UNIVERSITÉ DE MONTRÉAL

TRAITEMENT DE DONNÉES GPS ET DE COMPTAGE EN TRANSPORT EN COMMUN

DANIEL RIENDEAU

DÉPARTEMENT DES GÉNIES

CIVIL, GÉOLOGIQUE ET DES MINES

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE CIVIL)

AOÛT 2004

© Daniel Riendeau, 2004

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-97978-4
Our file Notre référence
ISBN: 0-612-97978-4

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

TRAITEMENT DE DONNÉES GPS ET DE COMPTAGE EN TRANSPORT EN COMMUN

présenté par : RIENDEAU Daniel

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

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

M. BAASS Karsten, Ph.D., president

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

M. LAVIGUEUR Pierre, M.Sc.A., membre

REMERCIEMENTS

L'auteur tient à remercier en premier lieu le professeur Robert Chapleau pour son soutien technique et académique, sa grande disponibilité, ses encouragements et, surtout, pour lui avoir donné l'occasion de travailler dans un domaine aussi enrichissant.

L'auteur est reconnaissant envers le Conseil de recherches en sciences naturelles et en génie (CRSNG), la Société de transport de la Ville de Laval (STL) et le professeur Robert Chapleau pour leur support financier.

Des remerciements vont à Pierre Lavigueur et Christine Gauvreau, respectivement directeur et conseillère en planification à la STL, pour leur courtoisie et leur soutien, pour avoir accordé un accès privilégié à leurs bases de données et pour avoir rendu possibles les expériences « sur le terrain ».

Des remerciements aussi à Catherine Morency et Bruno Allard du groupe MADITUC pour leur aide, ainsi qu'à Michel Prévost et Corinne Thomas du Réseau de transport de la Capitale (RTC) et à Daniel Bergeron et Sébastien Gagné de l'Agence métropolitaine de transport (AMT) pour leur amabilité et leurs informations pertinentes.

Finalement, l'auteur aimerait remercier tous ses collègues de travail pour l'avoir accompagné pendant la réalisation de son projet et pour les souvenirs mémorables qui en ont résulté.

RÉSUMÉ

Presque tous les transports de surface font dorénavant appel aux dispositifs GPS pour acquérir une information spatio-temporelle fiable sur les évolutions des véhicules d'une flotte de transport. Dans le contexte des réseaux de transport urbain collectif, la collecte des données est souvent couplée à des données d'achalandage montées-descentes et, pour des raisons d'économie de stockage, est surtout événementielle.

À ce stade, par des moyens hybrides (logiciels de traitement primaire et interactivité humaine) ainsi que par fusion avec des données de planification, on arrive à produire des ensembles de données qui permettent un bon suivi de situations particulières.

Toutefois, de nombreux problèmes reliés à la mesure du temps, au décompte des voyageurs, ainsi qu'au caractère événementiel des données et au processus de couplage avec les données de planification amènent des complexités de traitement qui entachent la capacité des données à être intégrées dans un système d'information plus complet.

Ce projet tente de définir les éléments de base d'un système spatio-temporel saisonnalisé de la demande de transport en commun par une organisation structurée et cumulative du traitement des données GPS et des comptes. Les données rendues disponibles par la STL, pour deux mois d'opération, sont examinées sous diverses perspectives, prétextes à quelques représentations graphiques opportunes : analyse compararative des véhicules, traces géographiques, analyse géostatistique des montants-descendants, profils de charge 2D et 3D en termes de tournées ou de temps.

ABSTRACT

Most ground transportation companies nowadays rely on GPS devices to get accurate space-time information in order to efficiently manage their fleets. In the context of urban transit networks, the positioning data is often acquired alongside the automated passenger counting (APC) data and is usually triggered by specific events in order to save memory space.

From this point, by using hybrid solutions (i.e., simple processing softwares combined with manual operations) and by merging GPS data with planned transit information, it is possible to create certain sets of data that enable a good monitoring of particular situations.

However, several problems arise from the measurement of time and space, the passenger counting devices, the non-continuous ("evently") nature of the acquired data as well as the merging process itself. Those problems prevent the data from being seamlessly integrated into a more comprehensive informational system.

This project attempts to define the basic elements of a space-time seasonal system of transit demand measurement based on a structured and cumulative organization of data from positioning and counting devices. Data collected during two months by the Société de transport de Laval (STL) are examined under various viewpoints, which are translated into relevant graphical representations such as vehicle comparisons, tracking points, geostatistical analysis of ins and outs, 2D and 3D load profiles with respect to time or bus passages.

TABLE DES MATIÈRES

REMERCIEMENTS ..... **iv**
RÉSUMÉ ..... **v**
ABSTRACT ..... **vi**
TABLE DES MATIÈRES ..... **vii**
LISTE DES TABLEAUX ..... **xi**
LISTE DES FIGURES ..... **xii**
LISTE DES ANNEXES ..... **xviii**
CHAPITRE 1 : Introduction **1**
1.1 Objectif 2
1.2 Contenu 2
1.3 Enjeux 4
CHAPITRE 2 : Revue de littérature **5**
2.1 Historique 5
2.2 Systèmes APC 6
2.2.1 Infodev 7
2.2.2 Init 10
2.2.3 Iris 11
2.3 Systèmes de localisation 13
2.3.1 Système de positionnement global 14
2.4 Systèmes d'information en temps réel 16
2.4.1 NextBus 16
2.5 Map Matching 18
2.6 Approche orientée-objet 18
2.6.1 Définitions 19
2.7 Origine des données utilisées dans ce projet 22
2.7.1 Société de transport de la Ville de Laval 22
CHAPITRE 3 : Collecte et traitement des données **24**
3.1 Introduction 24

3.2 Collecte des données 24
3.2.1 Mode de collecte 25
3.2.2 Les événements 25
3.3 Traitement des données à la STL 27
3.4 Organisation 29
3.5 Accès aux données 31
3.5.1 Variables de référence 31
3.6 Remarques 32
3.6.1 Collecte des données 32
3.6.2 Traitement des données 33
CHAPITRE 4 : Analyse de la validité des données 35
4.1 Identification des circuits 35
4.2 Identification des tournées 37
4.3 Couverture du réseau 38
4.3.1 Profil de charge 3D 41
4.4 Comptage des passagers 43
4.5 Données d'avril 2003 44
4.6 Balises virtuelles 48
4.6.1 Détection des balises virtuelles 49
CHAPITRE 5 : Expériences sur le terrain 53
5.1 Première expérience 53
5.1.1 Distribution des données 57
5.1.2 Profils de charge 58
5.2 Calcul des erreurs de comptage 59
5.3 Deuxième expérience 60
5.3.1 Résultats 62
5.4 Troisième expérience 65
5.4.1 Résultats 67
5.5 Conclusion 70
CHAPITRE 6 : Visualisation des données pour fins d'interprétation 72

6.1 Graphiques à bulles 72
6.1.1 Achalandage sur un parcours 72
6.1.2 Circonstances temporelles d'un parcours 73
6.2 Diagrammes espace-temps 75
6.3 Réalité virtuelle 75
CHAPITRE 7 : Reconnaissance des tournées 78
7.1 Introduction 78
7.2 Problématique 79
7.3 Principes 81
7.4 Sources d'information 84
7.4.1 Liste des nœuds 84
7.4.2 Liste des lignes de transport 85
7.4.3 Liste des horaires 85
7.4.4 Liste des jours fériés 85
7.5 Méthode 85
7.5.1 Identification des nœuds 86
7.5.2 Identification des circuits associés aux noeuds 86
7.5.3 Vérification des circuits et élimination 87
7.5.4 Identification des tournées potentielles 89
7.5.5 Sélection des tournées par élimination 90
7.6 Résultats 91
7.6.1 Comparaisons 92
7.6.2 Représentativité des données 94
7.7 Discussions 97
7.7.1 Améliorations possibles 97
7.7.2 Objectif (rappel) 97
CHAPITRE 8 : Détection de problèmes de comptage 99
8.1 Charges résiduelles 99
8.1.1 Boucles 103
8.2 Indices de fiabilité 103

8.2.1 Données d'avril 2003 110
8.3 Vue d'ensemble 112
CHAPITRE 9 : Modélisation saisonnalisée de la demande 115
9.1 Agrégation temporelle 116
9.2 Validation des données 118
9.3 Synthèse des données 119
9.4 Interpolations 121
9.5 Regroupement des tournées en intervalles 124
9.6 Quelques profils de charge 3D 129
CHAPITRE 10 : Conclusion 133
10.1 Comptage automatique des passagers 133
10.2 Reconnaissance des tournées 134
10.3 Autres considérations 136
RÉFÉRENCES 138
ANNEXES 142

LISTE DES TABLEAUX

Tableau 3-1 : Données collectées par les autobus de la STL munis de GPS 27
Tableau 4-1 : Échantillonage obtenu pour la ligne 28 nord en avril 2002 42
Tableau 5-1 : Intervalles entre chaque collecte de données par un GPS portatif 57
Tableau 7-1 : Établissement des horaires et assignations à partir des circuits parcourus 91
Tableau 7-2 : Tournées d'avril 2003 observées manuellement 93
Tableau 8-1 : Achalandage mesuré à bord de l'autobus SL6213 le 24 avril 2002 106
Tableau 8-2 : Achalandage mesuré à bord de l'autobus SL0012 le 15 avril 2002 106
Tableau 8-3 : Nombre moyen de montées et de descentes par ouverture de portes 108
Tableau 9-1 : Exemple de liste des échantillons classés par tournées 117
Tableau 9-2 : Variation de la charge du circuit 27 nord au terminus Laval (avril 2003) 126

LISTE DES FIGURES

Figure 1-1 : Résumé des enjeux reliés à la collecte et à l'analyse de données en transport 4
Figure 2-1 : Autobus muni d'équipements Infodev 8
Figure 2-2 : Interface graphique du logiciel BusStops 9
Figure 2-3 : Liste des produits d'Init 11
Figure 2-4 : Logo du détecteur d'Iris (« Miss IRMA ») 12
Figure 2-5 : Détecteur de mouvement directionnel IRMA 12
Figure 2-6 : Répartition des satellites GPS 14
Figure 2-7 : Contrôle au sol des satellites GPS 15
Figure 2-8 : Erreurs courantes dans le positionnement par satellites 16
Figure 2-9 : Page d'accueil de NextBus 17
Figure 2-10 : Aperçu de la notation utilisée en approche orientée-objet 19
Figure 2-11 : Exemples de relations "1 à *x*" 21
Figure 2-12 : Exemple de collections 21
Figure 3-1 : Autobus équipé de systèmes GPS et de comptage 24
Figure 3-2 : Mécanisme de collecte événementielle des données 26
Figure 3-3 : Fichier de données brutes (ouvert dans Excel) 28
Figure 3-4 : Couplage classique des données à la STL 28
Figure 3-5 : Diagramme schématique d'une base de données pour une société de transport 30
Figure 3-6 : Sélection d'un fichier de données brutes 31
Figure 3-7 : Traitement des données au RTC 34
Figure 4-1 : Distance parcourue par les systèmes Infodev en avril 2002 35
Figure 4-2 : Proportion de la distance identifiée par BusStops 36
Figure 4-3 : Quelques parcours faussement identifiés comme circuits 37
Figure 4-4 : Circuits reconnus avec un horaire validé par LADRA 38
Figure 4-5 : Représentativité des données collectées pendant avril 2002 39
Figure 4-6 : Utilisation des systèmes Infodev pendant avril 2002 40

Figure 4-7 : Profil de charge 3D - ligne 28 nord, avril 2002 41
Figure 4-8 : Exemples de profils de charge erronés 43
Figure 4-9 : Circuits reconnus avec un profil de charge valide 44
Figure 4-10 : Circuits reconnus avec un horaire et un profil de charge valides 44
Figure 4-11 : Distance parcourue par les systèmes Infodev en avril 2003 45
Figure 4-12 : Circuits reconnus avec un horaire et un profil de charge valides (avril 2003) 46
Figure 4-13 : Représentativité des données collectées pendant avril 2003 47
Figure 4-14 : Utilisation des systèmes Infodev pendant avril 2003 48
Figure 4-15 : Localisation des arrêts d'autobus et des balises virtuelles de la STL 49
Figure 4-16 : Nombre moyen de balises détectées par circuit 49
Figure 4-17 : Taux de détection de balise pour chaque autobus 50
Figure 4-18 : Taux de détection de chaque balise virtuelle 50
Figure 4-19 : Occurrences de la balise 5500 (circuit 55, près du parc Légaré) 51
Figure 4-21 : Occurrences de la balise 236 52
Figure 4-20 : Occurrences de la balise 9105 52
Figure 5-1 : Parcours effectué pendant l'expérience 1 réalisée le 18 juillet 2003 54
Figure 5-2 : Expérience 1 – données collectées par l'ordinateur embarqué 55
Figure 5-3 : Expérience 1 – données collectées par un appareil GPS portatif 56
Figure 5-4 : Distribution spatiale et temporelle des données prises par le GPS portatif 57
Figure 5-5 : Distribution spatio-temporelle des données prises par le GPS portatif 58
Figure 5-6 : Expérience 1 - profils de charge (détectés par l'ordinateur embarqué) 58

Figure 5-7 : Parcours effectué pendant l'expérience 2 réalisée le 14 avril 2004 61
Figure 5-8 : Expérience 2 – données collectées par l'ordinateur embarqué 62
Figure 5-9 : Expérience 2 - montées, descentes et charge réelles à chaque arrêt (comptage manuel) 63
Figure 5-10 : Expérience 2 - montées et de descentes détecté et réel à chaque arrêt 63
Figure 5-11 : Expérience 2 - profils de charge (pointillé = détecté, continu = réel) 64
Figure 5-12 : Parcours effectué pendant l'expérience 3 réalisée le 13 mai 2004 66
Figure 5-13 : Expérience 3 – données collectées par l'ordinateur embarqué 66
Figure 5-14 : Expérience 3 - montées, descentes et charge réelles à chaque arrêt (comptage manuel) 67
Figure 5-15 : Expérience 3 - montées et de descentes détecté et réel à chaque arrêt 68
Figure 5-16 : Expérience 3 - profils de charge (pointillé = détecté, continu = réel) 69
Figure 5-17 : Expériences 2 et 3 – données collectées par un appareil GPS portatif 70
Figure 6-1 : Achalandage sur les circuits effectués pendant l'expérience 1 72
Figure 6-2 : Extrait modifié d'un fichier de données LADRA ouvert dans Excel 73
Figure 6-3 : Durée des arrêts sur le parcours effectué pendant l'expérience 1 74
Figure 6-4 : Durée des arrêts sur le parcours effectué pendant l'expérience 2 74
Figure 6-5 : Expérience 1 - diagramme espace-temps 75

