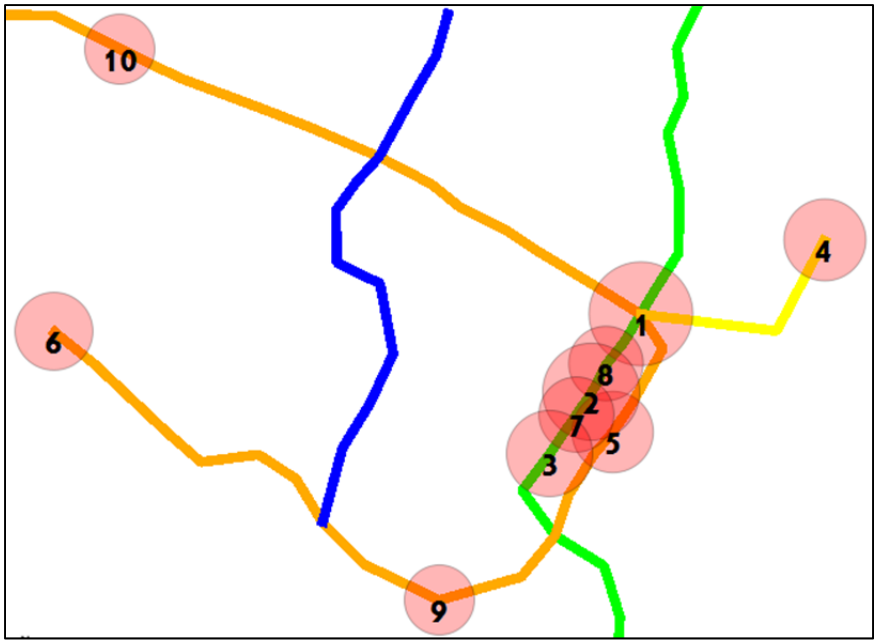


Figure 6.3 – Les 10 stations les plus fréquentées

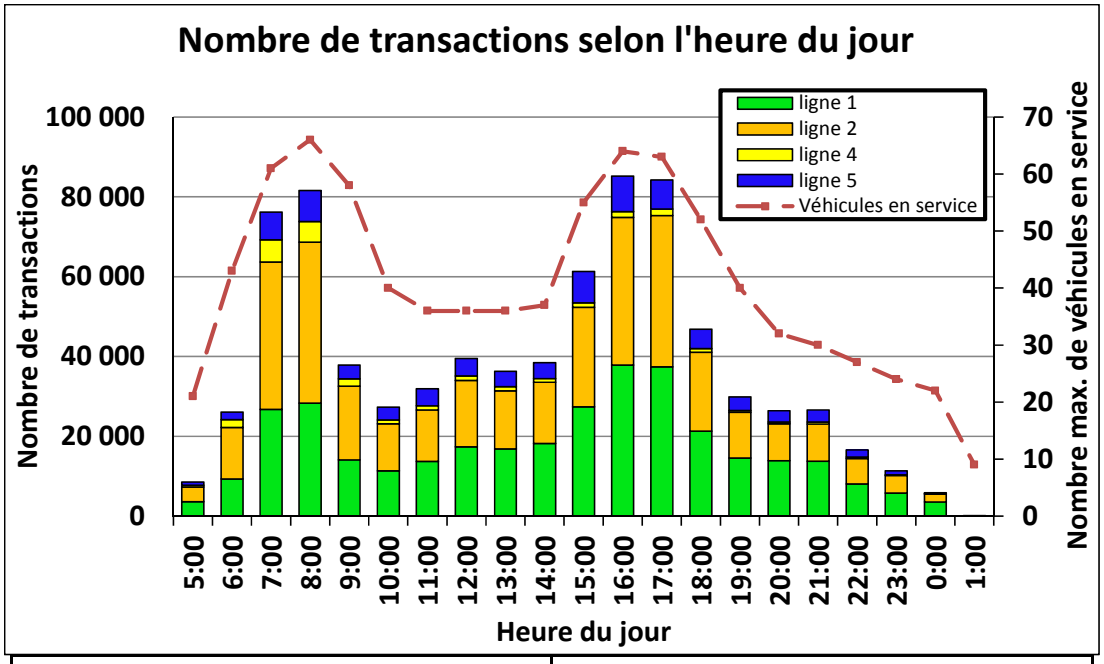
À titre d'information, les 10 stations les moins achalandées pour la journée du 17 mars 2010 sont compilées au Tableau 6.4.

Tableau 6.4 – Les 10 stations les moins fréquentées le 17 mars 2010

RANG	CODE	STATION	NB TRANSAC	%
59	154	Préfontaine	4426	0.6
60	554	d'Iberville	4278	0.5
61	162	l'Assomption	4231	0.5
62	120	Monk	3981	0.5
63	286	de la Concorde	3897	0.5
64	148	Beaudry	3790	0.5
65	542	Acadie	3292	0.4
66	228	de la Savane	2848	0.4
67	248	Georges-Vanier	2539	0.3
68	452	Jean-Drapeau	1400	0.2
		Total	34682	4

6.4 Distribution temporelle des transactions

La distribution temporelle des transactions permet d'identifier la variabilité de l'achalandage au courant de la journée. D'abord, les données peuvent être groupées par heure, afin de voir le profil général de la demande temporelle pour une journée de semaine. En y jumelant les données de service planifié, on peut s'assurer de la conformité de la planification aux différentes périodes d'achalandage de la journée. C'est ce qui est illustré à la Figure 6.4. La figure démontre la présence de périodes d'achalandage plus élevé le matin entre 7 h et 9 h, et le soir, entre 15 h et 18 h. Entre ces deux périodes de pointe, on retrouve un nombre de transactions assez semblable d'une heure à l'autre. Le nombre de transactions maximum est de 85 183 entre 16 h et 17 h, ce qui correspond à environ 1300 transactions/station/heure, ou 21 transactions/station/minute. Comme le taux d'utilisation aux stations n'est jamais le même partout, on observe en réalité un taux plus élevé aux stations centrales, et un taux plus faible aux stations de périphérie.



Pointe AM (7:00 – 9:00)	20% des transactions
Pointe PM (15:00 – 18:00)	30% des transactions

Figure 6.4 - Nombre de transactions selon l'heure du jour

Le Tableau 6.5 présente le nombre de transactions par heure par ligne. Les périodes de pointe s'étendent de 7 h à 9 h le matin, et de 15 h à 18 h le soir, et contribuent pour 20% et 30% respectivement du total de transactions pour la journée. De plus le nombre maximal de transactions sur une ligne en une heure est de 40 374, sur la ligne 2 entre 8 h et 9 h. Enfin, le nombre maximal de transactions pour une période d'une heure dans tout le réseau est de 85 183, entre 16 h et 17 h.

Tableau 6.5 - Nombre de transactions par heure sur chaque ligne

Heure/Ligne	1	2	4	5	TOTAL
5	3 593	3 669	496	756	8 514
6	9 320	12 914	1 931	1 921	26 086
7	26 718	36 957	5 537	6 935	76 147
8	28 279	40 374	5 153	7 765	81 571
9	14 060	18 467	1 878	3 479	37 884
10	11 360	11 771	992	3 197	27 320
11	13 712	12 874	1 022	4 301	31 909
12	17 361	16 629	1 129	4 395	39 514
13	16 850	14 515	1 066	3 856	36 287
14	18 203	15 297	922	4 054	38 476
15	27 384	24 958	1 120	7 862	61 324
16	37 832	37 005	1 424	8 922	85 183
17	37 413	37 887	1 658	7 278	84 236
18	21 276	19 787	940	4 839	46 842
19	14 536	11 451	494	3 409	29 890
20	13 913	9 222	448	2 843	26 426
21	13 755	9 295	499	3 021	26 570
22	8 059	6 354	413	1 744	16 570
23	5 772	4 376	219	960	11 327
24	3 558	1 950	112	209	5 829
25	30	39	1	0	70
Total	**342 984**	**345 791**	**27 454**	**81 746**	**797 975**
Max	37 832	40 374	5 537	8 922	85 183
Moyenne	16 333	16 466	1 307	3 893	37 999
Écart-Type	10 444	12 316	1 448	2 613	25 964

En illustrant le profil de charge pour différentes parties de la journée, on obtient les résultats présentés à la Figure 6.5. L'échantillon utilisé pour générer les profils de charge tient seulement compte des déplacements pour lesquels une destination a pu être dérivée. Les proportions du nombre de déplacements total pour les 2 périodes modélisées sont très semblables à la somme des déplacements illustrés dans le Tableau 6.5 (1% de différence).

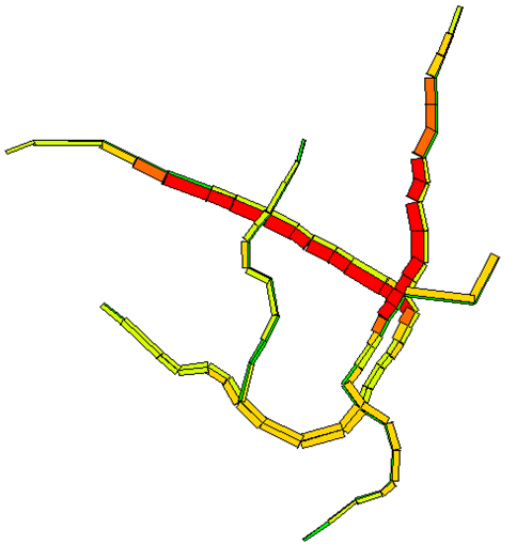
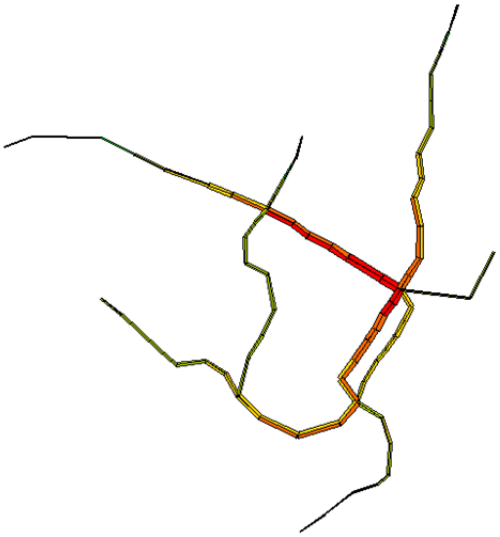
	Heures modélisées : **7 h – 8 h** Nombre de déplacements modélisés : **58 503** Passagers-heures : **13 124** Temps moyen de déplacement : **13 minutes**
	Heures modélisées : **13 h – 14 h** Nombre de déplacements modélisés : **28 282** Passagers-heures : **5 020** Temps moyen de déplacement : **10 minutes**

Figure 6.5 Profil de charge à différents moments de la journée

6.4.1 Distribution temporelle selon le numéro de transaction

Une personne qui désire pratiquer une activité doit d'abord se rendre à celle-ci, puis, une fois l'activité terminée, revient à son domicile, ou pratique une autre activité. Ainsi, toute activité pratiquée implique au moins 2 déplacements. Le nombre d'activités pratiquées dans une journée dictera le nombre de déplacements effectués. La Figure 6.6 illustre la distribution des transactions

dans la journée en fonction du numéro de séquence de la transaction au cours de cette même journée. Seules les transactions effectuées par cartes à puce ont été retenues, considérant que la majorité des supports papier sont utilisés pour faire une seule transaction. De plus, puisqu'on désire illustrer la distribution des déplacements selon le numéro de séquence de la transaction, il est plus représentatif d'utiliser seulement les numéros de support utilisés pour effectuer 2 déplacements et plus. Le sous-ensemble ainsi créé contient 643 080 transactions.

Pour les usagers qui se déplacent deux fois ou plus, la figure montre que le premier déplacement de la journée est généralement fait en pointe matinale, puis distribué temporellement de façon décroissante sur le restant de la journée. Le second déplacement se fait graduellement à partir de 7:00, et atteint son maximum entre 16:00 et 17:00. Enfin, les 3e et 4e déplacements se font à partir de 10:00, et s'étendent jusqu'à minuit. On retrouve 276 926 premiers embarquements, 277 172 deuxièmes, 63 144 troisièmes, et 25 838 quatrièmes. On explique le nombre plus élevé de secondes transactions par le fait que seuls les numéros de carte ayant effectué 2 transactions et plus ont été retenus, et que certaines transactions ont été refusées par le système.