Figure 6-6 : Représentation en VRML du parcours de l'expérience 1 (z = temps) 76
Figure 6-7 : Représentation en VRML du parcours de l'expérience 1 (z = heure de départ) 76
Figure 7-1 : Exemple de trace GPS 78
Figure 7-2 : Exemple de réseau de transport 78
Figure 7-3 : Exemple de fusion des données GPS avec les données planifiées 79
Figure 7-4 : Procédure de reconnaissance des tournées 80
Figure 7-5 : Relations lien-noeuds routiers et ligne-noeuds TC 81
Figure 7-6 : Relations ligne TC - lien routier 82
Figure 7-7 : Reconnaissance des tournées avec les nœuds TC 82
Figure 7-8 : Reconnaissance des tournées avec les liens routiers 83
Figure 7-9 : Reconnaissance des tournées par les assignations 83
Figure 7-10 : Reconnaissance des tournées par un lien avec le système d'affichage 83
Figure 7-11 : Reconnaissance des tournées - vue d'ensemble 84
Figure 7-12 : Exemple d'assignations à la STL pour un chauffeur 90
Figure 7-13 : Nombre total de circuits identifiés par la STL en avril 2002 (selon ITTC) 92
Figure 7-14 : Nombre total de circuits effectués en avril 2002 (selon ITTC) 93
Figure 7-15 : Nombre de tournées identifiés par LADRA et ITTC en avril 2002 94
Figure 7-16 : Représentativité des données traitées avec LADRA et ITTC en avril 2002 95
Figure 7-17 : Nombre de tournées identifiés par LADRA et ITTC en avril 2002 95
Figure 7-18 : Représentativité des données traitées avec LADRA et ITTC en avril 2003 96
Figure 8-1 : Exemple de profil de charge non-équilibré 101

Figure 8-2 : Exemple de correction de profil de charge 102
Figure 8-3 : Exemple de profil de charge corrigé 102
Figure 8-4 : Circuits aller-retour contenant une boucle à une extrémité 103
Figure 8-5 : Profils de charge mesurés de la ligne 24 sud à 6h59, 7h07 et 7h21 104
Figure 8-6 : Achalandage mesuré à bord de l'autobus SL6213 pendant un jour 105
Figure 8-7 : Achalandage mesuré à bord de l'autobus SL0011 pendant un jour 105
Figure 8-8 : Moyenne et écart-type des charges résiduelles des 19 autobus en avril 2002 108
Figure 8-9 : Écart moyen des 19 autobus en avril 2002 109
Figure 8-10 : Achalandage moyen détecté à chaque arrêt par les 19 autobus en avril 2002 109
Figure 8-11 : Achalandage mesuré à bord de l'autobus SL0005 pendant un jour 110
Figure 8-12 : Moyenne et écart-type des charges résiduelles des 39 autobus en avril 2003 111
Figure 8-13 : Écart moyen des 39 autobus en avril 2003 111
Figure 8-14 : Achalandage moyen détecté à chaque arrêt par les 39 autobus en avril 2003 111
Figure 8-15 : Vue d'ensemble des données collectées pendant avril 2003 113
Figure 8-16 : Vue d'ensemble des données collectées du 6 au 12 avril 2003 114
Figure 9-1 : Profils de charge du circuit 27 nord à 17h26 119
Figure 9-2 : Profils de charge du circuit 27 nord à 24h05 120
Figure 9-3 : Profils de charge filtrées du circuit 27 nord à 24h05 121
Figure 9-4 : Profils de charge du circuit 27 nord à 14h05 121
Figure 9-5 : Tracé du circuit 39 sud de la STL 123
Figure 9-6 : Variations du profil de charge de la ligne 39 sud 122

Figure 9-7 : Variation de la charge à différents endroits sur le circuit 39 sud 123
Figure 9-8 : Variation de l'achalandage à différents endroits sur le circuit 25 sud (2002) 124
Figure 9-9 : Tracé du circuit 27 nord de la STL 124
Figure 9-10 : Profil de charge 3D du circuit 27 en fonction des tournées – vue de dessus 125
Figure 9-11 : Profil de charge 3D du circuit 27 en fonction des tournées – vue en plongée 125
Figure 9-12 : Variation de la charge du circuit 27 nord au terminus Laval (avril 2003) 127
Figure 9-13 : Profil de charge 3D du circuit 27 nord en fonction du temps – vue de dessus 127
Figure 9-14 : Profil de charge 3D du circuit 27 nord en fonction du temps – vue en plongée 128
Figure 9-15 : Profil de charge 3D du circuit 27 sud – vue de dessus 128
Figure 9-16 : Profils de charge 3D du circuit 39 aller-retour 129
Figure 9-17 : Tracé du circuit 55 130
Figure 9-18 : Profils de charge 3D du circuit 55 aller-retour 130
Figure 9-19 : Tracé du circuit 60 131
Figure 9-20 : Profils de charge 3D du circuit 60 aller-retour 131
Figure 9-21 : Tracé du circuit 70 132
Figure 9-22 : Profils de charge 3D du circuit 70 aller-retour 132

LISTE DES ANNEXES

ANNEXE 1 : fichiers de données 142
Données brutes 142
Données BusStops et LADRA 143
ANNEXE 2 : algorithme de visualisation de tournées 146
ANNEXE 3 : Identificateur de trajets en transport collectif 150
Logiciel 150
Données 155

CHAPITRE 1 : Introduction

Le temps où l'on devait compter sur des humains pour mesurer et apprécier bien imparfaitement les évolutions spatio-temporelles, tant des véhicules que des clients du transport collectif urbain, n'est pas si loin. En effet, les comptes stationnaires aux « points de charge maximale » et les comptes à bord de véhicules ont exigé un fort tribut de pointeurs qui, à la longue, devenaient fort peu fiables ou trop prévisibles – ou les deux. Force est d'admettre que, les nouvelles technologies devenant disponibles à un coût raisonnable, la volonté d'obtenir des données objectives sur la qualité du service offert a amené quelques sociétés de transport au Québec à équiper une portion de leur flotte d'autobus de microprocesseurs embarqués reliés à des GPS et des systèmes de comptage.

Il faut souligner entre autres les expérimentations du Réseau de transport de la Capitale (RTC), du Réseau de transport de Longueuil (RTL) et de la Société de transport de la ville de Laval (STL). Selon les premières préoccupations manifestées, l'accent a été mis d'abord sur la reconnaissance des parcours (RTL, 2001), puis sur la validation de la ponctualité (LAVIGUEUR, 2001), pour ensuite devoir être porté sur des questions d'achalandage.

Cette technologie est expérimentée au Québec depuis plusieurs années sans que l'on semble en tirer quelque chose de plus substantiel. Cette affirmation est peut-être gratuite, mais il se trouve que les promesses (prétentions ou anticipations) de cette technologie, qui devait faire partie des STI (systèmes de transport intelligent), ne comportent que bien peu de réalisations tangibles.

1.1 Objectif

Ce document se veut une introduction à la collecte, au traitement et à l'analyse des données GPS et de comptage en transport en commun tel qu'il est pratiqué actuellement par certaines sociétés de transport utilisant les systèmes automatiques de localisation et de comptage. Ces données sont un outil fort avantageux pour déterminer l'offre de service d'une société de transport; plus précisément, elles permettent d'évaluer l'achalandage sur une partie ou la totalité du réseau ainsi que la qualité du service offert.

Ce projet ne se lance pas dans de telles analyses mais aborde plutôt la question de l'efficacité matérielle et informatique des systèmes utilisées pour obtenir ces données tout en faisant quelque peu la lumière sur les divers obstacles limitant la disponibilité des données et les erreurs techniques remettant en question l'intégrité des résultats obtenus.

Les problématiques abordées vont de la collecte événementielle des données à l'élaboration d'un système spatio-temporel saisonnalisé de l'évaluation de la demande, en passant par les divers problèmes de comptage, le couplage avec les données planifiées de trajets et d'horaire et l'interprétation des résultats par visualisation graphique. Plusieurs solutions sont proposées le long de cet ouvrage, bien que son but principal soit surtout de mettre en évidence, de façon imagée, les problèmes actuels.

1.2 Contenu

Le chapitre suivant présente une revue de littérature sur certains concepts de base et quelques joueurs importants dans le domaine du comptage automatique, de la localisation spatiale et du traitement de données. Le chapitre 3 explique en détail le fonctionnement des systèmes GPS et de

comptage à la STL, suivi au chapitre 4 d'une première analyse des données qui ont été collectées par la STL en avril 2002 et en avril 2003.

Les résultats de quelques expériences à bord d'autobus de la STL sont montrés au chapitre 5, lesquels sont prétextes à diverses représentations graphiques, dont une brève incursion dans la réalité virtuelle, au chapitre 6.

Le chapitre 7 aborde la problématique du couplage des données GPS avec les données planifiées décrivant le réseau et les horaires, et esquisse une méthode de couplage permettant d'obtenir de meilleurs résultats que ceux obtenus jusqu'ici. Le chapitre 8 aborde la problématique du comptage des passagers et présente quelques techniques pour détecter les problèmes de fonctionnement des systèmes embarqués.

Finalement, le chapitre 9 présente brièvement une méthode de discrétisation des données dans le but d'obtenir une description complète de l'achalandage pour une ligne de transport.

1.3 Enjeux

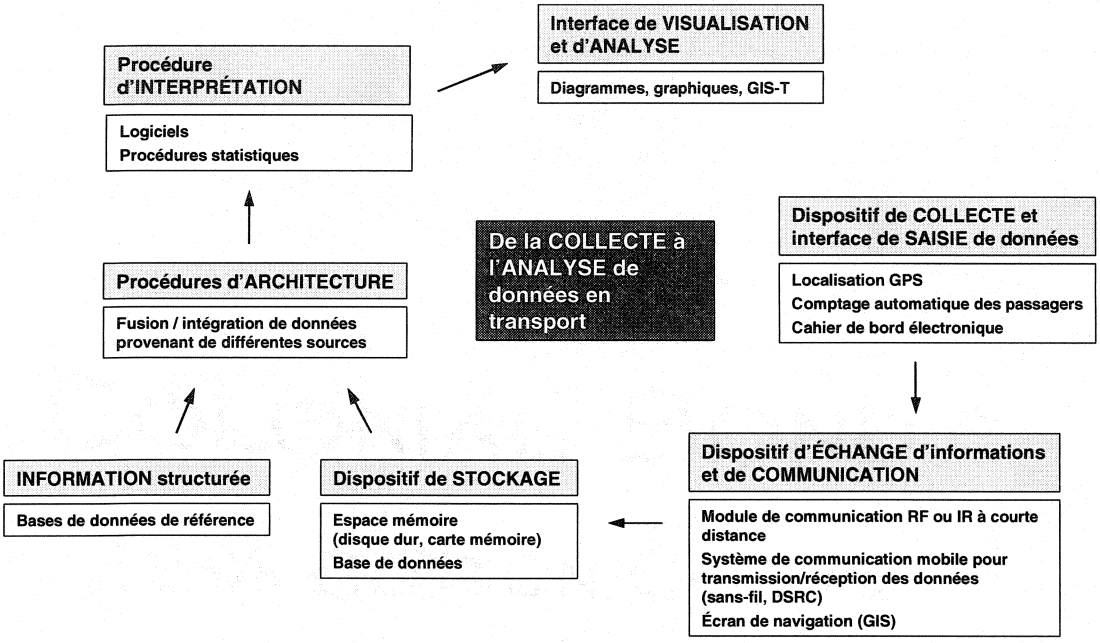


Figure 1-1 : Enjeux reliés à la collecte et à l'analyse de données en transport

(Inspiré de MORENCY, 2001)

CHAPITRE 2 : Revue de littérature

2.1 Historique

Il y a moins de 15 ans, comptes stationnaires et comptes-à-bord étaient en vigueur. Un exemple typique est le préposé à bord d'un véhicule (un *pointeur*) disposant d'un formulaire contenant la liste des arrêts que le véhicule devait effectuer. Le pointeur devait noter le nombre de montées et de descentes pour chaque arrêt ainsi que le temps correspondant (composante temporelle). Ces informations étaient acheminées à un centre de gestion et, lorsqu'on connaissait la localisation des arrêts (composante spatiale), on pouvait se lancer dans certaines analyses sur le profil de charge, la demande en fonction du lieu et du temps, la ponctualité, etc.

Des résultats similaires pouvaient être obtenus en disposant des pointeurs à différents points de service (lieux d'embarquement et de débarquement, ex. arrêts, gares, terminus, etc.). Ces préposés notaient alors le nombre de montants et de descendants pour chaque autobus qui passait, ainsi que le moment de leur arrivée.

L'informatique et l'électronique ont permis de remplacer ces pointeurs par des systèmes de comptage automatique (*Automated Passenger Counting*, APC) qui permettent l'accumulation continue et à coût minime de données sur l'achalandage. (En parallèle, ces technologies ont aussi permis à certaines compagnies plus conservatrices de munir leurs préposés d'ordinateurs de poche pour effectuer leurs comptages).

Pour fournir à ces données les composantes spatiales et temporelles nécessaires à l'analyse, certaines sociétés ont installé des balises à des endroits stratégiques – habituellement des points de service (TRB, 1998).

Cette solution a comme désavantages une résolution limitée et le manque de flexibilité : on ignore se qui se passe entre les balises, et l'augmentation de la résolution par l'ajout de balises est une solution coûteuse. On a ainsi remplacé les pointeurs, mais pas le caractère rigide de la prise des données…

Le système GPS présente donc plusieurs avantages. Un véhicule muni d'un tel système peut être localisé (position géographique) à peu près en n'importe quel lieu et temps, dans la mesure où il n'y a aucun obstacle entre le véhicule et le ciel. On peut ainsi étudier tous les mouvements de véhicules avec une précision de quelques dizaines de mètres et de quelques secondes. Il est maintenant possible d'obtenir des résultats d'analyses beaucoup plus précises, la précision étant maintenant limitée par les ressources informatiques et les erreurs dans les prises de mesures.

Mais justement, lorsque les ressources informatiques sont très limitées (c'était particulièrement le cas il y a quelques années), on doit se limiter aux données les plus importantes. Et les données les plus importantes dans un réseau de transport collectif se trouvent… aux points de service.

2.2 Systèmes APC

Le comptage des passagers est effectué pour de nombreuses raisons parmi lesquelles on retrouve (TRB, 1998) :

- Évaluation de l'achalandage par ligne de transport.
- Évaluation de l'achalandage par tournée.
- Évaluation de l'achalandage sur l'ensemble d'un réseau.
- Mesure des montées et des descentes aux arrêts.
- Mesure du point de charge maximal sur les lignes de transport.
- Évaluation de l'achalandage par segment sur le réseau routier.
- Évaluation de l'achalandage selon la période de la journée.

- Vérification de la ponctualité des autobus.
- Analyse des déplacements origine-destination.