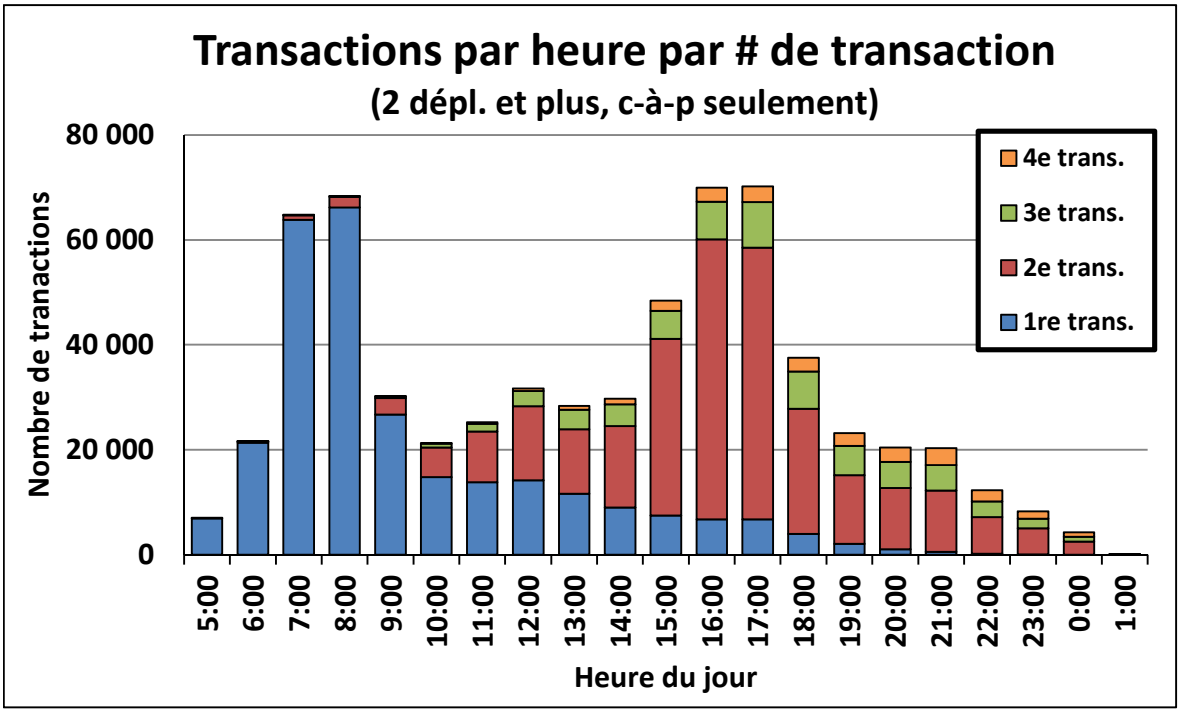


Figure 6.6 - Transactions par heure selon le numéro de transaction

Cette distribution donne un premier aperçu des patrons de déplacements effectués par les usagers. Puisqu'il n'y a pas d'information liée au motif de déplacement dans la base de données, celui-ci doit être supposé en posant différentes hypothèses. À ce titre, une méthode est proposée et appliquée par Devillaine, Munizaga et Trépanier (2012), aux villes de Santiago et Gatineau. La méthode requiert, entre autres, le calcul de la durée des activités pratiquées.

6.4.1.1 Durée des activités

Dans le cas des transactions effectuées dans le métro, contrairement à celles effectuées dans les bus, il n'y a pas de confusion possible entre une activité et une correspondance. En effet, la carte est validée à l'entrée dans le réseau seulement : il n'est pas nécessaire de la revalider lors d'un changement de ligne. On peut donc affirmer que toute activité est bornée par 2 transactions. Le temps qui sépare deux transactions successives d'un même numéro de carte correspond à la durée consacrée à une activité, et cette durée inclue le temps de parcours de la station d'origine à la destination.

Pour déterminer la durée d'une activité, il faut qu'un même numéro de support ait été utilisé pour compléter au moins 2 déplacements. Le calcul des durées d'activités est basé sur le numéro de séquence de transaction, et sur les heures auxquelles se font 2 transactions successives. Une fois la base de données triée selon le numéro de carte et l'heure de transaction, le calcul est le suivant : $t = h_{n+1} - h_n$, où t = durée de l'activité, h = heure de la transaction et n = numéro de séquence de la transaction. Si une transaction est refusée à une station, et par la suite acceptée à cette même station, le temps entre les deux transactions doit être ignoré, afin de ne pas considérer le temps pris à régler un problème d'équipement comme une activité.

Les durées d'activités en fonction du numéro de séquence des activités sont illustrées à la Figure 6.7. Celle-ci rend compte de deux caractéristiques communes aux durées d'activités. D'abord, plus une activité commence tard, plus sa durée est courte. De ce fait, il est possible d'affirmer que les 2e et 3e activités de la journée auront aussi une durée plus courte que la 1re, puisqu'elles sont nécessairement réalisées plus tard que celle-ci. Par ailleurs, il semble qu'en moyenne, les 2e et

3e activités de la journée ont pratiquement la même durée. Afin de comprendre pourquoi certaines activités sont plus longues que d'autres, ce qui revient à déterminer les motifs de déplacements, il pourrait être intéressant de coupler les destinations et les types de titres utilisés aux durées d'activités.

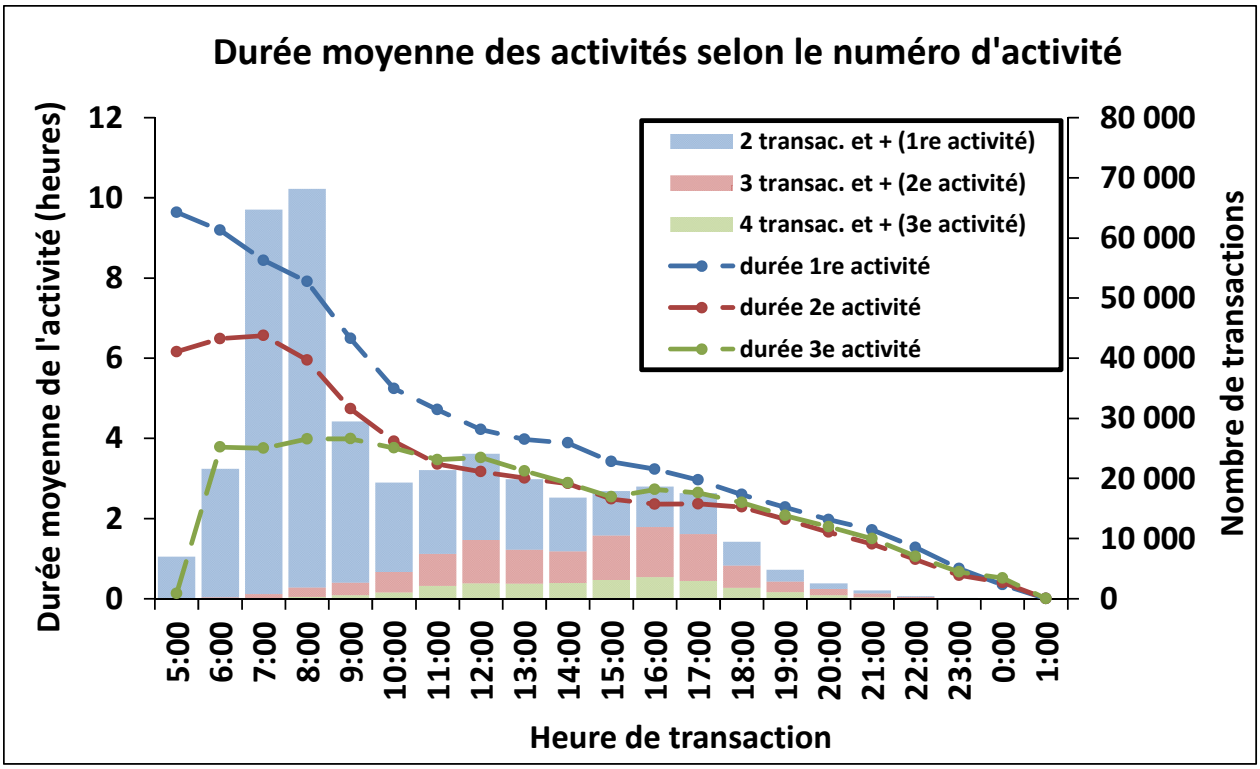


Figure 6.7 - Durée des activités selon le numéro de séquence de l'activité

6.4.2 Nombre de transactions par heure/origine

Dans un autre ordre d'idées, il est possible d'analyser l'achalandage d'une seule station pour chaque heure de la journée. On peut s'intéresser aux transactions à la station d'origine, ou bien à la station de destination dérivée. Ce genre d'analyse permet d'identifier le type de générateur de déplacements à proximité de la station. En effet, si la station possède une période de pointe matinale, alors on peut supposer qu'elle se retrouve dans une zone résidentielle, d'où les usagers amorcent leur déplacement le matin pour se rendre au travail ou à l'école. De la même façon, si la

station possède une pointe en soirée, c'est qu'elle est située près d'un lieu de travail, d'une école, ou d'une zone commerciale.

La station la plus utilisée du réseau, soit Berri-UQAM, compte environ 41 000 transactions quotidiennes. La répartition de celles-ci à chaque heure du jour est illustrée à la Figure 6.8. Ces transactions ont donc comme origine la station Berri-UQAM.

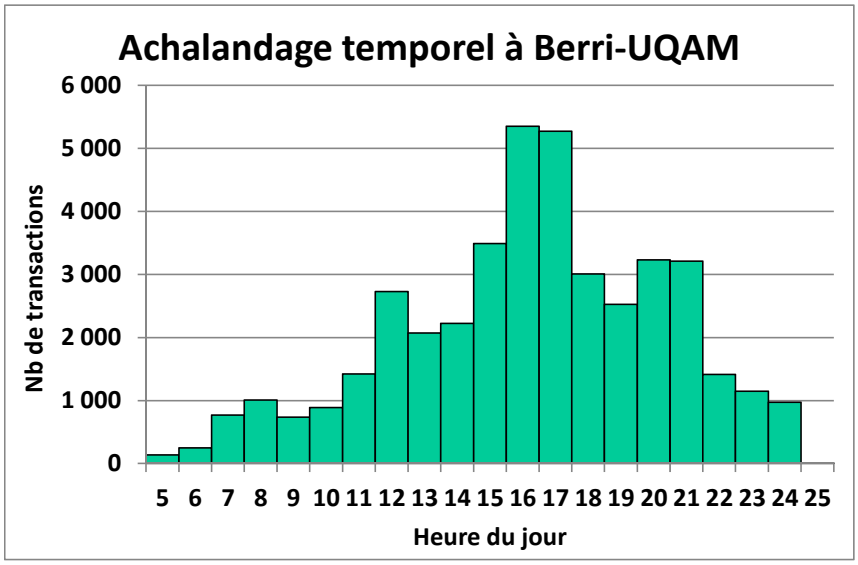


Figure 6.8 - Achalandage temporel à Berri-UQAM

La figure ci-dessus montre une période de pointe du soir débutant à 15 h, ce qui nous permet d'affirmer que cette station est un générateur de déplacements. Donc, puisque Berri-UQAM est une destination en début de journée, une question survient quant à la provenance des usagers qui s'y destinent. La Figure 6.9 présente le profile d'achalandage temporel de quelques stations d'où proviennent les usagers en direction de la station Berri-UQAM, pour les transactions dont la destination a pu être dérivée. En excluant les transactions dont la station d'origine est la même que la station de destination, on retrouve au total 34 379 transactions dont la destination est Berri-UQAM pour la journée du 17 mars 2010. D'entre elles, 3257 proviennent de la station Longueuil-Université de Sherbrooke, 1400 de la station Bonaventure et 1328 de la station McGill.