Les méthodes de comptages sont variées mais peuvent être classées en trois catégories (TRB, 1998) :

- Comptage manuel (avec un papier et un crayon, ou un ordinateur portable)
- Système de perception des tarifs (argent/billets ou carte « intelligente »)
- Système automatique (capteurs optiques aux portes, détecteurs sous le plancher, etc.)

Le comptage automatique est de plus en plus populaire et plusieurs compagnies se démarquent dans ce domaine.

2.2.1 Infodev

2.2.1.1 Historique

> *Infodev fut fondé en 1993 lors de la fusion de deux compagnies complémentaires. L'une oeuvrait dans l'acquisition de données physiques et la conception de systèmes électroniques depuis 1985, alors que l'autre concevait des logiciels d'applications depuis 1974. Ce fut le mariage parfait entre l'équipement et les logiciels.* (INFODEV, www.infodev.ca)

À la suite de la fusion, cette entreprise québécoise a consacré un an exclusivement en recherche et développement, effectuant des expériences avec différentes technologies de comptage : rayons latéraux, ultrasons, pression, bruit, champs électriques, micro-ondes, optique passive, empreinte thermique, optique active et images vidéos.

Le capteur le plus apte à compter à l'intérieur de centres commerciaux fut mis sur le marché. Il créa plusieurs précédents :

- Le premier système de comptage directionnel,
- Le premier système ayant un logiciel d'applications basé sur Windows,
- Le premier dispositif de comptage portatif,
- Le premier système sans fil,
- Le premier système avec un support à distance via les télécommunications.

Une opportunité de compter les passagers d'autobus s'est présentée et Infodev a développé des capteurs spécifiques pour cette situation. Ce système a été adopté par plusieurs sociétés de transport canadiennes et, au fil des années, s'est vendu ailleurs dans le monde : États-Unis, Mexique, Europe, Hong-Kong, Australie.

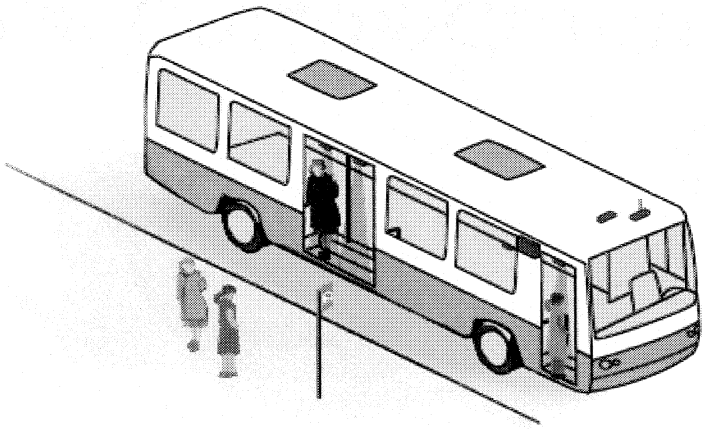


Figure 2-1 : Autobus muni d'équipements Infodev

(Source : INFODEV, www.infodev.ca)

Au Québec, Infodev fournit ses équipements notamment à la Société de transport de la Ville de Laval (STL), au Réseau de transport de Longueuil (RTL) et au Réseau de transport de la Capitale (RTC).

2.2.1.2 Fonctionnement

Les usagers du transport en commun sont détectés lorsqu'ils montent ou descendent d'un autobus à l'aide de capteurs directionnels électro-optiques installés au-dessus des portes. Un ordinateur embarqué, c'est-à-dire installé à bord du véhicule, analyse les signaux des capteurs et enregistre les résultats dans sa mémoire. L'ordinateur enregistre également la localisation du véhicule à l'aide d'une antenne GPS. Les données enregistrées dans l'ordinateur sont téléchargées dans un ordinateur externe à l'aide d'un module de communication radio ou infrarouge, ou à l'aide d'une unité de collecte de données.

Infodev a également conçu le logiciel d'analyse BusStops qui effectue la reconnaissance des circuits effectués par un autobus à l'aide des données collectées et permet à l'opérateur d'effectuer des analyses diverses : tableau noir, distance parcourue, achalandage, vitesses, arrêts…

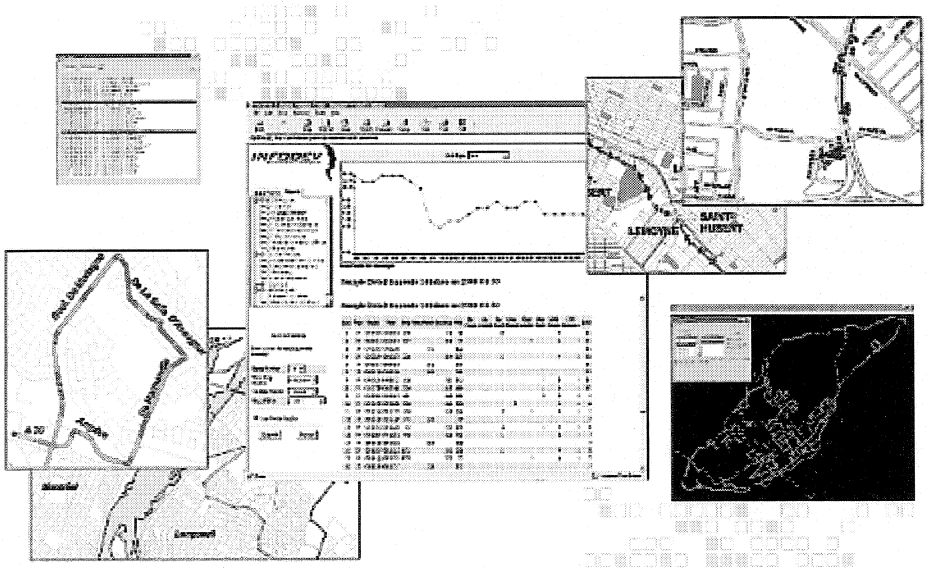


Figure 2-2 : Interface graphique du logiciel BusStops

(Source : INFODEV, www.infodev.ca)

2.2.2 Init

> *Init is one of the world´s technology and market leaders in the area of telematics[1] and electronic fare collection systems for public transport.* (INIT, www.init-ka.de)

La compagnie Init, fondée en 1983 par Gottfried Greschner, est la continuité d'un projet universitaire. En plus de son siège social à Karlsruhe (Allemagne), Init possède des succursales à Berlin et Essen (Allemagne) ainsi qu'une filiale, Init Innovations, basée à Chesapeake (Virginie, États-Unis).

Init se spécialise dans la fabrication de systèmes électroniques de perception des tarifs et de systèmes de transport intelligent (STI) dans le transport en commun. Init conçoit des appareils et des logiciels pour des activités aussi variées que :

- La répartition des véhicules assistée à l'ordinateur
- La gestion du transport « sur demande »
- L'information en temps réel aux usagers
- L'analyse de données et le calcul de statistiques
- La perception des tarifs
- La gestion des garages et dépôts
- Le comptage automatique des passagers
- La communication sans fil
- Les distributeurs de droit de passage
- La préemption aux feux de circulation

Les produits Init sont modulaires, i.e. ils peuvent être utilisés seuls ou agencés entre eux pour former des systèmes plus complets.

---

[1] Le terme « télématique » est utilisé ici comme synonyme de « système de transport intelligent ».

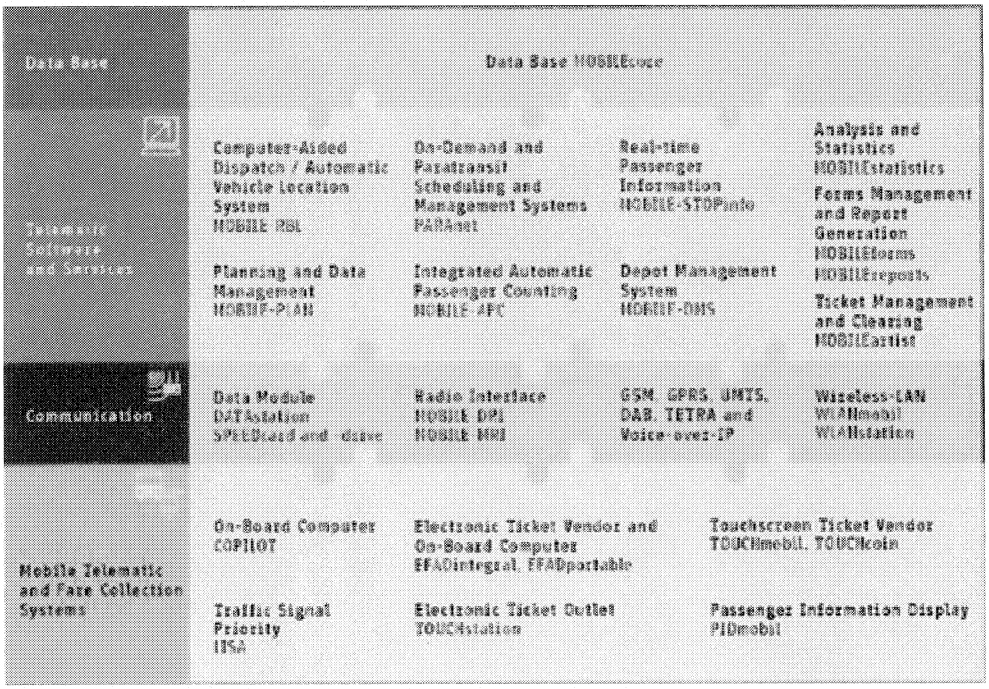


Figure 2-3 : Liste des produits d'Init

(Source : INIT, www.init-ka.de)

L'expertise d'Init est reconnue en Europe et en Amérique du Nord. Au Canada, la Société de transport de Montréal (STM) a muni 171 autobus de systèmes de comptage automatique des passagers.

Les systèmes de comptage automatique d'Init sont fabriqués avec les capteurs de mouvement de la compagnie allemande Iris.

2.2.3 Iris

> *IRIS Infrared & Intelligent Sensors, Inc. is the leading manufacturer of Automatic Passenger Counter (APC) equipment. Successful throughout Europe since 1991, our APC systems have recently made their breakthrough on the North American Market, both in APC Stand-Alone solutions and Intelligent Transportation Systems (ITS).* (IRIS, www.apc-irma.biz)

La compagnie Iris fabrique les détecteurs de mouvement IRMA (Infrared Motion Analyser). Ces détecteurs sont dotés d'un élément passif et d'un élément actif. L'élément actif émet une onde infrarouge et détecte la réflexion causée par une personne ou un obstacle. L'élément passif détecte la radiation thermique émise par le corps humain. La combinaison de ces deux techniques permet d'éliminer les désavantages de chacun et d'assurer une efficacité optimale dans le comptage des passagers.

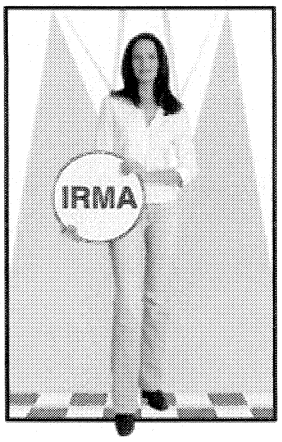


Figure 2-4 : Logo du détecteur d'Iris (« Miss IRMA »)

(Source : IRIS, www.apc-irma.biz)

Le détecteur IRMA est utilisé par de nombreuses sociétés de transport en Europe et en Amérique du Nord.

Selon les compagnies Iris et Init, le détecteur IRMA compte le nombre de passagers à bord avec une précision de 95 %.

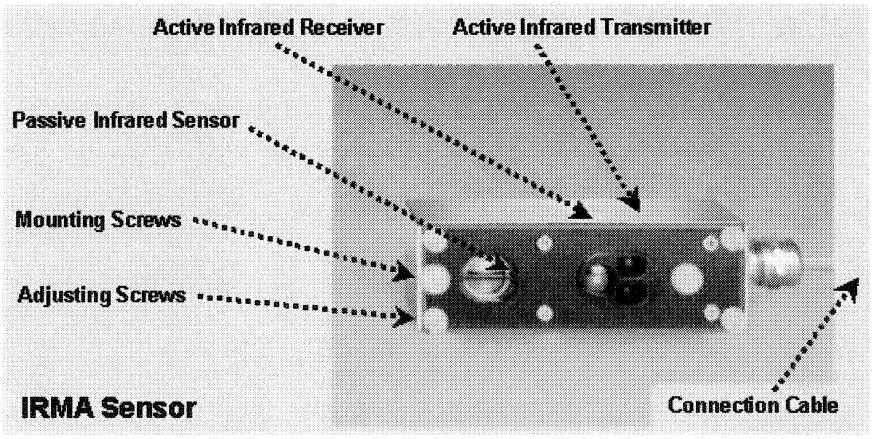


Figure 2-5 : Détecteur de mouvement directionnel IRMA

(Source : IRIS, www.apc-irma.biz)

2.3 Systèmes de localisation

> Robot # 5 : *"MALFUNCTION. NEED INPUT.*
> Howard : - *What does it needs input for?*
> Newton : - *I don't know. I guess it can't triangulate its position.*
> Howard : - *That's a simple function!*
> Newton : - *Can you triangulate* your *position, Howard?*
> Howard : - *No, I can't.*
> Newton : - *Well then, there you go."*
>
> Extrait du film *Short Circuit* (1986)

Les données obtenues par les systèmes de comptage sont incomplètes sans information sur les endroits et les moments où elles ont été obtenues. Certaines entreprises se contentent de ces données afin de vérifier l'achalandage de leur réseau de façon globale, mais à elles seules ces données ne permettront pas d'évaluer la demande spatiale et encore moins d'analyser le comportement de leurs véhicules.

Les systèmes de localisation automatique de véhicule (*Automated Vehicle Location*, AVL) peuvent être divisés en trois familles (DRANE & RIZOS, 1998) :

- Balises (*signposts*)
- Télécommunication (*wave-based*)
- À l'estime ou à l'aveugle (*dead reckoning*)

Le système de positionnement global fait partie de la deuxième famille car il est basé exclusivement sur les ondes radio émises par les satellites. D'un autre côté, la préemption aux feux de circulation (RTL, 2001 et 2004) fait appel à des balises installées sur lesdits feux et à des modules de communication installés à bord des autobus. Finalement, l'utilisation d'un odomètre et d'un compas pour localiser un véhicule est un exemple de système à l'estime rudimentaire.

2.3.1 Système de positionnement global

Pendant les années 1970, l'armée américaine décide de concevoir un système de positionnement par satellites afin de répondre à des besoins de positionnement pour la défense et, à titre de sous-produit, pour servir la communauté civile. Ce système allait être baptisé NAVSTAR the Global Positioning System (PARKINSON & SPILKER, 1996). Les premiers satellites furent lancés en 1978 à des fins expérimentales et le système devint pleinement opérationnel en 1994 (NCHRP, 1998).