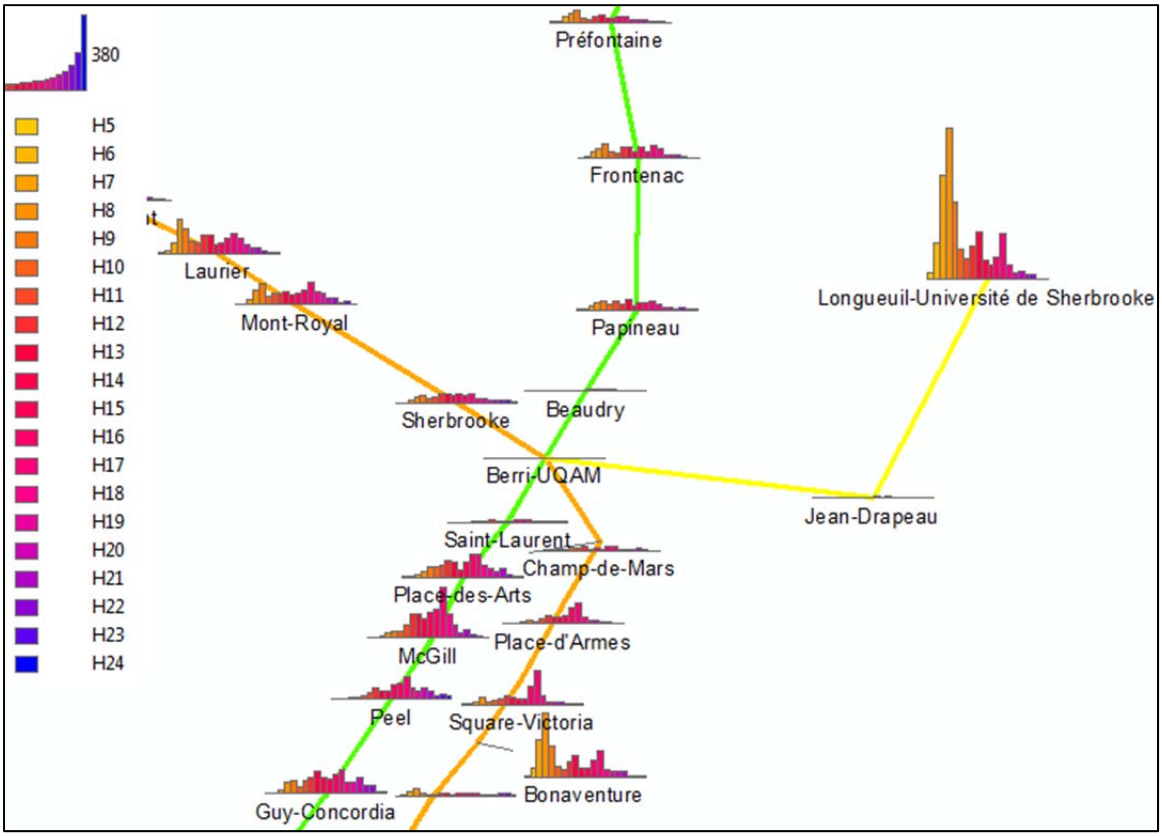


Figure 6.9 - Provenance des usagers dont la destination est Berri-UQAM selon l'heure du jour

6.4.3 Distribution des transactions à la seconde

Les données de transactions par carte à puce sont enregistrées aux tourniquets à la seconde près. L'utilisation des tourniquets varie entre 2 et 6 418 transactions de validation pour une journée complète de service. L'achalandage peut être illustré à une station particulière à chaque seconde de la journée, tel que présenté à la Figure 6.10. La figure reprend les données de la station Vendôme, située sur la portion ouest de la ligne 2. Cette station possède 8 tourniquets, c'est pourquoi il est possible que 7 transactions soient effectuées au même moment. Le profil de transactions illustré est typique, avec une période de pointe du matin plus forte, étant donné que cette station est en périphérie du centre-ville.

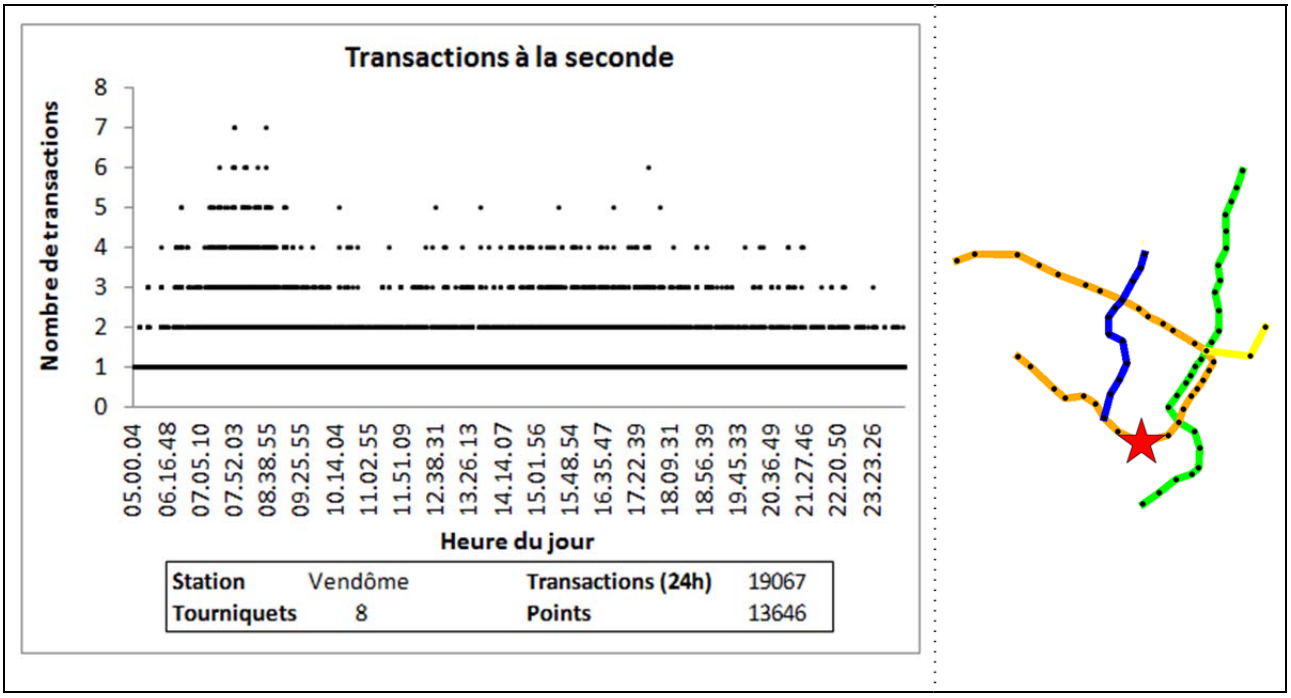


Figure 6.10 - Transactions à la seconde à la station Vendôme

En dressant le même genre de profil pour chacune des stations du réseau, on observe certaines situations particulières. Par exemple, à la station Henri-Bourassa (Figure 6.11), on atteint un total de 70 transactions en 1 seconde aux environs de 8 h. Puisqu'on compte seulement 11 tourniquets à cette station, il est impossible que la situation illustrée par les données rende compte de la situation réelle. En agrégeant les données par tourniquet pour cette station, on réalise que les données récoltées par le tourniquet identifié « 32PVM0280MZ0301E101 » sont groupées par tranches de 15 minutes.

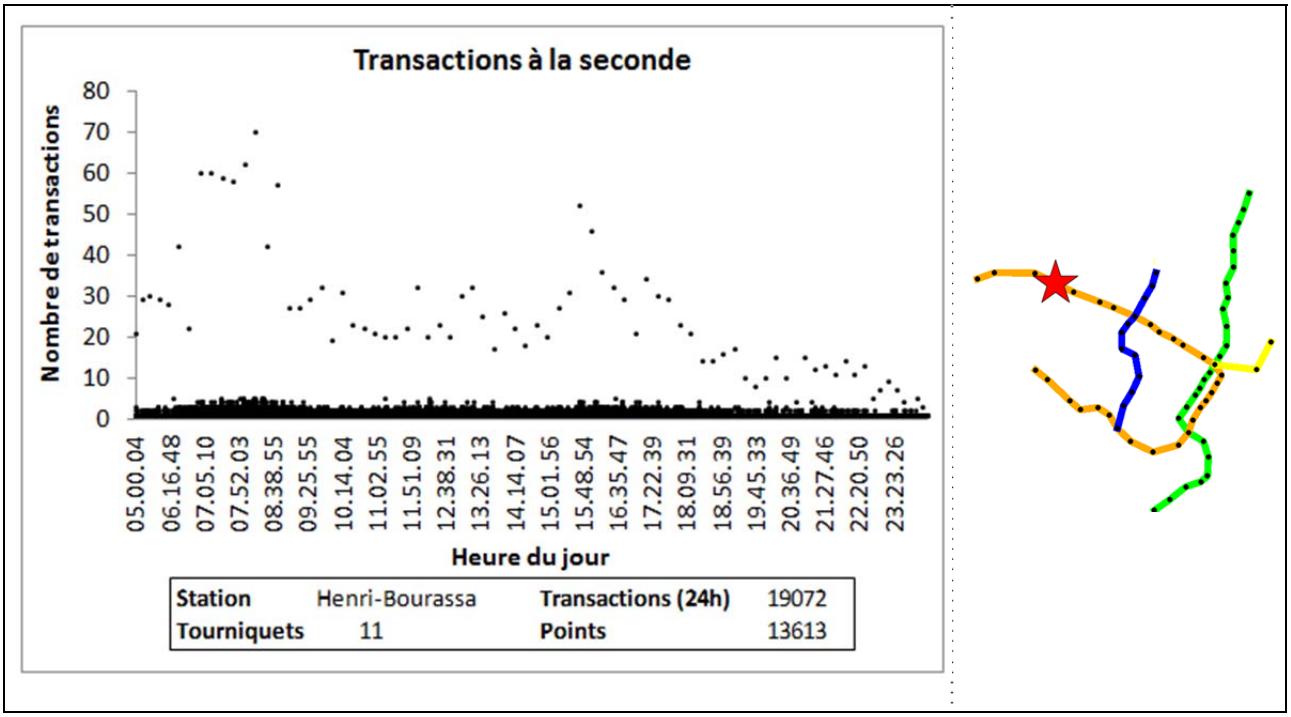


Figure 6.11 - Transactions à la seconde à la station Henri-Bourassa

Le même genre de problématique est rencontré avec le profil temporel de la station Crémazie, mais cette fois, ce sont 2 tourniquets pour lesquels les données sont groupées par tranches de 15 minutes. Le profil de transaction résultant est illustré à la Figure 6.12. Encore une fois, un total dépassant les 20 transactions en une seconde ne peut pas correspondre à la situation réelle, car la station Crémazie compte seulement 7 tourniquets.

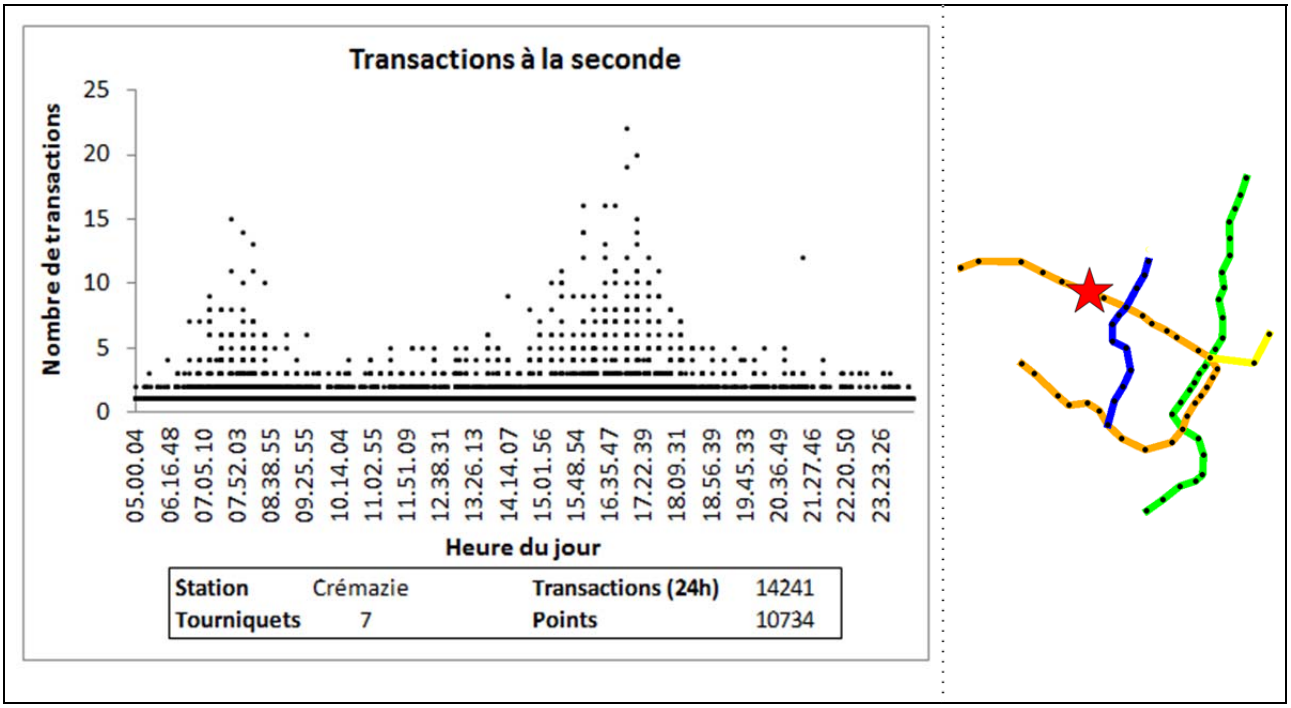


Figure 6.12 - Transactions à la seconde à la station Crémazie

L'analyse des données à la seconde permet ainsi de détecter des ensembles de données qui présentent un caractère particulier, et d'éviter la propagation d'erreurs dans d'autres analyses.

6.4.3.1 Redistribution des données groupées

Les exemples présentés ci-dessus montrent qu'il serait donc impossible de représenter précisément la charge à bord des rames à partir des données des tourniquets problématiques. En effet, les données sont groupées par tranches de 15 minutes, et il peut passer près de 5 rames dans cet intervalle. Pour corriger la situation, les transactions groupées doivent être redistribuées sur l'intervalle de temps.

Les données sont groupées par tranches de 15 minutes. En prenant par exemple une tranche de 15 minutes qui présente au total 70 transactions, il faut redistribuer celles-ci selon un intervalle de 4,67 transactions/minute, soit 1 transaction aux 13 secondes. Il faut donc lister l'heure du jour à la seconde, et identifier une transaction toutes les 13 secondes. On passe ensuite à la prochaine

tranche de 15 minutes et on répète l'opération. Le profil peut-être ainsi dressé pour toute la journée.

En utilisant cette méthode sur les données de la station Crémazie, on obtient le profil illustré à la Figure 6.13. Le total de transactions obtenu par redistribution est légèrement inférieur au total réel (différence d'environ 1%), compte tenu de l'arrondissement des intervalles et des fréquences lors des calculs. La Figure 6.13 diffère de la précédente en deux aspects : les transactions apparaissent en continu au lieu d'être groupées, et le maximum de transactions en une seconde est de 5 au lieu de 20. Le profil ressemble davantage à celui de la station Vendôme (Figure 6.10), qui lui, ne comportait aucune donnée groupée.

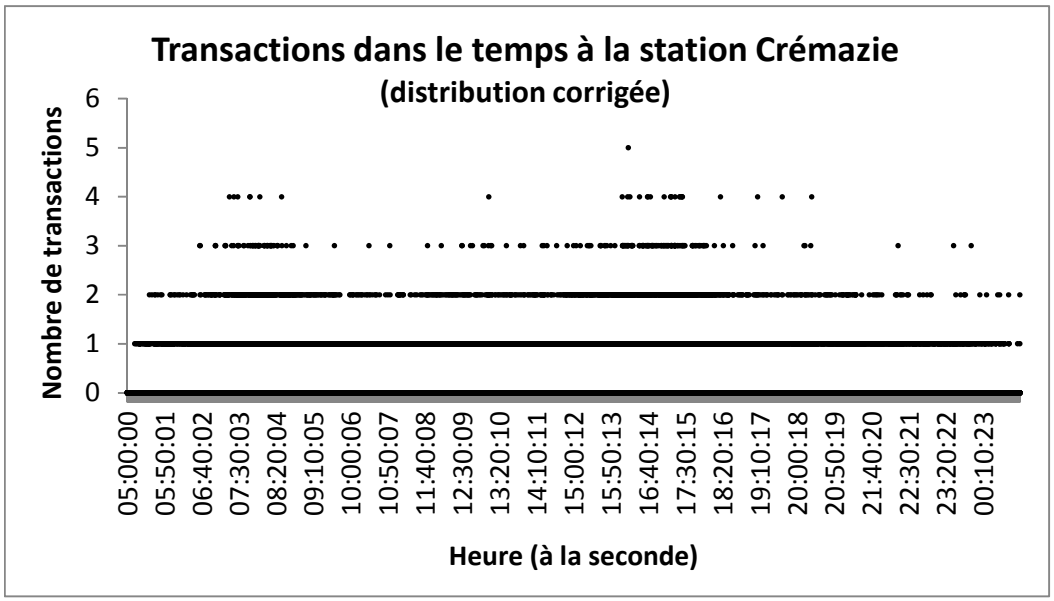


Figure 6.13 - Distribution corrigée à la station Crémazie

6.5 Discussion

Ce chapitre présentait quelques pistes d'analyses rendues possibles grâce aux technologies d'acquisition de données. Cette thématique porte davantage sur l'identification des besoins, et sur la méthode déployée pour combler ces besoins. De ce fait, le choix des données, la structure des

données, les relations entre les différentes bases de données disponibles, et les outils utilisés sont les aspects fondamentaux de la construction d'analyses.

Par ailleurs, avant même de passer à l'analyse des données, il faut s'assurer que celles-ci présentent bien le caractère qu'on leur attribue. Par exemple, la distribution des données à la seconde a permis de se rendre compte que certaines stations agrégeaient leurs ensembles de données aux 15 minutes. Cette particularité a un impact sur l'achalandage des rames à la seconde : celui-ci est surestimé toutes les 15 minutes, à moins que les distributions aux stations problématiques ne soient corrigées.

Finalement, la connaissance des objets en transports en commun, et plus particulièrement dans le cas du métro, permet de valider les résultats obtenus. La validité des résultats est limitée d'une part par l'intégrité des données, et, d'autre part, par la capacité à dériver une station de destination pour chaque transaction. La connaissance des ordres de grandeur liée à chaque objet du réseau de métro et de la logique du processus de transaction nous assurent de pouvoir reconnaître les erreurs à même un ensemble de données, ou une analyse.

CHAPITRE 7 CONCLUSION

7.1 Résumé des thématiques abordées

La première étape de ce projet consistait à formuler un rappel des concepts liés aux différentes composantes d'un réseau de transport en commun, en mettant l'accent sur un réseau de métro. Les notions de systèmes de transport intelligents et leurs implications ont été abordées. Ensuite, le rôle de la carte à puce en transport en commun a été défini, et ses bénéfices potentiels d'utilisation ont été identifiés. Le traitement de l'information récoltée et la méthode de dérivation d'information supplémentaire ont été présentés. Puis, la contribution des technologies de l'information à l'analyse de données en transport a été explorée. La pertinence de l'intégration de différentes sources de données de transport aux SIG aux fins d'analyse a été illustrée. Aussi, un aperçu des données GTFS a permis la familiarisation avec ce format standard développé par Google. Enfin, l'utilité de la modélisation en transport a été cernée dans le cadre du projet actuel.

La deuxième étape était la description du métro de Montréal : historique, composantes, planification, coûts, pannes et expérimentations effectuées sur le métro. Ce chapitre a permis d'apprécier les ordres de grandeur en lien avec la demande, l'offre et l'étendue du service du réseau de métro de Montréal.

Une fois les notions de base introduites et le contexte actuel du métro de Montréal présenté, le projet se penchait sur la gestion des données archivées en transport. L'ensemble de données de transactions est d'abord passé par des procédures d'identification et de description des objets. La cohérence des valeurs observées a été analysée, et la présence de cas particuliers a été soulignée. Par la suite, des procédures de traitements ont été programmées et appliquées sur l'ensemble de données de transactions afin d'en valider l'intégrité, l'uniformité et l'utilité. Puis, des procédures d'enrichissement ont été appliquées aux données. Celles-ci ont permis d'ajouter de l'information pertinente telle que le nombre de transactions par usager, et le point de débarquement associé à un déplacement. De plus, un portrait spatio-temporel des supports dont on ne possède qu'une seule transaction pour la journée a été dressé. Celui-ci permettra éventuellement de déterminer

une liste de critères pour la modélisation des stations de destinations des supports à transaction unique. Ensuite, dans un cadre expérimental, une méthode permettant de raffiner le niveau de résolution des données GTFS a été proposée, permettant de passer d'un niveau de précision minute à seconde. Enfin, quelques applications issues du croisement entre les données de service planifié à la seconde et les données de transactions ont été illustrées.

Le chapitre 5 était une introduction au logiciel de modélisation TRANSIMS. La quantité et la finesse du niveau de résolution des données disponibles donnaient lieu de croire que la modélisation du métro était possible à la seconde près. Ainsi, un historique du logiciel a d'abord été dressé, suivi d'une schématisation de son fonctionnement général, et de la présentation d'études de cas menées à l'aide du logiciel, afin d'évaluer le potentiel de l'utilisation de TRANSIMS. Puis, une description sommaire de l'expérimentation à réaliser avec le logiciel a été exposée. Les modules ont ensuite été analysés individuellement : codification, affectation, microsimulation et visualisation. Les intrants, extrants, paramètres, contraintes et formats pour chacun des modules ont été étudiés. Les résultats de la microsimulation ont été comparés aux résultats obtenus de façon manuelle au chapitre 4. En dernier lieu, un aperçu du module d'animation a été présenté en passant par son fonctionnement, les possibilités offertes, et un bref exemple.

Finalement, le chapitre 6 proposait différentes pistes d'analyses exploratoires des données traitées et enrichies. L'échantillon analysé a d'abord été décrit, et les analyses ont été conduites en illustrant un point de repère spatial et temporel pour chacune d'elles. Les thèmes d'analyses abordés étaient l'achalandage par ligne, par station, les durées d'activités, et la distribution des transactions à la seconde. Les résultats ont été présentés sous forme de graphiques, de cartes, et de tableaux.

7.2 Réflexivité et perspectives

« Plus généralement, une démarche réflexive en science consiste en une prise de conscience et en un examen approfondi de sa propre démarche scientifique. » (Wikipédia)

L'objet de ce projet était l'analyse exploratoire du potentiel théorique de l'assemblage de données archivées en transport en commun. En cours de réalisation, il est devenu évident que la grande problématique d'un tel projet est d'abord la rigueur de l'organisation, c'est-à-dire la définition du problème, le choix des données à utiliser, le format et la disposition des données pour les moyens techniques employés, et le niveau de résolution du résultat visé. Ainsi, la recherche s'est concentrée sur ces aspects critiques plutôt que sur les résultats eux-mêmes. Cela a permis d'établir une méthodologie générale, qui elle, peut être reprise dans tout contexte d'analyse de données archivées.

Par ailleurs, les données ont été enrichies à différents niveaux. D'abord, de l'information a été ajoutée à la base de données de transactions, notamment les points de débarquement des usagers. Pour y arriver, certaines hypothèses, proposées et vérifiées dans la littérature, ont été respectées. Celles-ci ont permis de dériver un certain nombre de destinations, à un niveau donné d'incertitude. Certaines limitations de cette démarche auraient pu être évitées en analysant un échantillon qui s'étend sur plus d'une journée. En effet, il aurait été possible de confirmer certaines hypothèses, dont le retour à l'origine lors d'une journée de plus de 2 déplacements. Cependant, tel que mentionné précédemment, le projet était axé davantage sur la méthode que sur les résultats. De ce fait, la conscience de ce genre de limitation s'avère indispensable, mais n'a pas d'impact sur l'issue du projet.

Puis, vu la disponibilité d'ensembles de données de service planifié et offert (GTFS et cantons respectivement), une autre démarche d'enrichissement a été suggérée, soit le raffinement des données de service planifié GTFS à la seconde. Un tel raffinement engendre une description systématique du service, et respecte la logique du déplacement des véhicules sur le réseau. De

plus, un logiciel comme TRANSIMS admet des ensembles de données aussi précises en matière de référence temporelle, ce qui permet d'obtenir des résultats plus fidèles à une situation réelle.

Enfin, plusieurs questions peuvent être soulevées quant à l'utilisation du logiciel TRANSIMS, à commencer par le choix de ce logiciel plutôt qu'un autre. TRANSIMS est un logiciel connu depuis les années '90, mais ce sont les développements les plus récents au niveau de l'interface qui ont permis une utilisation plus conviviale du logiciel. De plus, c'est un logiciel ouvert, programmable, qui fonctionne avec des fichiers intrants et extrants en format texte. Il est possible d'y intégrer des ensembles de données aussi denses que les transactions par cartes à puce, et les données de service planifié. Aussi, le logiciel a été développé aux États-Unis, où la priorité en transport est axée sur l'automobile. De ce fait, la portion qui touche le transport en commun dans TRANSIMS n'a pas été pleinement approfondie. Finalement, le logiciel a été choisi compte tenu du fait qu'il existe une documentation complète dédiée à son utilisation. Toutefois, cette documentation est excessivement lourde, et se doit d'être résumée et illustrée à des fins de reproductibilité.

Un autre aspect de l'utilisation de TRANSIMS qui pourrait être remis en question est le choix de la méthode de modélisation. D'abord, la codification a été faite suivant le modèle d'un réseau de transport en commun : arrêts, lignes et horaires. De plus, les paramètres ont été définis afin de représenter le plus fidèlement possible le métro de Montréal. Enfin, étant donné qu'il y avait peu d'exemples traitant du transport en commun dans la littérature, une méthode de modélisation du métro a été proposée de façon à profiter du potentiel des données archivées.

L'introduction de TRANSIMS dans ce projet avait pour but d'explorer les possibilités d'intégration des données pour le réseau de métro. En définitive, il serait souhaitable d'arriver à modéliser le réseau de transport en entier, incluant l'auto et les bus. Toutefois, il faudrait s'assurer de maîtriser la procédure de modélisation pour chacun de ces modes avant de les fusionne. Pour cela, il faudrait d'une part arriver à programmer les procédures de formatage des données dans une forme compatible à TRANSIMS, pour éviter la répétition de manipulations

fastidieuses. D'autre part, des procédures de validation automatisées devraient aussi être développées, afin de vérifier plus rapidement les résultats générés par TRANSIMS.

Dans un autre ordre d'idées, les analyses réalisées dans le contexte de ce projet étaient davantage un prétexte à l'élaboration d'une approche efficace se préoccupant surtout des détails liés à l'organisation des données, plutôt qu'à l'interprétation des résultats. De ce fait, il pourrait être intéressant de se pencher davantage sur des questions en lien avec le comportement des usagers, telles que le type de clientèle qui fréquente chaque station selon le titre utilisé, les temps de déplacements, ou encore l'itinéraire emprunté.

Finalement, l'analyse d'un ensemble de données qui s'étend sur plus d'une journée constituerait une perspective de recherche intéressante. En premier lieu, ce genre d'analyse permettrait de complémenter la méthode de travail proposée ici, puisque l'ensemble de données serait beaucoup plus grand. Puis, il serait possible de valider les hypothèses utilisées ici pour dériver de l'information supplémentaire. La proportion de déplacements pour lesquels une station de destination est dérivée pourrait augmenter significativement. Enfin, des études sur les comportements à long terme des usagers pourraient être conduites, telles que les heures typiques de déplacement, les stations habituelles des usagers, ou l'évolution du comportement des usagers dans le temps.