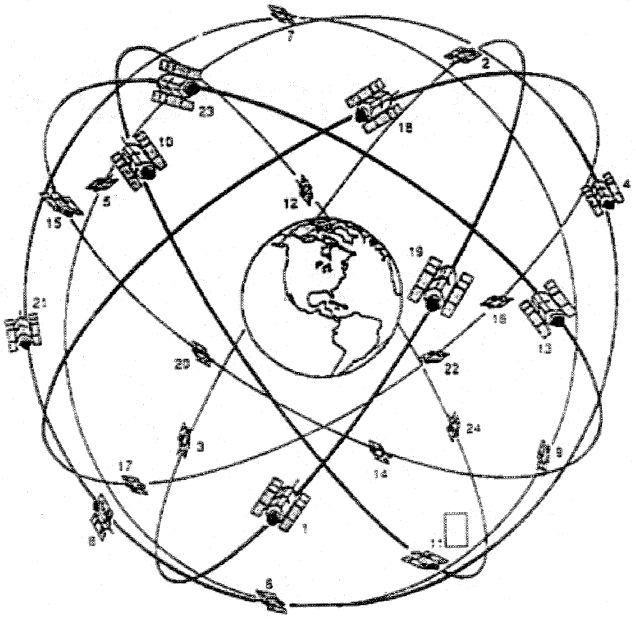


Figure 2-6 : Répartition des satellites GPS

(Source : GÉOMATIQUE CANADA, 1998)

La constellation de satellites comprend 21 satellites, en plus de 3 satellites actifs de rechange, positionnés à 20 000 km (environ trois fois le rayon terrestre) au-dessus de la surface de la Terre. Les satellites sont répartis de telle manière qu'au moins quatre d'entre eux soient visibles presque n'importe où dans le monde à tout moment (Figure 2-6). Chaque satellite reçoit et stocke des données en provenance du segment de contrôle, tient le temps

avec grande précision grâce à ses horloges atomiques précises et transmet des signaux à la Terre (GÉOMATIQUE CANADA, 1998).

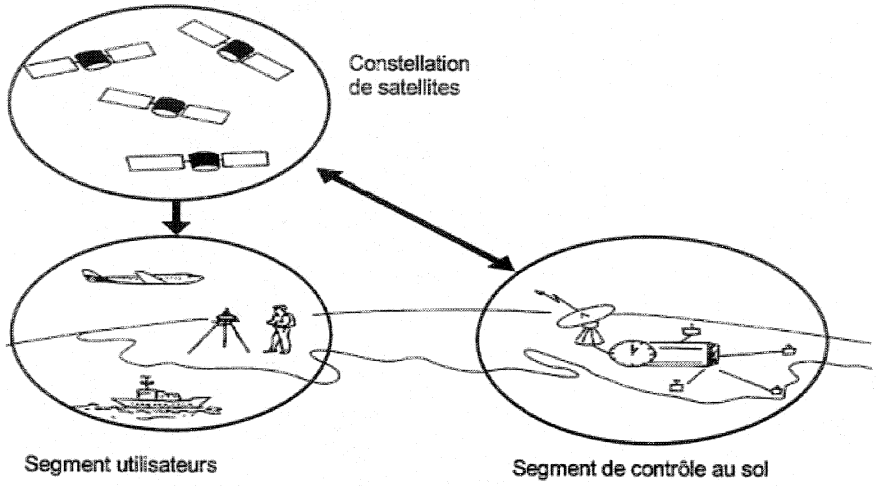


Figure 2-7 : Contrôle au sol des satellites GPS

(Source : GÉOMATIQUE CANADA, 1998)

Le segment de contrôle au sol (Figure 2-7) gère le système de satellites sur une base permanente. Il compte cinq stations de poursuite réparties tout autour de la Terre, dont l'une, située à Colorado Springs, constitue la station maîtresse. Ce segment de contrôle suit tous les satellites, veille à ce qu'ils fonctionnent adéquatement et calcule leurs positions dans l'espace.

De 1991 à 2002, les messages transmis par les satellites contenaient en permanence une dégradation de l'information sur la synchronisation des horloges des satellites. Cette dégradation délibérée, appelée « accès sélectif » (*selective availability*, SA) fut imposée pour des raisons militaires et avait pour but de limiter l'exactitude du positionnement autonome. La précision était alors de 100 m. L'accès sélectif fut aboli en 2002. La précision sans l'accès sélectif est d'environ 20 m (GÉOMATIQUE CANADA, 1998).

Différents types d'erreurs affectent les observations GPS :

- Orbite du satellite (5 à 25 m)
- Ionosphère (1 à 100 m)
- Troposphère (0,5 m)
- Multitrajets [interférences dues à la réflection de signaux par des objets voisins] (0,05 à 20 m)
- Bruits du récepteur (0 à 3 m)

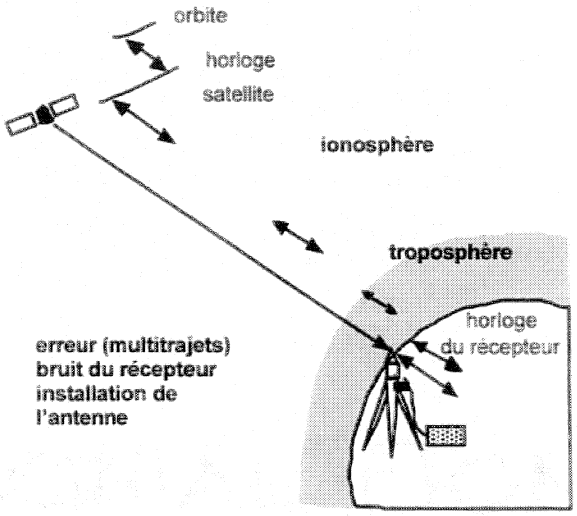


Figure 2-8 : Erreurs courantes dans le positionnement par satellites

(Source : GÉOMATIQUE CANADA, 1998)

2.4 Systèmes d'information en temps réel

2.4.1 NextBus

La technologie NextBus utilise le GPS conjointement à un logiciel de prédiction afin d'informer l'usager du moment d'arrivée de son autobus.

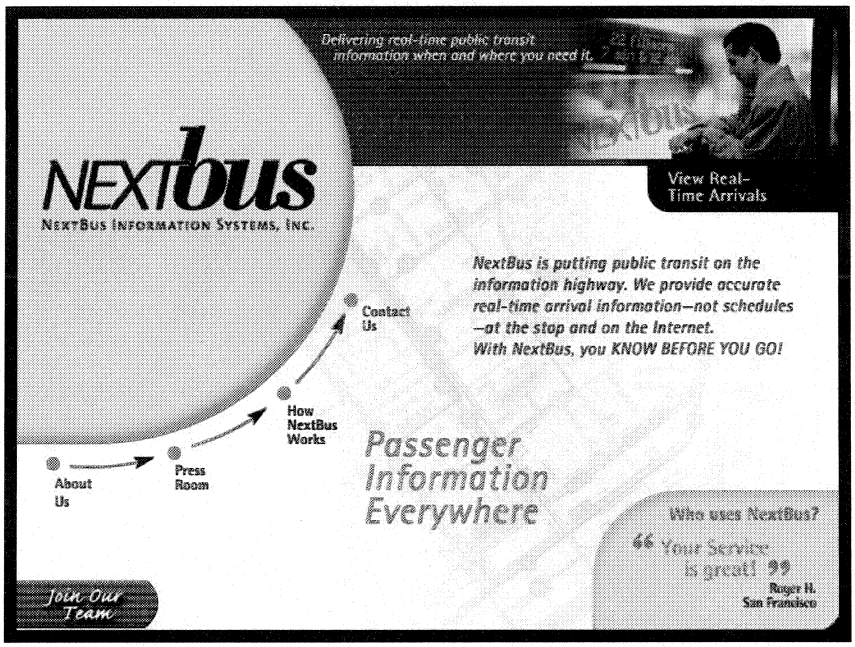


Figure 2-9 : Page d'accueil de NextBus (www.nextbus.com)

> *NextBus overcomes the single greatest barrier to the use of mass transit – unnecessary waiting. Armed with the real-time arrival information that only NextBus can provide you can:*
>
> - *Choose to take another route.*
> - *Run errands in local stores.*
> - *Grab a cup of coffee.* (NEXTBUS, www.nextbus.com)

En tenant compte de la localisation des véhicules de transport, de leurs arrêts prévus ainsi que de l'état de la circulation, le centre d'information NextBus parvient à estimer les heures d'arrivées des véhicules avec un haut niveau de précision.

Ces prédictions sont alors mises en disponibilité sur Internet ainsi que sur les panneaux d'affichages installés aux abribus, terminus, gares et autres lieux publics. Ces informations peuvent aussi être accessibles par des appareils portatifs tels : téléphones cellulaires, ordinateurs de poche, etc.

La compagnie NextBus est située à Emeryville (Californie, Etats-Unis). Sa technologie est utilisée dans plusieurs villes américaines, dont San Francisco, Santa Barbara, Fairfax et Oklahoma City.

2.5 Map Matching

Le *Map Matching* fut d'abord utilisé dans le but d'améliorer la précision des données de positionnement obtenues à bord d'un véhicule. L'hypothèse de base est que le véhicule se trouve sur le réseau routier en tout temps et qu'en connaissant la localisation des routes et des intersections on peut par déduction déterminer beaucoup plus précisément la position du véhicule. Cette technique est très utile sur les systèmes de localisation à l'estime afin de corriger les erreurs de dérive et est aussi parfois utilisée avec un appareil GPS[2]. La principale faiblesse d'une telle procédure est sa tendance à causer des erreurs lorsque le véhicule quitte la route (ex. stationnement, sentier hors-route). Pour plus d'informations à ce sujet, lire ANDERSSON & FJELLSTRÖM (2004) et ZHAO (1997).

Le principe du *Map Matching* est indirectement utilisé dans le cadre de ce projet dans l'élaboration d'une technique de reconnaissance des circuits d'autobus à partir de traces GPS (Chapitre 7).

2.6 Approche orientée-objet

L'utilisation d'une approche orientée-objet en transport constitue une suite logique à l'évolution des techniques d'analyse et de modélisation utilisée par les planificateurs de transport. Le passage à l'ère des systèmes d'information a favorisé l'utilisation de langages informatiques orientés-objet dans le design et l'organisation des logiciels de planification.

---

[2] Avec l'accès sélectif en opération avant 2002, l'idée du *Map Matching* était particulièrement pertinente.

Cette section résume quelques concepts de base de l'approche orienté-objet adaptée au transport, tels que décrits par TRÉPANIER (1999).

2.6.1 Définitions

L'objet de transport est une entité virtuelle, un élément d'information qui est relatif à la modélisation, l'observation, la planification, la simulation et l'analyse des systèmes de transport. Généré à partir de classes de transport, il possède des propriétés et des méthodes propres à l'exercice du transport.

Il existe de nombreuses notations pour illustrer les relations entre objets, tout dépendant de l'approche préconisée. La notation unifiée (*unified*) présentée ici est une nomenclature fondée sur la définition explicite des classes par des boîtes où sont inscrites les propriétés et méthodes et des relations par des droites avec des symboles qui indiquent le nombre d'instances impliquées.

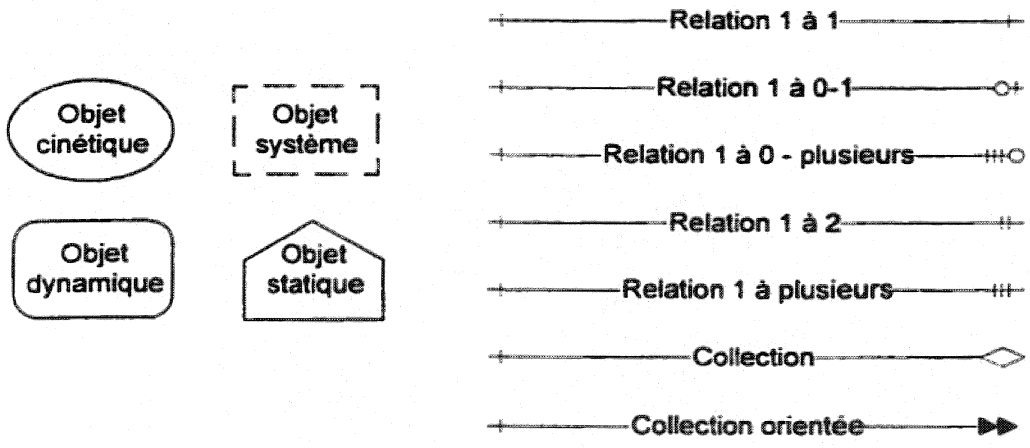


Figure 2-10 : Aperçu de la notation utilisée en approche orientée-objet

(Source : TRÉPANIER, 1999)

2.6.1.1 Objets

Dans un système orienté-objet, chaque objet est généré à partir d'une classe qui lui est propre et qui lui transmet par héritage ses propriétés et ses méthodes. Ces différentes classes sont ici regroupées en quatre métaclasses, ce qui permet de mieux caractériser les types de propriétés et de méthodes qui leur sont associés.

- Objets dynamiques : ce sont les acteurs du transport, ils décident du déplacement ou du moins y participent. L'objet dynamique peut être une personne, un groupe de personnes (ex. un ménage), un véhicule mouvant ou un produit déplacé.
- Objets cinétiques : ce sont les descripteurs et les « canalisateurs » du transport. Ces objets ne décident pas du déplacement mais ils le décrivent. L'objet cinétique peut être simple (lien entre deux points) ou complexe (séquence de points – ex. ligne de transport collectif).
- Objets statiques : ces objets sont normalement fixes dans le temps et supportent les activités de transport. Avec leur topologie bien définie, ils sont faciles à représenter par un système d'information géographique. On retrouve les objets de type point (générateur, arrêt, intersection), ligne (tronçon routier, frontière de zone) et surface (parcs, secteurs, municipalités).
- Objets systémiques : ce sont des ensembles cohérents d'objets élémentaires et des relations qui les lient. On retrouve des ensembles dynamiques (recensement, enquête origine-destination, données GPS), cinématiques (réseau de transport collectif, matrice origine-destination) et statiques (réseau routier, découpage territorial).

2.6.1.2 Relations

Les extrémités des lignes qui symbolisent les relations sont marquées d'un cercle pour le chiffre 0, d'une ou deux lignes pour les chiffres 1 et 2 et de trois lignes pour « 1 ou plusieurs ». Un losange symbolise une collection (ensemble d'objets) non-orientée alors qu'une double flèche symbolise une collection orientée (séquence).

Par exemple chaque ligne de transport relie deux terminus et chaque secteur contient 0, 1 ou plusieurs municipalités (Figure 2-11).