BIBLIOGRAPHIE

AECOM (2009). Revisiting the Portland GEN2 Modeling Process with TRANSIMS Version 4.0 Software Methods (DTFH61-04-D-00015). AECOM. Consulté le 16 janvier 2012, tiré de http://transims.googlecode.com/svn/v4/trunk/implementations/Portland%20Phase%202%20Technical%20Report.pdf

Agard, B., Morency, C., & Trépanier, M. (2006). Mining public transport user behaviour from smart card data. 12th IFAC Symposium on Information Control Problems in Manufacturing. (Vol. 12, Partie #1). École des Mines, Saint-Étienne, France.

Agence métropolitaine de transport (2008). Portrait de la mobilité. Faits saillants révélés par l'enquête Origine-Destination 2008 : Agence Métropolitaine de Transport. Consulté le 21 juillet 2011, tiré de http://www.enquete-od.qc.ca/docs/EnqOD08_FaitsSaillants.pdf

Argonne National Laboratory (2010). TRANSIMS Studio and the RTE Run Time Environment. Consulté le 8 décembre 2011, tiré de http://www.anl.gov/TRACC/Training/TMS/04_8-9_10/H_Ley/Transims%20Studio.pdf

Bagchi, M., & White, P.R. (2005). The potential of public transport smart card data. *Transport Policy*, *12*(5), 464-474.

Barry, J.J., Newhouser, R., Rahbee, A., & Sayeda, S. (2002). Origin and destination estimation in New York City with automated fare system data, *Transportation Research Record* (1817), 183-187.

Beckman, R.J. (1997). TRansportation ANalysis SIMulation System (TRANSIMS) The Dallas-Ft. Worth Case Study. *Los Alamos National Laboratory* (LAUR-97-4502). Consulté le 16 janvier 2012, tiré de http://transims.googlecode.com/svn/v4/trunk/implementations/TRANSIMSBuffaloFinalReportNov2009.pdf

Bergeron, M. (2008). Vocabulaire des systèmes de transport intelligents : Vocabulaire français-anglais, *Gouvernement du Québec*, ISBN 978-2-550-52829-6. Tiré de http://www.oqlf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/voc_sti_20080602.pdf

Bisson, B. (2011). Le coût d'entretien du métro s'emballe. Cyberpresse.ca. Consulté le 5 juillet 2011. Tiré de http://www.cyberpresse.ca/actualites/regional/montreal/201107/03/01-4414672-le-cout-dentretien-du-metro-semballe.php

Bisson, B. (2012). Métro de Montréal : les pannes sont là pour rester. Cyberpresse.ca. Consulté le 24 avril 2012. Tiré de http://www.cyberpresse.ca/actualites/regional/montreal/201203/20/01-4507250-metro-de-montreal-les-pannes-sont-la-pour-rester.php

Blais, S. (2007). Modélisation espace-temps du métro de Montréal. M. Sc. A., École Polytechnique de Montréal, Montréal, Québec, Canada.

Blythe, P.T. (2004). Improving public transport ticketing through smart cards. *Proceedings of the Institution of Civil Engineers.* Municipal Engineer (Vol. 157, Issue 1, pp. 47-54).

Center for Urban Transportation Research. (2011). *Expanding the Google Transit Feed Specification to Support Operations and Planning* (FDOT BDK85 # 977-15). Tallahassee, Floride : Center for Urban Transportation Research. Consulté le 1er mars 2012, tiré de http://www.nctr.usf.edu/wp-content/uploads/2012/02/77902.pdf

Chapleau, R. (1997). Les pannes de métro et ses clients victimes. *32e congrès annuel de l'Association Québécoise du Transport et des routes,* Trois-Rivières, Qc.

Chapleau, R. (2009). Cours CIV6708 – Transport en commun, matériel électornique disponible aux étudiants, Polytechnique Montréal.

Chapleau, R., & Chu, K.K. (2007). Modeling transit travel patterns from location-stamped smart card data using a disaggregate approach, *11th World Conference on Transportation Research*, Berkeley, California.

Chu, K.K. (2010). *Leveraging data from a smart card automatic fare collection system for public transit planning*. (Ph.D.). Consulté le 25 juillet 2011, tiré de PolyPublie (http://publications.polymtl.ca/441/1/2010_KaKeeChu.pdf)

Chu, K.K., & Chapleau, R. (2008). Enriching archived smart card transaction data for transit demand modeling. *Transportation Research Record : Journal of the Transportation Research Board* (2063), 63-72.

Communauté Urbaine De Montréal, Bureau De Transport Métropolitain (1983). Le Métro de Montréal. *Montréal : Bureau de transport métropolitain.*

Corriveau, J. (2011). Métro de Montréal : feu vert à Bombardier-Alstom. LeDevoir.com. Consulté le 8 août 2011. Tiré de http://www.ledevoir.com/politique/montreal/297508/metro-de-montreal-feu-vert-a-bombardier-alstom

Corriveau, J. (2012). Les cartes à puce expireront plus tôt que prévu. LeDevoir.com. Consulté le 1er mars 2012. Tiré de http://www.ledevoir.com/politique/montreal/344023/les-cartes-opus-expireront-plus-tot-que-prevu

Devillaine, F., Munizaga, M. & Trépanier, M. (2012). Detection of public transport user activites through the analysis of smartcard data. Transportation Research Board – 91st Annual Meeting, Washington, D.C. (17 p.).

Dueker, K.J., & Peng, Z.-R. (2008). Geographic information systems for transport (Gis-T). In Hensher & Button (éd.), *Handbook of transport modelling* (pp. 303-328). Oxford : Elsevier.

Ferris et al. (2010). OneBusAaway : Results from providing real-time arrival information for public transit. *Proceedings of CHI 2010.* Atlanta, GA, USA. Consulté le 2 mai 2012, tiré de http://wiki.onebusaway.org/bin/download/Main/Research/chi-2010-onebusaway-paper.pdf

FHWA TMIP (2005). Chapter 1 TRANSIMS OVERVIEW, *Federal Highway Administration Travel Modeling Improvement Program*. Consulté le 9 décembre 2011, tiré de http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/transims_fundamentals/ch1.pdf

GTFS Google Code (2012). Consulté le 15 février 2012, tiré de https://developers.google.com/transit/gtfs/reference

Google Transit (2012). Consulté le 2 mai 2012, tiré de http://www.google.com/intl/en/landing/transit/#mdy

Hensher, D.A. & Button, K.J. (2008). Introduction. In Hensher & Button (éd.), *Handbook of transport modelling* (pp. 1-9). Oxford : Elsevier.

Klein, O. (2007). *Modélisation et représentations spatio-temporelles des déplacements quotidiens urbains*. (Ph.D.). Consulté le 20 février 2012, tiré du serveur des thèses de l'Université de Strasbourg (http://scd-theses.u-strasbg.fr/1387/01/KLEIN_Olivier_2007.pdf)

Klein, O. (2011). Visualizing Daily Mobility : Towards Other Modes of Representation. In Banos & Thévenin (éd.), *Geographical Information and Urban Transport Systems* (pp. 167-220). Great Britain : Wiley.

Manheim, M. L. (1979). *Fundamentals of transportation systems analysis*. Cambridge, Mass. : MIT Press.

Métro de Montréal (2012). Consulté le 1er juillet 2012, tiré de http://metrodemontreal.com/index-f.html

Mouvement Collectif (2009). La carte OPUS. Consulté le 09 août 2011, tiré de http://www.mouvementcollectif.org/fr/grand_projets/la-carte-opus.

Munizaga, M., Palma, C., & Fischer, D. (2011). Estimation of a disaggregate multimodal public transport od matrix from passive smart card data from Santiago, Chile. Transportation Research Board – 90th Annual Meeting, Washington, D.C. (15 p.).

National Transportation Library (1995). TRANSIMS: Project Summary and Status MAY 1995. Consulté le 8 décembre 2011, tiré de http://ntl.bts.gov/DOCS/466.html

Okamura, T., Zhang, J., & Akimasa, F. (2004). Finding Behavioral Rules of Urban Public Transport Passengers by Using Boarding Records of Integrated Stored Fare Card System. *10th World Conference on Transportation Research,* Istanbul, Turquie, 2004.

OQLF. (2011). Le grand dictionnaire terminologique. Consulté le 26 septembre 2011, tiré de http://www.oqlf.gouv.qc.ca/ressources/gdt.html

Piché, D. & Chapleau, R. (2007). Flux dans le métro de Montréal : concepts et mesures. *75e congrès de l'Association francophone pour le savoir*, Trois-Rivières, Qc.

QGIS (2011a). À propos de QGIS. Consulté le 7 décembre 2011, tiré de http://qgis.org/fr/a-propos-de-qgis.html

QGIS (2011b). Fonctionnalités de QGIS. Consulté le 7 décembre 2011, tiré de http://www.qgis.org/fr/a-propos-de-qgis/fonctionnalites.html

Schelfer, M., & Procaccino, J.D. (2002). Smart Card Evolution, *Communications of the ACM*, *45*(7), 83-88.

SourceForge (2010a). TRANSIMS studio. Consulté le 9 décembre 2011, tiré de http://sourceforge.net/projects/transimsstudio/

SourceForge (2010b). TRANSIMS SCM repositories. Consulté le 9 décembre 2011, tiré de http://transims.svn.sourceforge.net/viewvc/transims?view=revision&revision=1191

STM (2010). Rapport d'activité 2010. Consulté le 1er mai 2012, tiré de http://www.stm.info/en-bref/ra2010.pdf

STM (2011a). Plan interactif. Consulté le 6 décembre 2011, tiré de http://stm.info/metro/mapmetro.htm

STM (2011b). FAQ Tarification. Consultée le 6 octobre 2011, tiré de http://www.stm.info/tarification/faq.htm

STM (2011c). Rapport d'activité 2011. Consulté le 1er mai 2012, tiré de http://www.stm.info/en-bref/ra2011.pdf

STM (2012a). Budget 2012. Consulté le 1er mai 2012, tiré de http://www.stm.info/en-bref/budget2012.pdf

STM (2012b). Grille tarifaire 2012. Consultée le 17 janvier 2012, tiré de http://stm.info/tarification/grilletarif.htm

STM (2012c). Tout sur la STM. Consulté le 9 mai 2012, tiré de http://stm.info/en-bref/toutsurlaSTM.htm#q14 .

STM EN BREF (2011). Tout sur la STM. Consulté le 2 juin 2011, tiré de http://www.stm.info/en-bref/toutsurlaSTM.htm

Thévenin, T. (2011). Modeling Transport Systems on an Intra-Urban Scale. In Banos & Thévenin (éd.), *Geographical Information and Urban Transport Systems* (pp. 167-220). Great Britain : Wiley.

TRANSIMS Google Code (2011a). TRANSIMS documentation. Consulté le 19 décembre 2011, tiré de http://code.google.com/p/transims/source/browse/documentation/NASA%20Open%20Source%20Agreement%201_3.htm

TRANSIMS Google Code (2011b). How To, Quick Reference, Training & User Manuals. Consulté le 9 décembre 2011, tiré de http://code.google.com/p/transims/source/browse/#svn%2Fv4%2Ftrunk%2Fdocumentation

Trépanier et al. (2004). Examen des potentialités d'analyse des données d'un système de paiement par carte à puce en transport urbain. *Utilisation des systèmes de transport intelligents (STI) à l'appui de la gestion de la circulation : Congrès annuel de l'Association des transports du Canada à Québec,* Québec.

Trépanier, M. & Chapleau, R. (2001a). Analyse orientée-objet des données d'enquêtes origine-destination. *Canadian Journal of Civil Engineering*, *28*(1), 48-58.

Trépanier, M. & Chapleau, R. (2001b). Linking transit operational data to road network with a transportation object-oriented GIS. *URISA Journal 13*(2), 23-30.

Trépanier, M., Chpaleau, R. & Tranchant, N. (2005). Cartes à puces en transport en commun : une exploitation informationnelle pour fins de planification. *40e congrès annuel de l'Association Québécoise du Transport et des routes,* Laval, Qc.

Utsunomiya, M., Attanucci, J., & Wilson, N. (2006). Potential uses of transit smart card registration and transaction data to improve transit planning. *Transportation Research Record : Journal of the Transportation Research Board* (1971), 119-126.

Vuchic, V.R. (2005). *Urban transit: operations, planning, and economics*. Hoboken, New Jersey : John Wiley & Sons, INC.