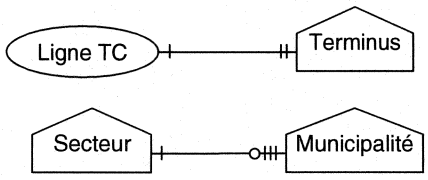


Figure 2-11 : Exemples de relations "1 à *x*"

Par ailleurs, un réseau de transport collectif est un ensemble de lignes de transport et de points de sevice (nœuds). Mais, du même coup, chaque ligne de transport est définie par une séquence de nœuds.

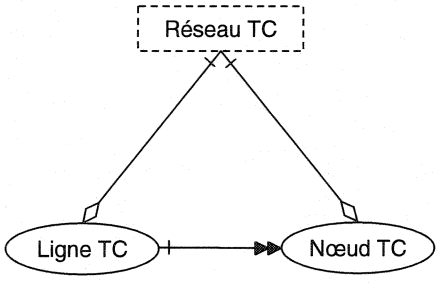


Figure 2-12 : Exemple de collections

2.7 Origine des données utilisées dans ce projet

Les données GPS et de comptage utilisées dans ce document sont fournies gracieusement par la Société de transport de la Ville de Laval (STL), laquelle a également fournie les données décrivant le réseau d'autobus – notamment la localisation des arrêts d'autobus et des *balises virtuelles* ainsi que la description des circuits et des horaires planifiés.

2.7.1 Société de transport de la Ville de Laval

La STL dessert le territoire de la Ville de Laval, « *la deuxième plus grande ville du Québec, d'une superficie de 245,84 km^2 et regroupant une population de près de 350 000 personnes. (Source : Gazette officielle du Québec)* » (STL, www.stl.laval.qc.ca)

2.7.1.1 Description

Son réseau est constitué de 33 circuits bidirectionnels couvrant une distance de 610 kilomètres, et de 2300 arrêts d'autobus dont 281 abribus. Elle dispose d'une flotte de 225 autobus et d'un unique garage. Son effectif est composé de 696 employés dont 463 chauffeurs. Ses deux terminus les plus importants sont le terminus Laval (terminus / station de métro Henri-Bourassa) et le terminus Le Carrefour (Carrefour Laval).

La STL fut créée en 1984 et succéda à la Commission de transport de la ville de Laval (CTL), laquelle fut fondée en 1971. Avant cela, le transport en commun était assuré par trois entreprises privées.

2.7.1.2 Financement

Budget annuel 2004 : 80 167 333 $

Sources de financement :

- Ville de Laval : 56 %
- Clientèle : 28 %
- Gouvernement du Québec : 6 %
- Agence métropolitaine de transport : 3 %
- Autres sources de revenu : 7 %

2.7.1.3 Clientèle

Achalandage annuel : de l'ordre de 18,6 millions d'usagers
(nombre de passages effectués).

Profil des usagers (âgés de 15 ans et plus) :

- 15 à 24 ans : 52 %
- 25 à 34 ans : 15 %
- 35 à 44 ans : 14 %
- 45 à 54 ans : 11 %
- 55 à 64 ans : 5 %
- 65 ans et plus : 3 %

- Hommes : 39 %
- Femmes : 61 %

- Propriétaires : 55 %
- Locataires : 45 %

Taux de pénétration d'utilisation du transport en commun :

27 % de la population lavalloise (15 ans et plus);

42 % des ménages lavallois.

(Source : www.stl.laval.qc.ca/information/coup_doeil.htm)

CHAPITRE 3 : Collecte et traitement des données

3.1 Introduction

Ce chapitre décrit le fonctionnement des systèmes de localisation et de comptage installés à bord de véhicules de transport, ainsi que les différentes étapes menant à la création et l'organisation d'une base de données inclusive et ordonnée.

Le cas décrit ici est celui de la Société de transport de Laval (STL) et les équipements installés à bord des autobus sont ceux d'Infodev.

3.2 Collecte des données

Plusieurs pièces d'équipement sont installées à bord d'autobus afin de permettre la collecte d'informations sur l'achalandage et la localisation. Un système Infodev typique comprend les équipements suivants :

- Un capteur de mouvement directionnel installé au-dessus de chaque porte;
- Une antenne GPS;
- Un ordinateur embarqué;
- Un module de communication radiofréquence ou infrarouge à courte portée.

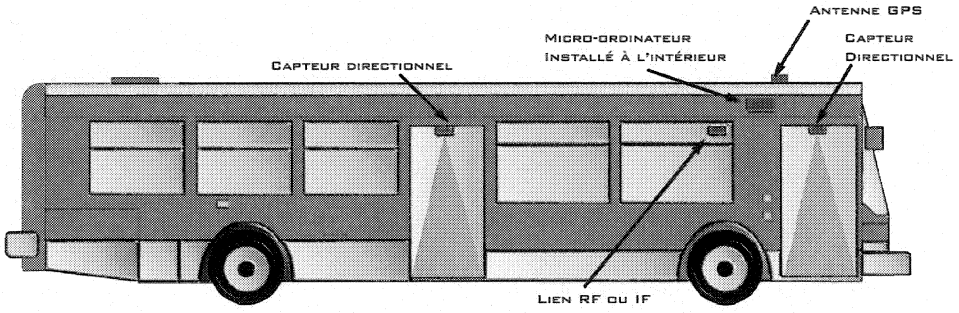


Figure 3-1 : Autobus équipé de systèmes GPS et de comptage

(Source : INFODEV, www.infodev.ca)

L'ordinateur embarqué, et dans certains cas l'antenne GPS, sont cachés du public tandis que les autres pièces sont conçues de façon à passer inaperçues.

3.2.1 Mode de collecte

Les véhicules sont équipés d'un ordinateur pour stocker les données pendant une certaine période, et d'un module de communication à courte distance (radiofréquence ou infrarouge) pour transférer ces informations vers un ordinateur stationnaire.

Les ordinateurs ne peuvent stocker une quantité infinie de données. Cette limite était certainement plus contraignante il y a quelques années lorsque la technologie GPS s'est introduite à bord des autobus. Il fallait alors limiter la collecte des données à l'essentiel, à commencer bien sûr par les points de service, c'est-à-dire les arrêts, les gares, les stations, etc.

L'ouverture des portes déclenche des opérations de collecte : l'ordinateur embarqué stocke en mémoire la localisation du véhicule obtenue par le GPS, le moment exact où les portes s'ouvrent et se referment, ainsi que le nombre de personnes montant et descendant. L'ordinateur obéit donc à un *événement*, et, dans ce cas-ci, l'événement en question s'appelle « ouverture des portes ».

D'autres événements peuvent être programmés pour collecter des données à des moments jugés pertinents. Cette façon de collecter des données est dite *événementielle*.

3.2.2 Les événements

Les ordinateurs de bord de la STL (LAVIGUEUR, 2001), du RTL (RTL, 2001) et du RTC effectuent donc la collecte de données selon un certain nombre d'événements programmés. Ces événements sont notamment : l'ouverture des

portes, l'arrêt du véhicule (halte), le démarrage et l'arrêt du moteur. Le passage à certains endroits précis est également considéré comme un événement. Ces lieux, particularisés par des paires de coordonnées géographiques reconnues par l'ordinateur embarqué, sont appelés balises *virtuelles*.

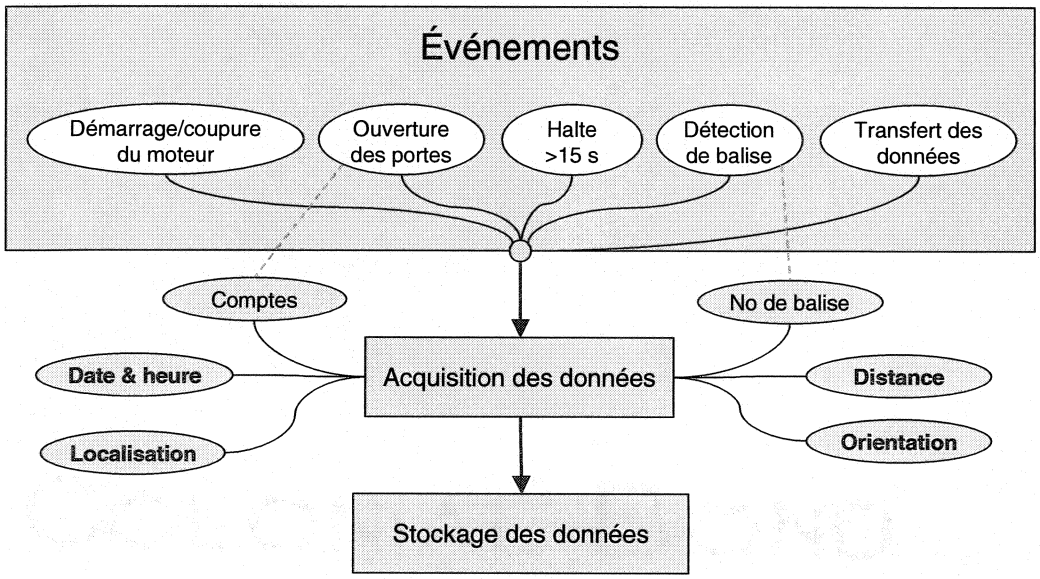


Figure 3-2 : Mécanisme de collecte événementielle des données

Chaque événement permet d'obtenir des informations particulières :

- Ouverture des portes : comptage des montants et des descendants, mesure de la ponctualité aux arrêts.
- Halte (sans ouv. des portes) : durée des arrêts dus à la signalisation et à la circulation routières.
- Passage près d'une balise virtuelle : confirmation que le véhicule passe à un endroit donné, précision sur son emplacement géographique.
- Démarrage/coupure du moteur : moments de mise en service et de mise hors service du véhicule.
- Transfert des données : moment où les données accumulées sont sauvegardées dans un fichier.

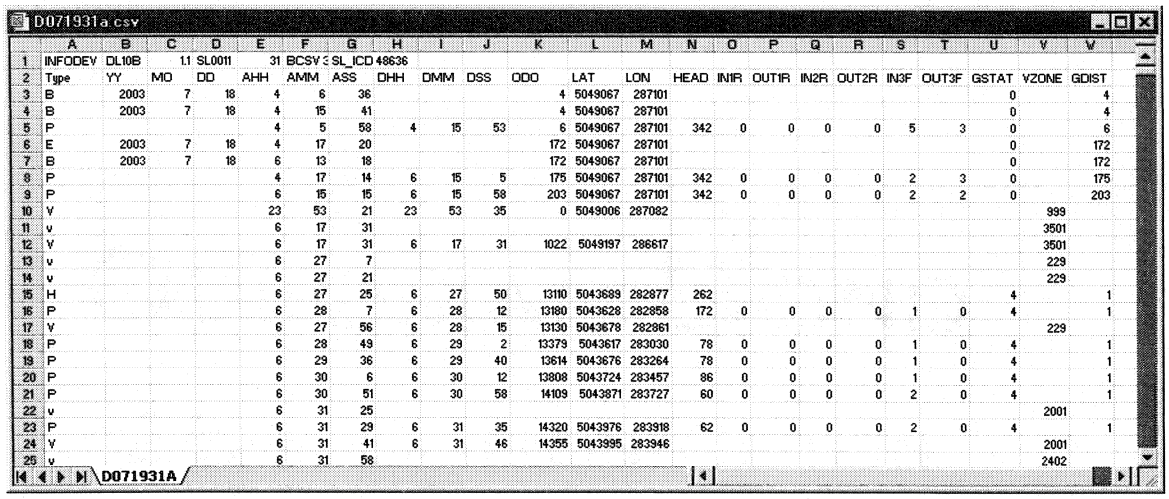
Lorsqu'un événement est détecté, plusieurs données sont stockées en mémoire. Outre le moment (date, heure) et la position géographique donnée par le GPS, des informations comme le relevé de l'odomètre, l'orientation du véhicule, ainsi que quelques données concernant la qualité des informations GPS sont aussi recueillies.

Tableau 3-1 : Données collectées par les autobus de la STL munis de GPS

Information	Événement déclencheur
Type d'événement B = démarrage du moteur E = coupure du moteur V = passage près d'une balise P = ouverture des portes H = halte de plus de 15 s sans ouverture de porte X = transfert des données	
Année, mois, jour	B, E, X
Heure, minute, seconde au début de l'événement	Tous
Heure, minute, seconde à la fin de l'événement	V, P, H, X
Relevé de l'odomètre	Tous sauf X
Latitude, longitude en coordonnées MTM (NAD83)	Tous sauf X
Orientation du véhicule (0 à 359°)	P, H
Nombre de montants et de descendants aux portes arrières et à la porte avant	P
Nombre de satellites GPS détectés	B, E, P, H
Numéro de balise	V

3.3 Traitement des données à la STL

Les données accumulées à bord des véhicules (*données brutes*) sont téléchargées et sauvegardées dans des fichiers identifiés en fonction de la date de sauvegarde et du numéro de véhicule. Ces données sont ensuite fusionnées avec des données de planification et les résultats sont sauvegardés à nouveau dans une autre série de fichiers.



D071931a.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	INFODEV	DL10B		1.1	SL0011		31	BCSV3	SL_ICD	48636													
2	Type	YY	MO	DD	AHH	AMM	ASS	DHH	DMM	DSS	ODO	LAT	LON	HEAD	IN1R	OUT1R	IN2R	OUT2R	IN3F	OUT3F	GSTAT	VZONE	GDIST
3	B	2003	7	18	4	6	36				4	5049067	287101								0		4
4	B	2003	7	18	4	15	41				4	5049067	287101								0		4
5	P				4	5	58	4	15	53	6	5049067	287101	342	0	0	0	0	5	3	0		6
6	E	2003	7	18	4	17	20				172	5049067	287101								0		172
7	B	2003	7	18	6	13	18				172	5049067	287101								0		172
8	P				4	17	14	6	15	5	175	5049067	287101	342	0	0	0	0	2	3	0		175
9	P				6	15	15	6	15	58	203	5049067	287101	342	0	0	0	0	2	2	0		203
10	V				23	53	21	23	53	35	0	5049006	287082									999	
11	v				6	17	31															3501	
12	V				6	17	31	6	17	31	1022	5049197	286617									3501	
13	v				6	27	7															229	
14	v				6	27	21															229	
15	H				6	27	25	6	27	50	13110	5043689	282877	262							4		1
16	P				6	28	7	6	28	12	13180	5043628	282858	172	0	0	0	0	1	0	4		1
17	V				6	27	56	6	28	15	13130	5043678	282861									229	
18	P				6	28	49	6	29	2	13379	5043617	283030	78	0	0	0	0	1	0	4		1
19	P				6	29	36	6	29	40	13614	5043676	283264	78	0	0	0	0	1	0	4		1
20	P				6	30	6	6	30	12	13808	5043724	283457	86	0	0	0	0	1	0	4		1
21	P				6	30	51	6	30	58	14109	5043871	283727	60	0	0	0	0	2	0	4		1
22	v				6	31	25															2001	
23	P				6	31	29	6	31	35	14320	5043976	283918	62	0	0	0	0	2	0	4		1
24	V				6	31	41	6	31	46	14355	5043995	283946									2001	
25	v				6	31	58															2402	

D071931A

Figure 3-3 : Fichier de données brutes (ouvert dans Excel)

L'approche classique de traitement à la STL se fait en deux étapes séparées :

- Fusion des données brutes avec des données sur les circuits planifiés.
- Fusion des données précédentes avec des données sur l'assignation des chauffeurs et les horaires des circuits.