Zhao, J. (2004). The planning and analysis implications of automatic data collection systems: Rail transit OD matrix inference and path choice modelling examples. (M.S. thesis, MIT, USA). Consulté le 25 juillet 2011, tiré de http://dspace.mit.edu/bitstream/handle/1721.1/28752/59820922.pdf?sequence=1

Annexe A : Procédures de traitement des données dans STATA

1. Doublons et activité BF et BG

```
*******************Correction du champ heure28 (: -> .)*******************
gen hre = regexs(2) if regexm(heure28, "(([0-9]+)[:]*([0-9]+)[:]*([0-9]+))")
gen min = regexs(3) if regexm(heure28, "(([0-9]+)[:]*([0-9]+)[:]*([0-9]+))")
gen sec = regexs(4) if regexm(heure28, "(([0-9]+)[:]*([0-9]+)[:]*([0-9]+))")

gen heure28corr = heure28
replace heure28corr = hre + "." + min + "." + sec if hre!=""
drop hre min sec
*****************************************************************************

*******************Retenir les enregistrements uniques*******************
gsort +support +heure28corr +activite +refus +id_transac +debcourse +fincourse +lieu +ceb

drop if (id_transac == id_transac[_n-1] & debcourse == debcourse[_n-1] & ///
fincourse == fincourse[_n-1] & course==course[_n-1] & ligne == ligne[_n-1] & ///
reseauc == reseauc[_n-1] & dir==dir[_n-1] & trajet==trajet[_n-1] & ///
lieu == lieu[_n-1] & tarif1==tarif1[_n-1]& reseau==reseau[_n-1] & dh==dh[_n-1] & ///
ceb==ceb[_n-1] & support ==support[_n-1] & produit==produit[_n-1] & refus==refus[_n-1] & ///
activite==activite[_n-1] & heure28corr==heure28corr[_n-1] & test==test[_n-1] & ///
date28==date28[_n-1])
*****************************************************************************

****************************Supprimer BF et BG****************************
drop if activite =="BF" | activite =="BG" | support=="non-valide"
*****************************************************************************
```

2. Refus consécutifs

```
**************Station de transaction****************
generate station = ""
replace station = substr(ceb,7,3)
**************************************************

**********************Suppression des refus consécutifs*********************************
gsort +support +heure28corr +activite +refus +id_transac +debcourse +fincourse +lieu +ceb

drop if (refus>0 & station==station[_n-1] & support==support[_n-1] & support != "non-valide")
*****************************************************************************************
```

3. Nombre de transactions par support

```
***************************Nombre de transactions et numéro de transaction****************************
egen sup_num = group(support)
by sup_num, sort : egen int nb_transac = count(sup_num)
gsort +nb_transac +support +heure28corr +activite +refus
by nb_transac sup_num, sort : gen no_trans = _n
gsort +nb_transac +support +heure28corr +activite +refus +id_transac +debcourse +fincourse +lieu +ceb
*******************************************************************************************************
```

4. Dérivation des stations de destination

```
**********************************Station de destination*************************************
gen destination = ""

*Destination de la première transaction
replace destination = station[_n+1] if no_trans == 1 & nb_transac >= 2 & support==support[_n+1]

*Destination du 2e voyage (2 transactions exactement -> déplacement symétrique)
replace destination = station[_n-1] if no_trans >= 2 & nb_transac == 2 & support==support[_n-1]

*Destinations des usagers effectuant plus de 2 transactions
replace destination = station[_n+1] if no_trans >= 2 & nb_transac > 2 & support==support[_n+1]

*Retour à la station d'origine des usagers effectuant plus de 2 transactions
replace destination = station[_n-nb_transac+1] if no_trans > 2 & nb_transac > 2 & support!=support[_n+1]

*Correction de la dernière destination de la série si la transaction est refusée
replace destination = "" if no_trans == nb_transac & nb_transac >= 2 & refus > 0

*********************************************************************************************
```

Annexe B : Commandes STATA

Commande	**Fonction**	**Exemple**
tabout*	Exporter un tableau de données	tabout station destination using C:/Stata/matrice.txt
distinct*	Compter le nombre de valeurs uniques que possède un champ	distinct ceb
cond(a,b,c)	Correspond au « IF » de MS Excel	replace serie = cond(condition, valeur si vrai, valeur si faux)

*La fonction doit être téléchargée du site d'extensions pour Stata

http://ideas.repec.org/c/boc/bocode/s447101.html

http://ideas.repec.org/c/boc/bocode/s424201.html

Annexe C : Paramètres des fichiers de contrôle TRANSIMS

../control/TransimsNet.ctl

```
#GÉNÉRATION DU RÉSEAU
TITLE                                   Synthetic TRANSIMS Network
DEFAULT_FILE_FORMAT                     TAB_DELIMITED
PROJECT_DIRECTORY                       ../

#---- FICHIERS D'INPUT ----#

#---- Obligatoire ----#
NET_DIRECTORY                           ../Input
NET_NODE_TABLE                          Input_Node.txt
NET_LINK_TABLE                          Input_Link.txt
#NET_ZONE_TABLE                         Input_Zones.txt

#---- Facultatif ----#
#NET_SHAPE_TABLE                        Input_Shape.txt
#NET_TURN_PROHIBITION_TABLE             Turn_Prohibition.txt
KEEP_NODE_LIST                          Keep_Node_List.txt

#---- FICHIERS D'OUTPUT ----#

#---- Obligatoire ----#
NEW_DIRECTORY                           ../network
NEW_NODE_TABLE                          Node
NEW_ZONE_TABLE                          Zone
NEW_LINK_TABLE                          Link
NEW_SHAPE_TABLE                         Shape
NEW_ACTIVITY_LOCATION_TABLE             Activity_Location
NEW_PARKING_TABLE                       Parking
NEW_PROCESS_LINK_TABLE                  Process_Link
NEW_LANE_CONNECTIVITY_TABLE             Lane_Connectivity
NEW_POCKET_LANE_TABLE                   Pocket_Lane
NEW_UNSIGNALIZED_NODE_TABLE             Sign_Warrants
NEW_SIGNALIZED_NODE_TABLE               Signal_Warrants

## --- Facultatif ---------
#NEW_TURN_PROHIBITION_TABLE             Turn_Prohibition
#LINK_NODE_EQUIVALENCE                  Link_Node
```

```
#---- Paramètres ----#

ACTIVITY_LOCATION_SIDE_OFFSET          15              //---- mètres
MAXIMUM_ACCESS_POINTS                  2
MINIMUM_SPLIT_LENGTHS                  40, 40, 40, 40, 40, 40, 40, 40
MINIMUM_LINK_LENGTH                    7.5             //---- mètres
MAXIMUM_CONNECTION_ANGLE               180
#MAXIMUM_LENGTH_TO_XY_RATIO            1.2
INTERSECTION_SETBACK_DISTANCE          7.5             //---- mètres
#FIRST_EXTERNAL_ZONE_NUMBER            10
#COLLAPSE_NODES_FLAG                   TRUE
#ADD_UTURN_TO_DEADEND_LINKS            YES
CREATE_NOTES_AND_NAME_FIELDS           YES
REPORT_FLAG                            FALSE
```

../control/TransitNet.ctl

```
#TRANSPORT EN COMMUN
TITLE                                   Convert Transit Network
DEFAULT_FILE_FORMAT                     TAB_DELIMITED
PROJECT_DIRECTORY                       ../

#---- Fichiers réseau ----#

NET_DIRECTORY                           ../network
NET_NODE_TABLE                          Node
NET_ZONE_TABLE                          Zone
NET_LINK_TABLE                          Link
NET_PARKING_TABLE                       Parking
NET_ACTIVITY_LOCATION_TABLE             Activity_Location
NET_PROCESS_LINK_TABLE                  Process_Link
NET_LANE_CONNECTIVITY_TABLE             Lane_Connectivity
ROUTE_HEADER_FILE                       ../input/Route_Header.txt
ROUTE_NODES_FILE                        ../input/Route_Nodes.txt
#PARK_AND_RIDE_FILE                     /network/Park_Ride.txt
#ZONE_EQUIVALENCE_FILE                  input/Fare_Zone.txt

#---- FICHIERS D'OUTPUT ----#

NEW_DIRECTORY                           ../network
#NEW_PARKING_TABLE                      Parking_2
NEW_ACTIVITY_LOCATION_TABLE             Activity_Location_2
NEW_PROCESS_LINK_TABLE                  Process_Link_2
NEW_TRANSIT_STOP_TABLE                  Transit_Stop
NEW_TRANSIT_ROUTE_TABLE                 Transit_Route
NEW_TRANSIT_SCHEDULE_TABLE              Transit_Schedule
NEW_TRANSIT_DRIVER_TABLE                Transit_Driver
CREATE_NOTES_AND_NAME_FIELDS            YES

#---- Paramètres ----#

STOP_SPACING_BY_AREATYPE                40, 40, 40, 40, 40, 40, 40, 40
TRANSIT_TIME_PERIODS                    5:00, 7:00, 9:00, 15:00, 18:00, 22:00
TRANSIT_TRAVEL_TIME_FACTOR              1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0
MINIMUM_DWELL_TIME                      22              //---- seconds ----
INTERSECTION_STOP_TYPE                  NEARSIDE
TRANSITNET_REPORT_1                     FARE_ZONE_EQUIVALENCE
```

../control/ArcNet.ctl

```
#GÉNÉRATION DES SHAPEFILES DU RÉSEAU
TITLE                                   Convert the Network to Shapefiles
DEFAULT_FILE_FORMAT                     TAB_DELIMITED
PROJECT_DIRECTORY                       ../

#---- Fichiers réseau ----#

NET_DIRECTORY                           ../network
NET_NODE_TABLE                          Node
NET_LINK_TABLE                          Link
NET_ZONE_TABLE                          Zone
NET_SHAPE_TABLE                         Shape
NET_ACTIVITY_LOCATION_TABLE             Activity_Location_2
NET_PARKING_TABLE                       Parking
NET_PROCESS_LINK_TABLE                  Process_Link_2
NET_POCKET_LANE_TABLE                   Pocket_Lane
NET_LANE_CONNECTIVITY_TABLE             Lane_Connectivity
NET_TRANSIT_STOP_TABLE                  Transit_Stop
NET_TRANSIT_ROUTE_TABLE                 Transit_Route
NET_TRANSIT_SCHEDULE_TABLE              Transit_Schedule
NET_TRANSIT_DRIVER_TABLE                Transit_Driver
NET_SIGNALIZED_NODE_TABLE               Signal_Warrants
NET_UNSIGNALIZED_NODE_TABLE             Sign_Warrants
#NET_TIMING_PLAN_TABLE                   Timing_Plan
#NET_PHASING_PLAN_TABLE                  Phasing_Plan
#NET_DETECTOR_TABLE                      Detector
#NET_SIGNAL_COORDINATOR_TABLE            Signal_Coordinator

#---- FICHIERS D'OUTPUT ----#

ARCVIEW_DIRECTORY                       ../network/arcview

#---- Paramètres ----#

DRAW_NETWORK_LANES                      YES
LANE_WIDTH                              3.5             //---- meters ----
LINK_DIRECTION_OFFSET                   0             //---- meters ----
ACTIVITY_LOCATION_SIDE_OFFSET           5              //---- meters ----
PARKING_SIDE_OFFSET                     5               //---- meters ----
```

../control/Router.ctl

```
#AFFECTATION DES DÉPLACEMENTS
TITLE                                   Route the Highway and Transit Trips
DEFAULT_FILE_FORMAT                     TAB_DELIMITED
PROJECT_DIRECTORY                       ../

#---- Fichiers Réseau ---#
NET_DIRECTORY                           ../network
NET_NODE_TABLE                          Node
NET_LINK_TABLE                          Link
NET_PARKING_TABLE                       Parking
NET_LANE_CONNECTIVITY_TABLE             Lane_Connectivity
NET_ACTIVITY_LOCATION_TABLE             Activity_Location_2
NET_PROCESS_LINK_TABLE                  Process_Link_2
NET_TRANSIT_STOP_TABLE                  Transit_Stop
#NET_TRANSIT_FARE_TABLE                  Transit_Fare
NET_TRANSIT_ROUTE_TABLE                 Transit_Route
NET_TRANSIT_SCHEDULE_TABLE              Transit_Schedule

#---- FICHIERS D'INPUT ----#
TRIP_FILE                               demand/trips.txt
TIME_OF_DAY_FORMAT                      24_HOUR_CLOCK
VEHICLE_FILE                            demand/vehicles.txt
#PLAN_FILE                               demand/TravelPlans

#---- FICHIERS D'OUTPUT ----#
NEW_PLAN_FILE                           results/Plans
NEW_PROBLEM_FILE                        results/ROUTER_PROBLEMS

#---- Options ----#
NODE_LIST_PATHS                         NO
LIMIT_PARKING_ACCESS                    NO
IGNORE_TIME_CONSTRAINTS                 YES
WALK_PATH_DETAILS                       NO

#---- Paramètres ----#
WALK_SPEED                              1.0        //---- meters / second ----
VEHICLE_TIME_VALUE                      10.0       //---- imped / second ----
FIRST_WAIT_VALUE                        10.0       //---- imped / second ----
TRANSFER_WAIT_VALUE                     10.0       //---- imped / second ----
DISTANCE_VALUE                          1.0        //---- imped / meter ----
COST_VALUE                              5.0        //---- imped / cent ----

#---- Contraintes ----#
MAX_WALK_DISTANCE                       500            //---- meters ----
MAX_WAIT_TIME                           60              //---- minutes ----
#LEFT_TURN_PENALTY                       300             //---- impedance ----
#UTURN_PENALTY                           5000            //---- impedance ----
```

../control/PlanSum.ctl

```
#RÉSUMER LES DÉPLACEMENTS
TITLE                                   Summarize Travel Plans
DEFAULT_FILE_FORMAT                     TAB_DELIMITED
PROJECT_DIRECTORY                       ../

#---- Fichiers réseau----#
NET_DIRECTORY                           ../network
NET_NODE_TABLE                          Node
NET_LINK_TABLE                          Link
NET_PARKING_TABLE                       Parking
NET_ACTIVITY_LOCATION_TABLE             Activity_Location_2
NET_LANE_CONNECTIVITY_TABLE             Lane_Connectivity
NET_PROCESS_LINK_TABLE                  Process_Link
NET_TRANSIT_STOP_TABLE                  Transit_Stop
NET_TRANSIT_ROUTE_TABLE                 Transit_Route
NET_TRANSIT_SCHEDULE_TABLE              Transit_Schedule
NET_TRANSIT_DRIVER_TABLE                Transit_Driver

#---- FICHIERS D'INPUT ----#
PLAN_FILE                               results/Plans
NODE_LIST_PATHS                         YES

#---- Paramètres ----#
SUMMARY_TIME_PERIODS                    4:00..27:00
SUMMARY_TIME_INCREMENT                  15              //---- minutes ----
#EQUATION_PARAMETERS_1                   BPR, 0.15, 4.0, 0.75
#EQUATION_PARAMETERS_2                   BPR, 0.10, 4.5, 0.75
```

```
#---- FICHIERS D'OUTPUT ----#

NEW_LINK_DELAY_FILE                     results/Performance
NEW_LINK_DELAY_FORMAT                   TAB_DELIMITED

NEW_RIDERSHIP_FILE                      results/Ridership
NEW_RIDERSHIP_FORMAT                    TAB_DELIMITED

NEW_LINK_VOLUME_FILE                    results/Volume
NEW_LINK_VOLUME_FORMAT                  TAB_DELIMITED

NEW_ZONE_SKIM_FILE                      results/Skim
NEW_ZONE_SKIM_FORMAT                    TAB_DELIMITED
SKIM_TOTAL_TIME                         YES

NEW_TRIP_TABLE_FILE                     results/TripData
NEW_TRIP_TABLE_FORMAT                   TAB_DELIMITED
#ZONE_EQUIVALENCE_FILE                   input/Zone_Equiv.txt

NEW_TRIP_TIME_FILE                      results/TripTimes
#NEW_TRIP_TIME_FORMAT                    TAB_DELIMITED

PLANSUM_REPORT_1                        TRAVEL_SUMMARY_REPORT
PLANSUM_REPORT_2                        TRANSIT_RIDERSHIP_SUMMARY
PLANSUM_REPORT_3                        TRANSIT_STOP_SUMMARY
PLANSUM_REPORT_4                        TRANSIT_PASSENGER_SUMMARY
PLANSUM_REPORT_5                        TRANSIT_TRANSFER_SUMMARY
PLANSUM_REPORT_6                        TRANSIT_TRANSFER_DETAILS
PLANSUM_REPORT_7                        TRIP_TIME_REPORT
```

../control/ArcPlan.ctl

```
#CRÉATION DES SHAPEFILES D'ITINÉRAIRES
TITLE                                   Display the Plans
PROJECT_DIRECTORY                       ../
DEFAULT_FILE_FORMAT                     TAB_DELIMITED

#---- Fichiers réseau ----#
NET_DIRECTORY                           ../network
NET_NODE_TABLE                          Node
NET_LINK_TABLE                          Link
NET_SHAPE_TABLE                         Shape
NET_PARKING_TABLE                       Parking
NET_ACTIVITY_LOCATION_TABLE             Activity_Location_2
NET_TRANSIT_STOP_TABLE                  Transit_Stop
NET_TRANSIT_ROUTE_TABLE                 Transit_Route
NET_TRANSIT_DRIVER_TABLE                Transit_Driver

#---- FICHIER D'INPUT ----#
PLAN_FILE                               results/Plans

#---- Paramètres ----#
NODE_LIST_PATHS                         NO
TIME_OF_DAY_FORMAT                      24_HOUR_CLOCK
SELECT_TIME_PERIODS                     4:00..27:00
#SELECT_RANDOM_PERCENTAGE               10              //---- percent ----
#RANDOM_NUMBER_SEED                     11010

#---- FICHIER D'OUTPUT ----#
ARCVIEW_PLAN_FILE                       results/arcview/Plans.shp
ARCVIEW_BANDWIDTH_FILE                  results/arcview/Bandwidth.shp
#ARCVIEW_TIME_CONTOUR                   results/arcview/TimeContour.shp
#ARCVIEW_DISTANCE_CONTOUR               results/arcview/DistContour.shp
#ARCVIEW_ACCESSIBILITY_FILE             results/arcview/Access.shp
#---- Distances ----#
LINK_DIRECTION_OFFSET                   2.0             //---- meters ----
ACTIVITY_LOCATION_SIDE_OFFSET           15.0            //---- meters ----
PARKING_SIDE_OFFSET                     5.0             //---- meters ----
TRANSIT_STOP_SIDE_OFFSET                8.0             //---- meters ----
TRANSIT_DIRECTION_OFFSET                4.0             //---- meters ----

#---- Visualisation ---#
BANDWIDTH_SCALING_FACTOR                5            //---- vehicles / meter ----
MAXIMUM_SHAPE_ANGLE                     45              //---- degrees ----
MINIMUM_SHAPE_LENGTH                    50               //---- meters ----
#CONTOUR_TIME_INCREMENTS                60, 120, 180, 240, *120
#CONTOUR_DISTANCE_INCREMENTS            *500         //---- meters ----
#OUTPUT_COORDINATE_SYSTEM               UTM, 18N, METERS
```

../control/ArcRider.ctl

```
#GÉNÉRATION DES SHAPEFILES D'ACHALANDAGE TC
TITLE                                   Display the Simulation Results
PROJECT_DIRECTORY                       ../
DEFAULT_FILE_FORMAT                     TAB_DELIMITED

#---- Fichiers réseau ----#
NET_DIRECTORY                           ../network
NET_NODE_TABLE                          Node
NET_LINK_TABLE                          Link
NET_SHAPE_TABLE                         Shape
#NET_LANE_USE_TABLE                      Lane_Use
NET_TRANSIT_STOP_TABLE                  Transit_Stop
NET_TRANSIT_ROUTE_TABLE                 Transit_Route
NET_TRANSIT_SCHEDULE_TABLE              Transit_Schedule
NET_TRANSIT_DRIVER_TABLE                Transit_Driver

#---- FICHIERS D'INPUT ----#
VEHICLE_TYPE_FILE                       demand/Vehicle_Type.txt
RIDERSHIP_FILE                          results/Ridership
LINE_EQUIVALENCE_FILE                   input/Line_Equiv.txt

#---- FICHIERS D'OUTPUT ----#
ARCVIEW_LINE_DEMAND_FILE                results/arcview/Line_Demand.shp
ARCVIEW_LINE_GROUP_FILE                 results/arcview/Line_Group.shp
ARCVIEW_RIDERSHIP_FILE                  results/arcview/Ridership.shp
ARCVIEW_STOP_DEMAND_FILE                results/arcview/Stop_Demand.shp

#---- Paramètres ----#
TIME_OF_DAY_FORMAT                      24_HOUR_CLOCK
SELECT_TIME_PERIODS                     5:00..27:00

##SELECT_TRANSIT_STOPS                   8990, 2990, 2099
##SELECT_TRANSIT_ROUTES                  23..26
##SELECT_TRANSIT_MODES                   BUS, RAPIDRAIL

TRANSIT_STOP_SIDE_OFFSET                8               //---- meters ----
TRANSIT_DIRECTION_OFFSET                4               //---- meters ----

#MAXIMUM_SHAPE_ANGLE                    45              //---- degrees ----
#MINIMUM_SHAPE_LENGTH                   1               //---- meters ----
RIDERSHIP_SCALING_FACTOR                50             //---- riders / meter ----
#MINIMUM_RIDERSHIP_VALUE                10              //---- riders ----
#MINIMUM_RIDERSHIP_SIZE                 1.0             //---- meters ----
#MAXIMUM_RIDERSHIP_SIZE                 400.0           //---- meters ----

##OUTPUT_XYM_SHAPES                     TRUE
OUTPUT_COORDINATE_SYSTEM                UTM, 18N, METERS
```

../control/PlanSort.ctl

```
#TRIER LES DÉPLACEMENTS
TITLE                                   Sort the Plan Files for Simulation
DEFAULT_FILE_FORMAT                     TAB_DELIMITED
PROJECT_DIRECTORY                       ../

INPUT_PLAN_FILE                         results/Plans
OUTPUT_PLAN_FILE                        results/TimePlans
PLAN_SORT_OPTION                        TIME
```

../control/Microsimulator.ctl

```
#MICROSIMULATION
TITLE                                   Simulate the Trips
DEFAULT_FILE_FORMAT                     TAB_DELIMITED
PROJECT_DIRECTORY                       ../

#---- Fichiers réseau ----#
NET_DIRECTORY                           ../network
NET_NODE_TABLE                          Node
NET_LINK_TABLE                          Link
NET_POCKET_LANE_TABLE                   Pocket_Lane
NET_PARKING_TABLE                       Parking
NET_LANE_CONNECTIVITY_TABLE             Lane_Connectivity
NET_ACTIVITY_LOCATION_TABLE             Activity_Location_2
NET_PROCESS_LINK_TABLE                  Process_Link_2
#NET_UNSIGNALIZED_NODE_TABLE            Unsignalized_Node
#NET_SIGNALIZED_NODE_TABLE              Signalized_Node
#NET_TIMING_PLAN_TABLE                  Timing_Plan
#NET_PHASING_PLAN_TABLE                 Phasing_Plan
#NET_DETECTOR_TABLE                     Detector
#NET_SIGNAL_COORDINATOR_TABLE           Signal_Coordinator
NET_TRANSIT_STOP_TABLE                  Transit_Stop
NET_TRANSIT_ROUTE_TABLE                 Transit_Route
NET_TRANSIT_SCHEDULE_TABLE              Transit_Schedule
NET_TRANSIT_DRIVER_TABLE                Transit_Driver

#---- FICHIERS D'INPUT et propriétés ----#
VEHICLE_FILE                            demand/Vehicles.txt
SORT_VEHICLES                           TRUE
VEHICLE_TYPE_FILE                       demand/Vehicle_Type.txt
PLAN_FILE                               results/TimePlans
NODE_LIST_PATHS                         No

#---- Paramètres ----#
CELL_SIZE                               7.5             //---- meters ----
TIME_STEPS_PER_SECOND                   1               //---- steps / second ----
TIME_OF_DAY_FORMAT                      24_HOUR_CLOCK
SIMULATION_START_TIME                   00:00
SIMULATION_END_TIME                     27:00
SPEED_CALCULATION_METHOD                CELL-BASED
```

```
DRIVER_REACTION_TIME                    0.7, 0.8, 0.9, 1.0
PERMISSION_PROBABILITY                  55              //---- percent ----
#SLOW_DOWN_PROBABILITY                   10, 15, 20
#SLOW_DOWN_PERCENTAGE                    10, 20
RANDOM_NUMBER_SEED                      0

MINIMUM_WAITING_TIME                    15              //---- seconds ----
MAXIMUM_WAITING_TIME                    9000             //---- seconds ----
MAX_ARRIVAL_TIME_VARIANCE               0              //---- minutes ----
MAX_DEPARTURE_TIME_VARIANCE             0              //---- minutes ----

NEW_PROBLEM_FILE                        results/MsimProblems
#MAX_SIMULATION_ERRORS                   100000

OUTPUT_SNAPSHOT_FILE_1                  results/Snapshots
OUTPUT_SNAPSHOT_TIME_FORMAT_1           24_HOUR_CLOCK
OUTPUT_SNAPSHOT_INCREMENT_1             0:15
OUTPUT_SNAPSHOT_TIME_RANGE_1            4:00..27:00
##OUTPUT_SNAPSHOT_LINK_RANGE_1           2..10, 14..16, 18, 20
##OUTPUT_SNAPSHOT_COORDINATES_1          3500, 3500, 4500, 4500

OUTPUT_SNAPSHOT_FILE_2                  results/SnapshotsSeconds
OUTPUT_SNAPSHOT_FORMAT_2                BINARY
OUTPUT_SNAPSHOT_INCREMENT_2             1
OUTPUT_SNAPSHOT_TIME_RANGE_2            4:00..26:00

OUTPUT_EVENT_TYPE_1                     START_TIME, END_TIME
OUTPUT_EVENT_FILE_1                     results/Events
OUTPUT_EVENT_FILTER_1                   60
OUTPUT_EVENT_TIME_FORMAT_1              24_HOUR_CLOCK
OUTPUT_EVENT_TIME_RANGE_1               4:00..26:00

OUTPUT_RIDERSHIP_FILE_1                 results/Ridership
OUTPUT_RIDERSHIP_FORMAT_1               TAB_DELIMITED
OUTPUT_RIDERSHIP_TIME_FORMAT_1          24_HOUR_CLOCK
OUTPUT_RIDERSHIP_TIME_RANGE_1           4:00..26:00
#OUTPUT_RIDERSHIP_ROUTE_RANGE_1          100..400
```