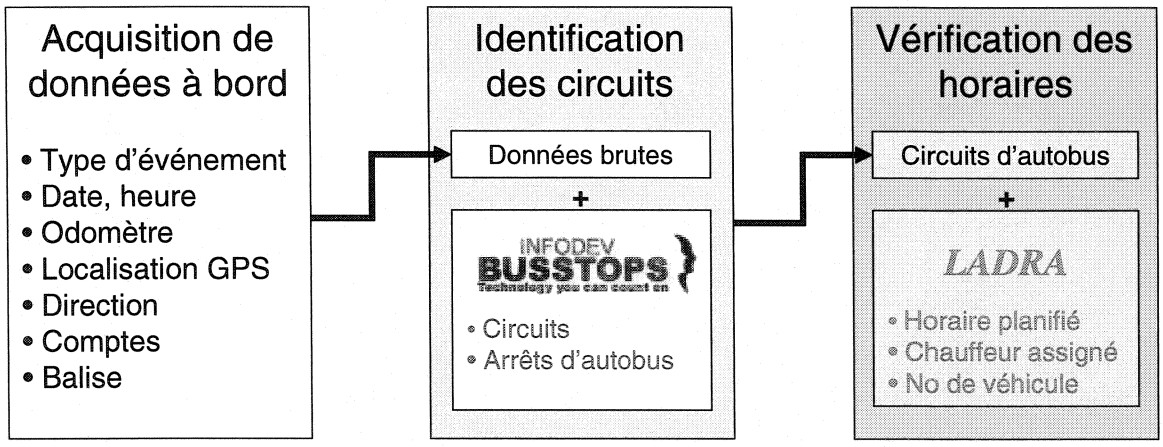


Figure 3-4 : Couplage classique des données à la STL

Typiquement, la première étape se fait avec le logiciel BusStops conçu par la compagnie Infodev. La seconde étape est effectuée à l'aide d'une application adaptée – le logiciel LADRA (Logiciel d'aide au développement du réseau

d'autobus) – avec des données provenant du logiciel d'habillage d'horaires Hastus.

Les fichiers produits par BusStops et LADRA sont eux aussi identifiés selon le numéro du véhicule et la date de sauvegarde des données brutes.

Le découpage des circuits effectué par BusStops est parfois incorrect et le logiciel LADRA ne fait aucune correction d'erreur (DUPUIS, 2000). Lorsqu'une information est incertaine, elle est tout simplement rejetée, et lorsqu'un groupe de données ne semble correspondre à aucun circuit d'autobus ou aucun horaire, les logiciels respectifs laissent des espaces vides. Cela évite que des analyses soient menées sur des informations erronées et conduisent à des conclusions fausses.

C'est pourquoi les données d'achalandage montées-descentes ne sont pas nécessairement associées à ces parcours après le couplage. Des erreurs de mesures ou d'opérations peuvent empêcher le couplage de certaines données.

3.4 Organisation

La gestion d'une base de données de grande taille nécessite certainement une organisation rigide et cohérente. Dans ce cas-ci, deux éléments-clé méritent notre attention : le jour où les données ont été sauvegardées et le véhicule qui les a fournies.

Une procédure de collecte et de sauvegarde des données peut se résumer ainsi :

Pour chaque véhicule…

À chaque n *jour(s)…*

Sauvegarder les données collectées dans un fichier

Pour faciliter la recherche de données d'un fichier à un autre, la nomenclature suivante est utilisée pour nommer les fichiers :

[préfixe] [année] [mois] [jour] [unité] [suffixe] . [extension]

Les préfixes et suffixes peuvent apporter des informations supplémentaires sur la nature des fichiers. En inscrivant l'année en premier suivi du mois et du jour, on facilite le classement en ordre chronologique.

Même si les données à bord des véhicules ne sont pas systématiquement sauvegardées à chaque jour, il est quand même facile de retrouver l'information désirée, les données à l'intérieur des fichiers étant classées en ordre chronologique. Le moment d'inscription (date et heure, à la seconde près) des données est considéré comme une clé primaire, c'est-à-dire une variable qui indique l'ordre original des données.

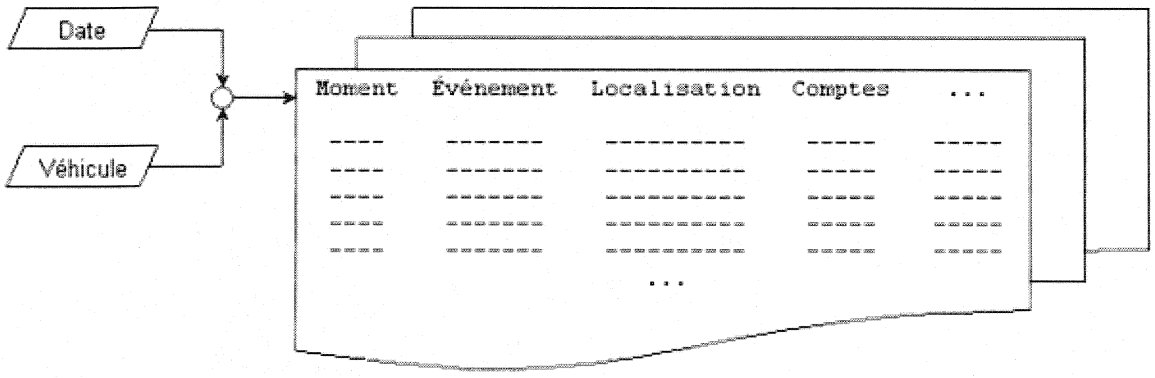


Figure 3-5 : Diagramme schématique d'une base de données pour une société de transport

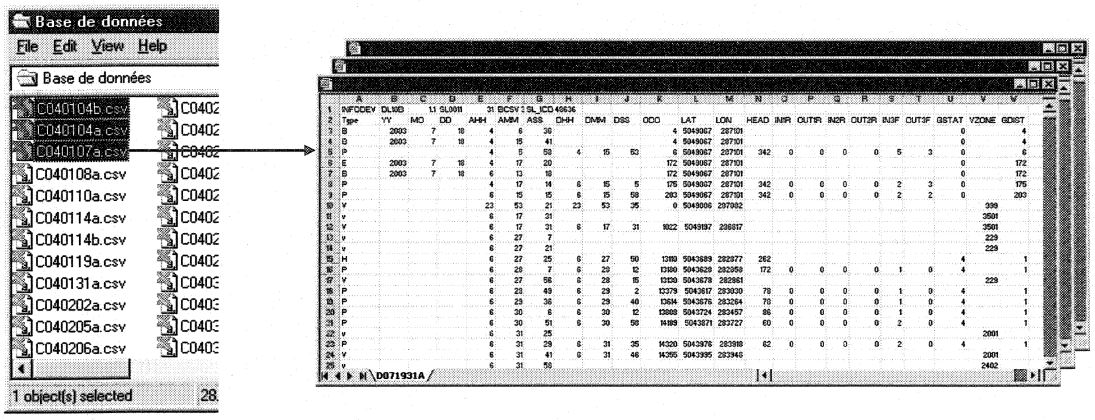


Figure 3-6 : Sélection d'un fichier de données brutes

3.5 Accès aux données

3.5.1 Variables de référence

Une analyse du transport peut se faire en fonction de diverses variables de référence : date, véhicule, chauffeur, endroit ou ligne, auxquelles on associe les valeurs que l'on veut mesurer : ponctualité, achalandage, vitesse, etc.

Trois variables en particulier se distinguent par leur importance :

- Véhicule : Chaque fichier de sauvegarde correspond déjà à un véhicule spécifique (unité). Il suffit alors, pour chaque véhicule, de réunir les données provenant de fichiers créés à partir du même véhicule (agrégation véhiculaire). Pour chaque véhicule, on obtient la description de ses déplacements pendant une période donnée et on peut calculer leurs coûts d'opérations (kilométrage, période d'exploitation).

- Jour : le jour est l'unité de temps idéal pour analyser l'évolution de la demande sur un réseau de transport; en général, les horaires des véhicules (et dans une certaine mesure, ceux des usagers) sont très semblables d'une journée à l'autre, et même si certains jours de la semaine sont différents (samedi ≠ mercredi), ceux-ci peuvent se

comparer sur une base hebdomadaire ($samedi_2 \approx samedi_1$). Chaque fichier correspond à une journée (ou quelques jours) d'opération d'un autobus.

- Ligne : les lignes de transport (circuits) sont le fondement du réseau de transport collectif : les usagers montent et descendent aux arrêts puis voyagent sur les lignes. C'est à partir des lignes de transport qu'on détermine les profils de charge, les diagrammes espace-temps (la position des véhicules effectuant une ligne donnée par rapport au temps) et d'autres indicateurs de performance et d'achalandage.

3.6 Remarques

3.6.1 Collecte des données

640K ought to be enough for anybody.

– Bill Gates en 1981

La nature événementielle de la collecte des données permet l'économie d'espace-mémoire mais le niveau de détail est considérablement affecté, comme on peut le constater au chapitre 5. On ne peut s'empêcher de remettre en question cette « économie » lorsqu'on observe l'augmentation graduelle de la capacité des nouveaux supports informatiques, ou encore lorsqu'on envisage la transmission à distance des données vers un serveur central.

Si le changement de technologie n'est pas envisageable à court terme, d'autres solutions plus modestes peuvent être essayées, comme l'ajout de balises virtuelles ou la programmation de nouveaux événements basés sur le temps, la distance, l'orientation du véhicule, la vitesse, etc. Certaines de ces idées sont déjà utilisées ailleurs; par exemple, voir UTRC2 (2003).

La gestion des données GPS en temps réel semble gagner de plus en plus en popularité. Les avantages sont réels : gestion plus directe du réseau de transport et possibilité d'informer le public sur le service en cours et les délais imprévus. L'information au public est justement la mission que s'est donnée la compagnie NextBus aux États-Unis (voir section 2.4). À Montréal, la gestion des trains de banlieue est effectuée en temps réel par l'Agence métropolitaine de transport (AMT). La localisation de chaque train obtenue par GPS est envoyée directement vers un serveur central via le réseau de téléphone cellulaire Telus. Le suivi des opérations est effectué au siège social de l'AMT et, au besoin, des annonces vocales peuvent être faites aux gares pour informer les usagers.

3.6.2 Traitement des données

Le traitement des données *par étape* tel que décrit précédemment est imparfait car les informations relatives aux circuits planifiés sont séparées des informations sur les horaires d'autobus et les assignations de chauffeurs. Le logiciel BusStops actuellement utilisé à la STL n'utilise que les informations décrivant les circuits planifiés pour effectuer le couplage avec les données GPS, ce qui rend l'identification des circuits plus difficile et sujet à l'erreur. Le logiciel LADRA, ne faisant aucune correction sur les données de BusStops, ne peut que déterminer l'horaire des circuits identifiés tout en écartant les erreurs d'identification. La productivité des systèmes de collecte de données est donc sérieusement diminuée.

Une méthode de traitement des données réunissant circuits et horaires est proposée au chapitre 7 et porte ses fruits. Similairement, la compagnie Infodev a produit une nouvelle version de BusStops permettant une approche *unifiée* du traitement des données. Cette nouvelle version, utilisée notamment par le RTC,

tire profit de trois grandes familles de données : collectées (GPS, comptage), planifiées (circuits, horaires) et opérationnelles (rapports d'opération).

À ce jour, la STL utilise encore l'ancienne version de BusStops en combinaison avec le logiciel LADRA.

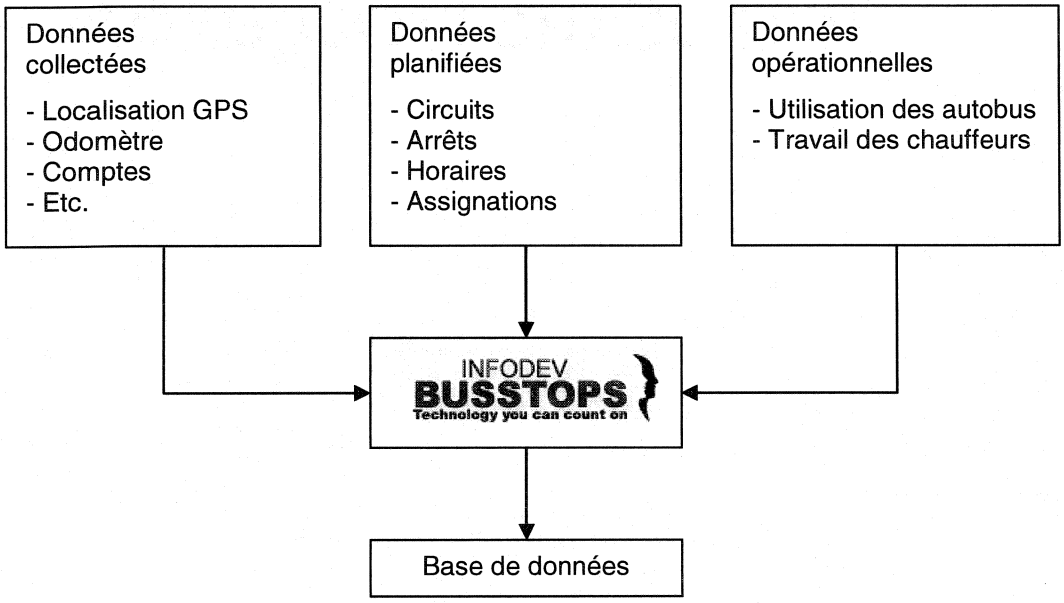


Figure 3-7 : Traitement des données au RTC

CHAPITRE 4 : Analyse de la validité des données

4.1 Identification des circuits

Les données GPS collectées par les autobus de la STL ne contiennent aucune information de planification (circuits, horaires...) et aucune collaboration n'est sollicitée de la part des chauffeurs pour compléter ces données. Une fusion avec les données planifiées – comme celles expliquées aux sections 3.3 et 3.6.2 – est donc nécessaire afin de déterminer les circuits parcourus et les horaires suivis.