../visualizer/Sim_Vis.ctl

```
#FICHIER DE CONTRÔLE - INPUT DU TRANSIMS VISUALIZER
NET_LINK_TABLE                      ../network/Link
LINK_SHAPE_FILE                     ../network/arcview/Link.shp
NET_NODE_TABLE                      ../network/Node
NET_SHAPE_TABLE                     ../network/Shape
NET_TRANSIT_SCHEDULE_TABLE          ../network/Transit_Schedule
INDEX                               ../results/index
SNAPSHOT_FILE                       ../results/snapshot_compressed
VEHICLE_TYPE_FILE                   ../demand/vehicle_type.txt
```

Annexe D : Fichiers input TRANSIMS

../Input/Input_Node.txt

Node	X_COORD	Y_COORD	Z_COORD	NOTES
1	603940.81	5045506.55	0	Station Henri-Bourassa
2	604913.24	5045033.47	0	Station Sauvé
3	606287.4	5044518.99	0	Station Crémazie
4	607044.7	5044224.61	0	Station Jarry
5	608154.32	5043751.47	0	Station Jean-Talon
6	608997.61	5043311.22	0	Station Beaubien
7	609466.95	5042933.78	0	Station Rosemont
8	610199.76	5042569.33	0	Station Laurier
9	610716.08	5042228.81	0	Station Mont-Royal
10	611802.36	5041568.07	0	Station Sherbrooke
11	612394.09	5041196.79	0	Station Berri-UQAM
12	612763.09	5040636.86	0	Station Champ-de-Mars
13	612516.53	5040212.24	0	Station Place-d'Armes
14	612240.14	5039713.86	0	Station Square-Victoria
15	611943.66	5039283.82	0	Station Bonaventure
16	611641.51	5038923.82	0	Station Lucien-L'Allier
17	611230.25	5038262.17	0	Station Georges-Vanier
18	614225.87	5050280.02	0	Station Honoré-Beaugrand
19	613936.3	5049400.43	0	Station Radisson
20	613653.98	5048721.93	0	Station Langelier
21	613369.8	5048056.97	0	Station Cadillac
22	613384.45	5047213.62	0	Station Assomption
23	613398.06	5046350.24	0	Station Viau
24	612984.49	5045517.82	0	Station Pie-IX
25	613095.66	5044727.55	0	Station Joliette
26	612837.46	5044131.67	0	Station Préfontaine
27	613024.87	5043218.86	0	Station Frontenac
28	613015.73	5042192.16	0	Station Papineau
29	612668.4	5041645.36	0	Station Beaudry
30	612143.62	5040763.67	0	Station Saint-Laurent
31	611846.07	5040387.4	0	Station Place-des-Arts
32	611606.81	5039930.28	0	Station McGill
33	611359.94	5039574.05	0	Station Peel
34	610924.93	5038934.67	0	Station Guy-Concordia
35	610480.44	5038335.73	0	Station Atwater

36	611011.1	5037601.62	0	Station Lionel-Groulx
37	611806.28	5037110.42	0	Station Charlevoix
38	612061.65	5036284.93	0	Station LaSalle
39	612014.16	5035300.01	0	Station De l'Église
40	611671.27	5034985.74	0	Station Verdun
41	610862.09	5034718.04	0	Station Jolicoeur
42	610000.62	5034029.68	0	Station Monk
43	609198.96	5033462.39	0	Station Angrignon
44	615379.65	5042398.03	0	Station Longueuil-Université de Sherbrooke
45	614607.66	5040925.46	0	Station Jean-Drapeau
46	610461.63	5036934.37	0	Station Place-Saint-Henri
47	609131.46	5036555.5	0	Station Vendôme
48	607916.66	5037135.16	0	Station Villa-Maria
49	607234.71	5037813.03	0	Station Snowdon
50	606807.58	5038521.3	0	Station Côte-Sainte-Catherine
51	606196.04	5038919.32	0	Station Plamondon
52	605262.5	5038797.14	0	Station Namur
53	604698.29	5039289.33	0	Station De la Savane
54	603529.38	5040390.43	0	Station Du Collège
55	607557.82	5039023.86	0	Station Côte-des-Neiges
56	607999.99	5039738.46	0	Station Université-de-Montréal
57	608381.63	5040553.85	0	Station Édouard-Montpetit
58	608169.22	5041693.15	0	Station Outremont
59	607468.2	5042032.71	0	Station Acadie
60	607451.45	5042882.01	0	Station Parc
61	607788.65	5043348.77	0	Station De Castelnau
62	608675.16	5044705.2	0	Station Fabre
63	609073.83	5045383.48	0	Station D'Iberville
64	609271.14	5046076.49	0	Station Saint-Michel
65	602877.94	5040911.23	0	Station Côte-Vertu
66	602863.93	5046046.41	0	Station Cartier
67	600702.44	5046065.95	0	Station De La Concorde
68	599794.42	5045732.59	0	Station Montmorency
900	609271.14	5046676.49	0	Parking SM
901	609321.14	5046676.49	0	Parking SM

../Input/Input_Link.txt

LINK	ANODE	BNODE	TYPE	LANES_AB	SPEED_AB	LANES_BA	SPEED_BA	USE
901	900	901	LOCAL	1	18	1	18	ANY
201	1	2	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
202	2	3	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
203	3	4	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
204	4	5	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
205	5	6	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
206	6	7	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
207	7	8	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
208	8	9	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
209	9	10	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
210	10	11	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
211	11	12	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
212	12	13	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
213	13	14	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
214	14	15	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
215	15	16	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
216	16	17	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
217	17	36	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
118	18	19	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
119	19	20	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
120	20	21	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
121	21	22	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
122	22	23	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
123	23	24	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
124	24	25	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
125	25	26	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
126	26	27	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
127	27	28	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
128	28	29	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
129	29	11	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
130	11	30	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
131	30	31	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
132	31	32	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
133	32	33	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
134	33	34	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
135	34	35	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
136	35	36	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL

137	36	37	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
138	37	38	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
139	38	39	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
140	39	40	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
141	40	41	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
142	41	42	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
143	42	43	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
401	44	45	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
402	45	11	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
218	36	46	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
219	46	47	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
220	47	48	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
221	48	49	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
222	49	50	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
223	50	51	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
224	51	52	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
225	52	53	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
226	53	54	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
227	54	65	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
501	49	55	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
502	55	56	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
503	56	57	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
504	57	58	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
505	58	59	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
506	59	60	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
507	60	61	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
508	61	5	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
509	5	62	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
510	62	63	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
511	63	64	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
228	66	1	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
229	67	66	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL
230	68	67	HEAVYRAIL	1	18	1	18	WALK/RAPIDRAIL

../Input/Keep_Node_List.txt

NODE
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
900
901

../Input/Route_Nodes.txt

ROUTE	NODE	…		…		…		…	
1	18								
1	19	2	33	3	46	5	1	7	49
1	20	2	32	3	47	5	66	7	48
1	21	2	31	3	48	5	67	7	47
1	22	2	30	3	49	5	68	7	46
1	23	2	11	3	50	6	65	7	36
1	24	2	29	3	51	6	54	7	17
1	25	2	28	3	52	6	53	7	16
1	26	2	27	3	53	6	52	7	15
1	27	2	26	3	54	6	51	7	14
1	28	2	25	3	65	6	50	7	13
1	29	2	24	4	1	6	49	7	12
1	11	2	23	4	2	6	48	7	11
1	30	2	22	4	3	6	47	7	10
1	31	2	21	4	4	6	46	7	9
1	32	2	20	4	5	6	36	7	8
1	33	2	19	4	6	6	17	7	7
1	34	2	18	4	7	6	16	7	6
1	35	3	68	4	8	6	15	7	5
1	36	3	67	4	9	6	14	7	4
1	37	3	66	4	10	6	13	7	3
1	38	3	1	4	11	6	12	7	2
1	39	3	2	4	12	6	11	7	1
1	40	3	3	4	13	6	10	7	66
1	41	3	4	4	14	6	9	7	67
1	42	3	5	4	15	6	8	7	68
1	43	3	6	4	16	6	7	8	44
2	43	3	7	4	17	6	6	8	45
2	42	3	8	4	36	6	5	8	11
2	41	3	9	4	46	6	4	9	11
2	40	3	10	4	47	6	3	9	45
2	39	3	11	4	48	6	2	9	44
2	38	3	12	4	49	6	1	10	64
2	37	3	13	4	50	7	65	10	63
2	36	3	14	4	51	7	54	10	62
2	35	3	15	4	52	7	53	10	5
2	34	3	16	4	53	7	52	10	61
		3	17	4	54	7	51	10	60
		3	36	4	65	7	50	10	59

10	58
10	57
10	56
10	55
10	49
11	49
11	55
11	56
11	57
11	58
11	59
11	60
11	61
11	5
11	62
11	63
11	64

../Input/Route_Header.txt

ROUTE	NAME	MODE	HEADWAY_1	HEADWAY_2	HEADWAY_3	HEADWAY_4	HEADWAY_5	HEADWAY_6	HEADWAY_7
1	L1_A	RAPIDRAIL	35	10	9	8	7	5	4
2	L1_HB	RAPIDRAIL	38	10	9	8	7	6	4
3	L2_CV	RAPIDRAIL	31	8	8	8	7	6	6
4	L2_Hb_CV	RAPIDRAIL	31	8	8	8	7	6	6
5	L2_Hb_Mo	RAPIDRAIL	60	10	10	10	10	6	5
6	L2_CV_HB	RAPIDRAIL	60	10	10	10	10	6	4
7	L2_Mo	RAPIDRAIL	35	10	10	7	5	4	4
8	L4_BU	RAPIDRAIL	35	8	8	7	6	5	4
9	L4_L	RAPIDRAIL	35	10	9	8	7	5	4
10	L5_S	RAPIDRAIL	38	10	9	8	7	6	4
11	L5_SM	RAPIDRAIL	31	8	8	8	7	6	6

OFFSET_1	OFFSET_2	OFFSET_3	OFFSET_4	OFFSET_5	OFFSET_6	OFFSET_7
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

../demand/vehicle_type.txt

TYPE	SUBTYPE	LENGTH	MAX_SPEED	MAX_ACCEL	MAX_DECEL	USE	CAPACITY	LOADING	UNLOADING
1	0	7.5	37.5	7.5	7.5	CAR	11	2	3
1	1	7.5	37.5	7.5	7.5	CAR	11	2	3
1	2	7.5	37.5	7.5	7.5	CAR	11	2	3
2	0	22.5	37.5	7.5	7.5	TRUCK	2	2	3
2	1	15	37.5	7.5	7.5	LIGHTTRUCK	2	2	3
2	2	15	37.5	1.1	1.1	LIGHTTRUCK	2	2	3
2	3	18.5	37.5	1.1	1.1	HEAVYTRUCK	2	2	3
2	4	30	37.5	0.8	0.8	HEAVYTRUCK	2	2	3
4	0	15	37.5	1.1	1.1	BUS	60	2	3
5	0	15	37.5	1.1	1.1	BUS	60	2	3
8	0	30	26.4	4	4	LIGHTRAIL	250	1	1
9	0	100	20	1.4	3.2	RAPIDRAIL	1200	0.06	0.06
10	0	100	20	1.4	3.2	RAPIDRAIL	1200	0.06	0.06
11	0	7.5	37.5	7.5	7.5	HOV3	11	2	3

METHOD	MIN_DWELL	MAX_DWELL
PARALLEL	0	0
PARALLEL	0	0
PARALLEL	0	0
PARALLEL	0	0
PARALLEL	0	0
PARALLEL	0	0
PARALLEL	0	0
PARALLEL	0	0
SERIAL	0	30
SERIAL	0	30
PARALLEL	5	30
PARALLEL	5	60
PARALLEL	5	60
PARALLEL	0	0

../demand/trips.txt (extrait)

HHOLD	PERSON	TRIP	PURPOSE	MODE	VEHICLE	START	ORIGIN	ARRIVE	DESTINATION
1	1	1	1	3	0	08:00:00	81	29:00:00	81
2	1	1	1	3	0	08:05:00	137	29:00:00	81
3	1	1	1	3	0	08:10:00	136	29:00:00	81
4	1	1	1	3	0	08:15:00	135	29:00:00	81
5	1	1	1	3	0	08:20:00	134	29:00:00	81
6	1	1	1	3	0	08:25:00	133	29:00:00	81
7	1	1	1	3	0	08:30:00	132	29:00:00	81
8	1	1	1	3	0	08:35:00	131	29:00:00	81
9	1	1	1	3	0	08:40:00	64	29:00:00	81
10	1	1	1	3	0	08:45:00	129	29:00:00	81