Dans un premier temps, la STL utilise BusStops pour déterminer quels circuits ont été effectués par les véhicules. Idéalement, cette procédure ne rejette que les voyages à vide, identifiant l'ensemble des tracés correspondant à des circuits réels. En pratique, malheureusement, un manque de données ou une mauvaise interprétation de celles-ci fait en sorte que certains circuits effectués ne sont pas identifiés ou sont identifiés incorrectement, et que certains voyages à vide sont identifiés comme des circuits numérotés (« faux positifs »).

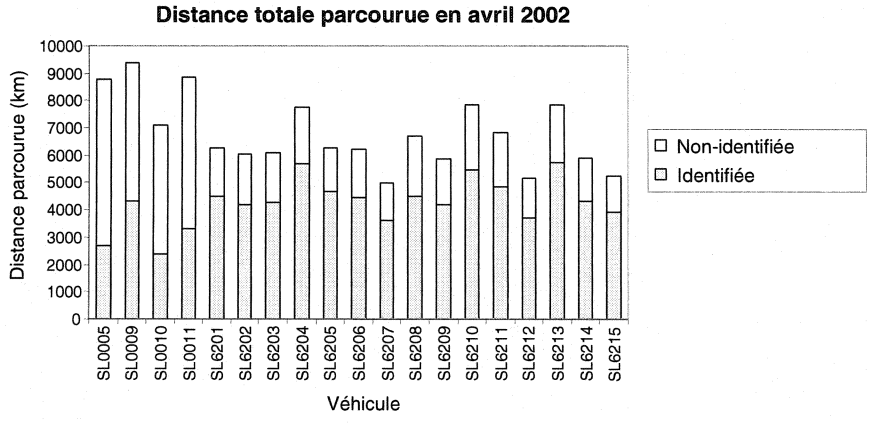


Figure 4-1 : Distance parcourue par les systèmes Infodev en avril 2002

La STL disposait en avril 2002 de 19 autobus équipés de GPS et de système de comptage. Selon les données accumulées pendant le mois d'avril 2002 par ces

autobus, un autobus parcourt en moyenne 275 km par jour (déviation standard = 94 km/j, maximum = 487 km en une journée, les déplacements inférieurs à 1 km par jour sont ignorés dans le calcul). Selon les mêmes données, le nombre de circuits identifiés est d'environ 34 par 1000 km, soient 9 circuits par jour. Cela correspond à 61 % de la distance totale parcourue, incluant les voyages à vide. Mais cette productivité n'est pas égale pour tous les véhicules (Figure 4-1 et Figure 4-2).

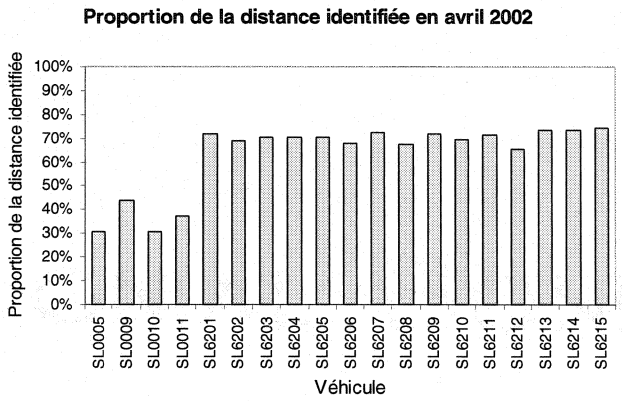


Figure 4-2 : Proportion de la distance identifiée par BusStops

Deux éléments importants doivent être pris en compte. La première concerne les distances non-identifiées : naturellement, la distance parcourue en circuits ne sera jamais 100 % de la distance totale parcourue puisque une partie de la distance totale (environ 20 %) est toujours effectuée à vide (haut-le-pied, ou *dead head*). Par contre, les distances non identifiées incluent aussi des circuits qui n'ont pas été reconnus comme tels (« faux négatifs »).

La seconde concerne les distances identifiées : une partie de ces distance inclut des parcours faussement identifiés comme étant des circuits (« faux positifs », Figure 4-3); ces faux circuits ont une importance particulière car ils altèrent les résultats des analyses. En principe néanmoins, ces erreurs sont éliminées lors de la fusion avec les horaires planifiés effectuée par LADRA.

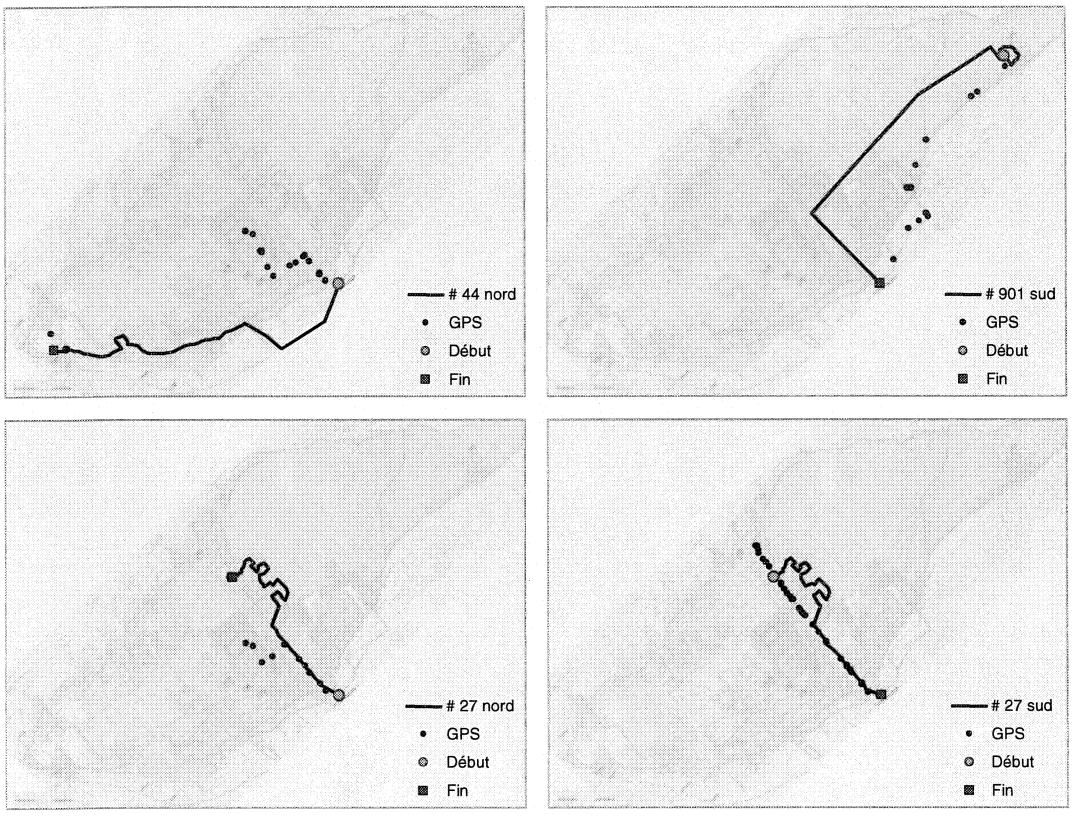


Figure 4-3 : Quelques parcours faussement identifiés comme circuits

4.2 Identification des tournées

Dans un deuxième temps, la STL utilise LADRA pour déterminer les heures de départ des circuits identifiés. Cette procédure permet en même temps de détecter et rejeter les circuits faux ou mal identifiés. Des observations (décrites à la section 7.6.1) montrent que LADRA réussit à éliminer les erreurs de BusStops mais qu'à son tour il ne réussit pas à valider tous les autres circuits identifiés.

Dans ce cas-ci, LADRA réussit à coupler 60 % des circuits détectés à un horaire planifié (Figure 4-4).

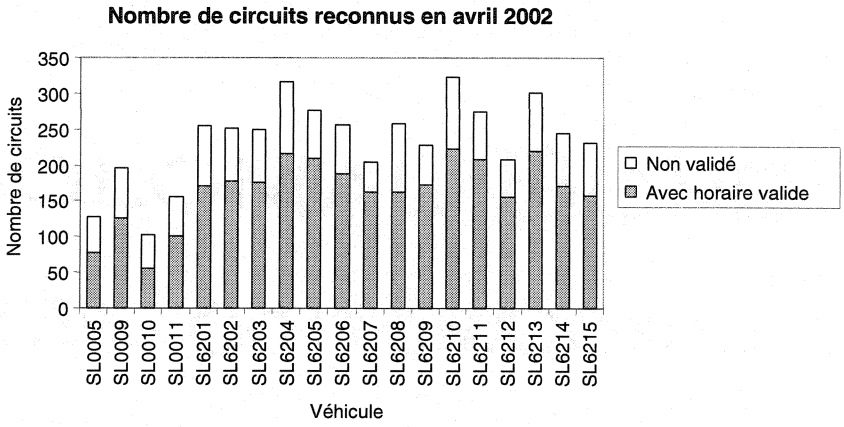


Figure 4-4 : Circuits reconnus avec un horaire validé par LADRA

4.3 Couverture du réseau

Un réseau est constitué d'un ensemble de lignes (circuits) sur lesquelles le service est établi en terme de tournées. Une tournée est un passage de véhicule planifié à un moment précis dans une journée, que l'on identifie par l'heure de départ.

À chaque jour, les 19 autobus munis de systèmes de collecte de données sillonnent le réseau routier et cumulent de l'information en effectuant diverses tournées sur différents circuits. Lorsque l'on dispose de données accumulées pendant plusieurs jours, on peut effectuer une agrégation temporelle (BLOUIN, 1984). Cette technique consiste à regrouper des données accumulées sur une certaine période en fonction des jours. À toutes fins pratiques, on distingue quatre sortes de jours : ouvrables, samedis, dimanches et fériés.

Pour un circuit donné, on souhaite que toutes les tournées soient représentées au moins une fois pendant une période donnée, ce qui permet une analyse complète de l'évolution de l'achalandage pendant une journée.

Une analyse de la représentativité des données de la STL accumulées pendant le mois d'avril 2002 donne les résultats suivants :

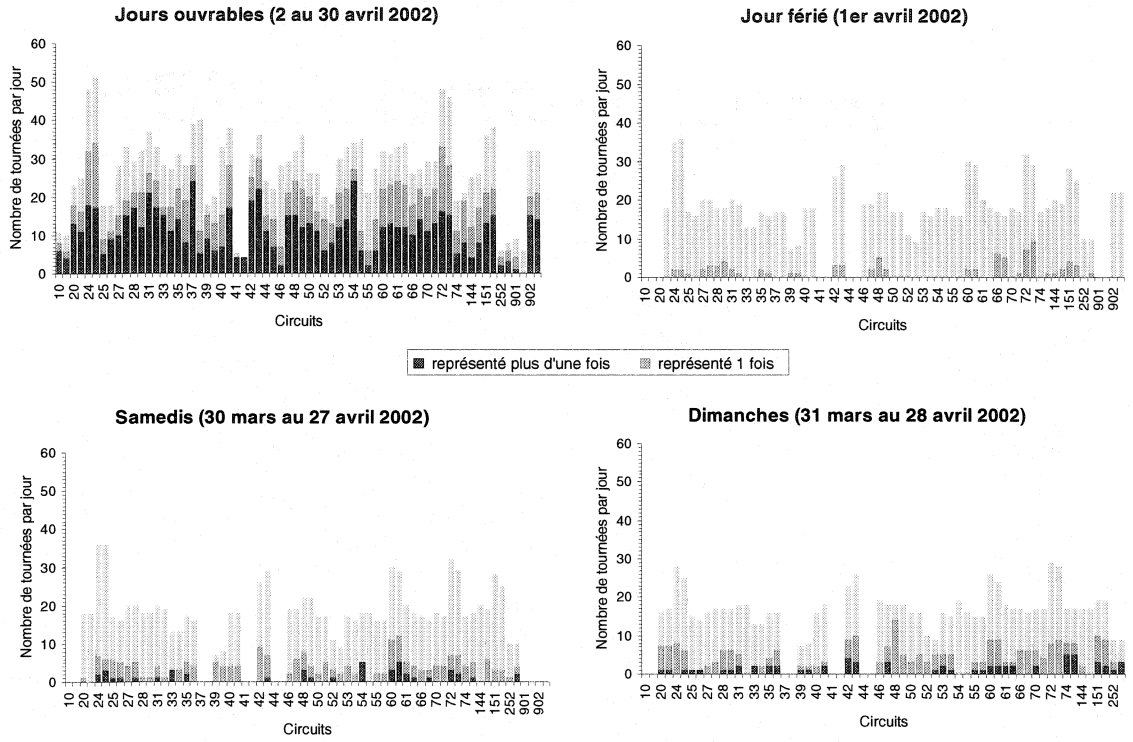


Figure 4-5 : Représentativité des données collectées pendant avril 2002

Dans la Figure 4-5, chaque colonne grise correspond à un nombre *planifié* de tournées pour chaque circuit et chaque direction (une absence de colonne signifie qu'il n'y a pas de service ce jour-là). Au total, on retrouve à chaque jour ouvrable 1769 tournées planifiées réparties sur 33 circuits allers-retours (66 colonnes).

Les circuits identifiés couvrent environ 65 % des tournées planifiées pendant les jours ouvrables (en turquoise et en bleu). Cette proportion est évidemment supérieure à celle des autres jours puisque il y a 21 jours ouvrables pendant le mois d'avril alors qu'il n'y a que 5 samedis (22 %), 5 dimanches (28 %) et 1 jour

férié (8 %). De plus, la STL ne semble pas favoriser particulièrement l'utilisation des véhicules munis de systèmes GPS. À preuve, comme on le constate à la Figure 4-6, ces véhicules sont moins utilisés les fins de semaine.

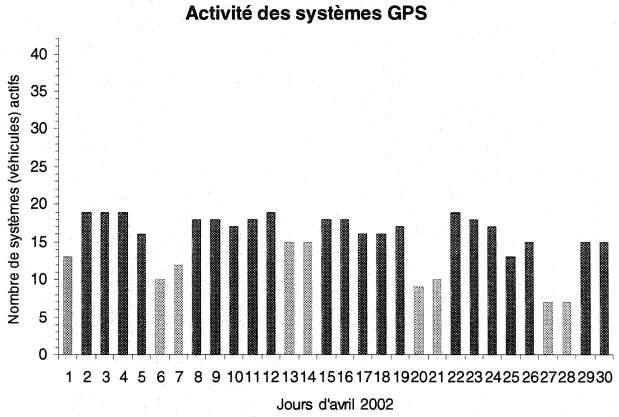


Figure 4-6 : Utilisation des systèmes Infodev pendant avril 2002

On doit toutefois souligner le caractère particulier des autobus utilisés ici. La flotte de la STL est constituée de deux modèles d'autobus : le « vieux » Classic et le plus récent Nova LFS (*Low Floor Series*, à plancher surbaissé). Les 19 autobus illustrés ici sont tous du modèle conventionnel Classic, lequel a une plus grande capacité de passagers et est donc utilisé surtout sur les circuits à achalandage élevé pendant les périodes de pointe. Cela pourrait expliquer la couverture inégale du réseau pendant les jours ouvrables et l'utilisation plus faible des systèmes Infodev la fin de semaine.

Cela dit, ces systèmes ont effectués dans l'ensemble plus de 2400 tournées pendant les jours ouvrables d'avril 2002, soit 33 % de plus que le nombre total de tournées planifiées. Il est donc théoriquement possible d'effectuer une couverture complète du réseau en un mois (jours ouvrables) avec seulement 19 autobus. La couverture du réseau pourrait être grandement améliorée avec un

algorithme de reconnaissance des circuits plus efficace et une priorité accordée aux véhicules munis de systèmes de collecte de données.

> *Si l'échantillonnage pouvait être contrôlé parfaitement, l'ensemble des voyages de semaine serait couvert au moins une fois par mois. Évidemment, ce calcul théorique se heurte aux contraintes opérationnelles (variation quotidienne des affectations de véhicules aux tournées, véhicules en réparation, pannes, etc.) et à la lourdeur de la gestion qui serait nécessaire afin d'en assurer le suivi.* (LAVIGUEUR, 2001).

4.3.1 Profil de charge 3D

La représentation graphique des résultats occupe une place importante de par sa capacité à transmettre une grande quantité d'informations à la fois et aussi parce que l'on y aperçoit parfois des détails (formes suspectes, tendances) qui, sous forme numérique, auraient échappés à un œil attentif.

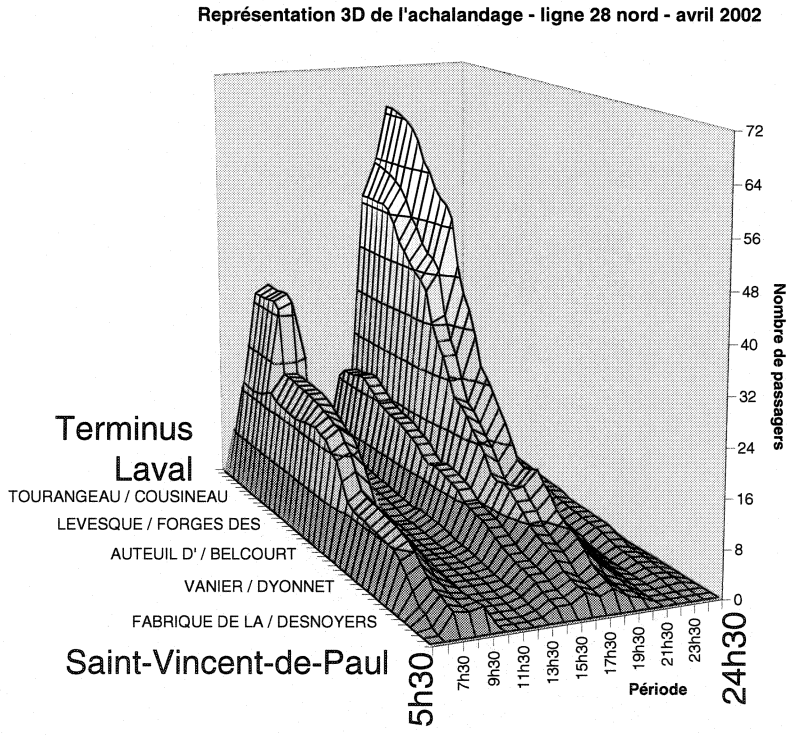


Figure 4-7 : Profil de charge 3D - ligne 28 nord, avril 2002

Le profil de charge 3D présenté à la Figure 4-7 illustre le circuit 28 partant du terminus Laval (métro Henri-Bourassa) et se rendant au pénitencier Saint-Vincent-de-Paul pendant le mois d'avril 2002. L'achalandage des tournées échantillonnées par des véhicules munis d'un système GPS est exprimée en fonction du temps, ce qui permet de mettre clairement en évidence les périodes de pointe et hors-pointe.

La représentation du portrait réel de l'achalandage exigerait la pondération des tournées échantillonnées afin de tenir compte de l'ensemble des tournées effectivement opérées. En effet, toutes les tournées n'ont pas été visitées par les véhicules munis de système GPS pendant le mois d'avril 2002. Le Tableau 4-1 donne un aperçu du niveau d'échantillonnage obtenu.

Tableau 4-1 : Échantillonage obtenu pour la ligne 28 nord en avril 2002

Période	N. de tournées retrouvées	N. total de tournées
1h00 à 5h59	0	0
6h00 à 6h59	1	2
7h00 à 7h59	2	2
8h00 à 8h59	2	2
9h00 à 9h59	1	1
10h00 à 10h59	0	1
11h00 à 11h59	1	2
12h00 à 12h59	0	1
13h00 à 13h59	0	2
14h00 à 14h59	1	1
15h00 à 15h59	2	2
16h00 à 16h59	0	2
17h00 à 17h59	3	3
18h00 à 18h59	2	2
19h00 à 19h59	1	1
20h00 à 20h59	1	1
21h00 à 21h59	1	1
22h00 à 22h59	1	1
23h00 à 23h59	1	1
0h00 à 0h59	1	1

4.4 Comptage des passagers

La compagnie Infodev a innové en concevant un détecteur très robuste qui, en théorie, n'est pas influencé par la vitesse de passage des gens sous le capteur ni par les passagers immobiles dans sa zone de comptage, qui ne compte pas les objets ou les sacs, qui sépare bien les passagers se suivant de près et qui demeure insensible aux variations de température et d'éclairage (INFODEV, www.infodev.ca).

Néanmoins, des erreurs se glissent quand même dans le comptage des passagers le long des circuits. Une accumulation d'erreurs se traduit par une charge résiduelle en fin de parcours. Deux exemples de profil de charge erroné sont présentés à la Figure 4-8.

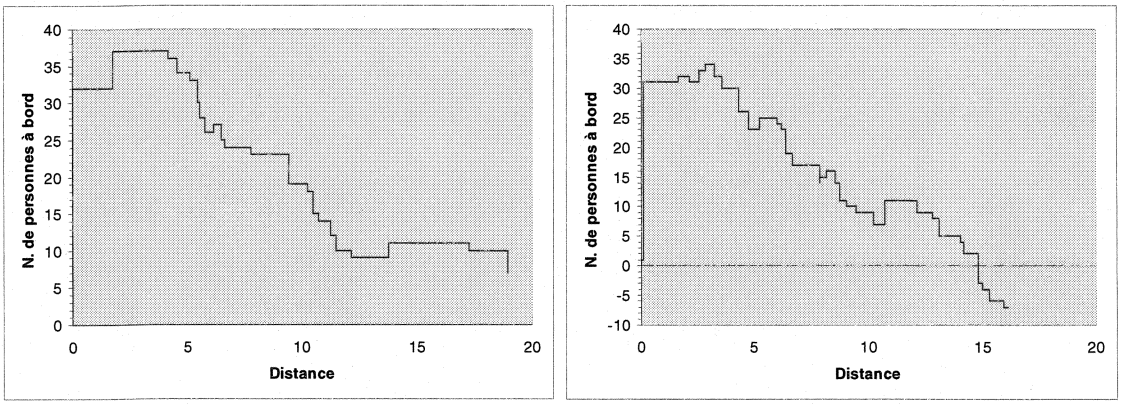


Figure 4-8 : Exemples de profils de charge erronés

Pour fins de simplicité et pour limiter les erreurs d'analyse, un filtre est appliqué : un profil de charge est considéré valide si la charge finale (charge résiduelle en fin de parcours) se situe entre 5 et -5. Les autres tournées sont rejetées.

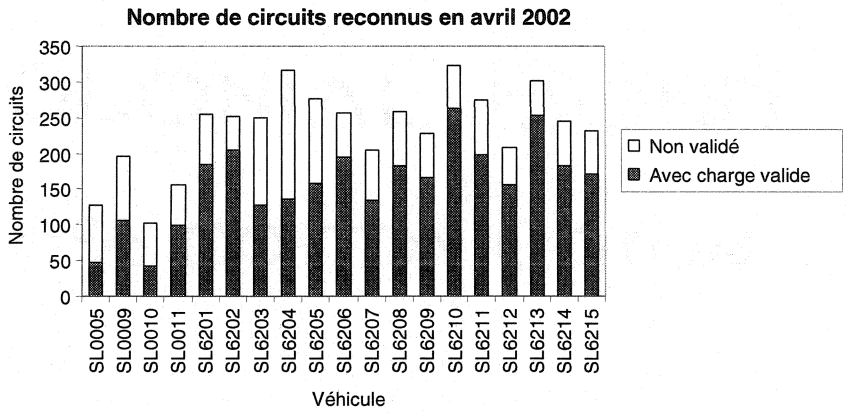


Figure 4-9 : Circuits reconnus avec un profil de charge valide

Il s'avère alors que 67 % des circuits détectés possèdent un profil de charge valide. Les circuits possédant à la fois un horaire et un profil de charge valides représentent 42 % des circuits identifiés (Figure 4-10).

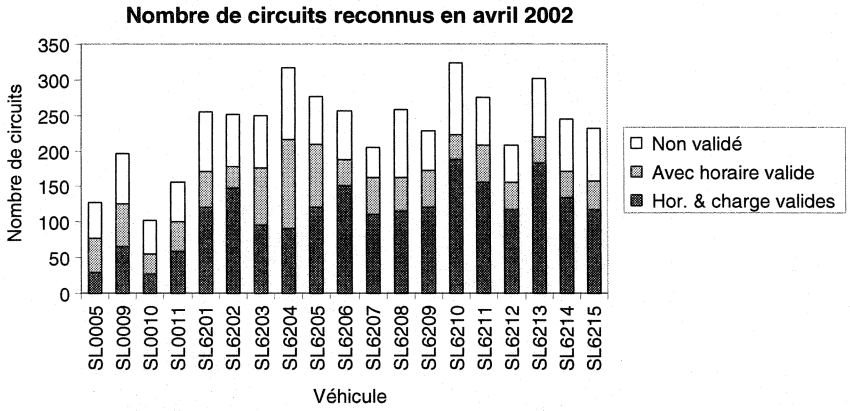


Figure 4-10 : Circuits reconnus avec un horaire et un profil de charge valides

4.5 Données d'avril 2003

En avril 2003, quarante-deux autobus sont équipés de systèmes Infodev.

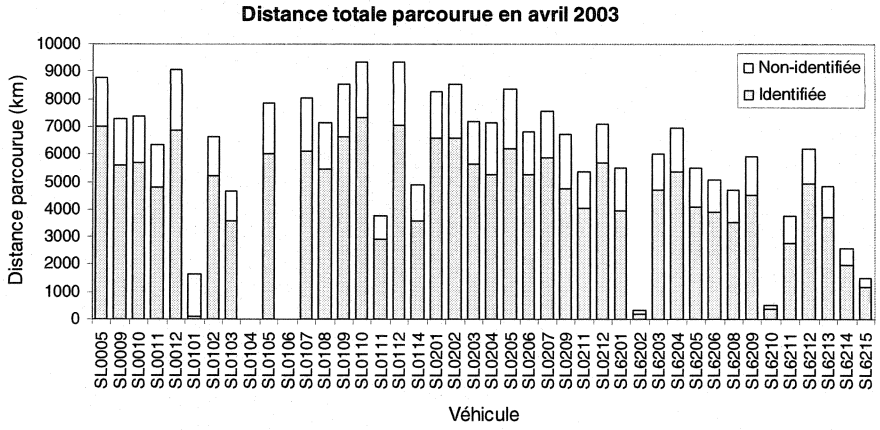


Figure 4-11 : Distance parcourue par les systèmes Infodev en avril 2003

Selon les calculs, et en mettant de côté les systèmes problématiques, un autobus effectue en moyenne 270 km par jour (écart-type = 20 km/j, maximum = 300 km en une journée, les déplacements inférieurs à 1 km par jour sont ignorés dans le calcul). Environ 11 circuits par jour sont identifiés, ce qui correspond à 77% de la distance totale parcourue, incluant les voyages à vide.

Plusieurs systèmes ont connu des problèmes techniques affectant la prise de données, ce qui explique les quelques colonnes vides. Les problèmes rencontrés sont de trois sortes :

- Horloge déréglée; les données collectées sont associées à une mauvaise date et peut-être aussi une mauvaise heure (véhicules SL6210 et SL6215).
- GPS défectueux; aucune donnée GPS n'est collectée (véhicules SL0101 et SL6202).
- Absence de données provenant de l'odomètre; apparemment, ce problème en cause un autre : aucune ouverture de porte ou halte n'est enregistrée alors que les données GPS indiquent des déplacements (véhicules SL0106 et possiblement SL6202).

En plus de ces problèmes, d'autres non apparents à première vue sont mis en évidence au chapitre 8.

Comme il a été expliqué précédemment, ce ne sont pas tous les circuits qui sont validés par LADRA et certains présenteront un profil de charge erroné. Dans ce cas-ci, 74 % des circuits détectés ont un horaire valide et 49 % ont à la fois un horaire et un profil de charge valides, ce qui est une légère amélioration par rapport à avril 2002.

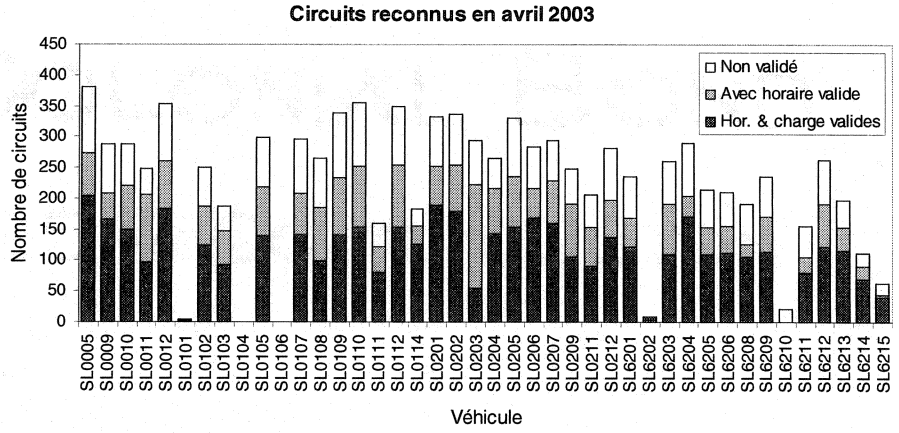


Figure 4-12 : Circuits reconnus avec un horaire et un profil de charge valides (avril 2003)

Les autobus munis de systèmes GPS ont effectué plus de 6100 tournées pendant les jours ouvrables, soit plus de trois fois le nombre total de tournées planifiées. On s'attend donc à ce que la couverture du réseau soit meilleure, et c'est le cas. Les 21 jours ouvrables sont couverts à 86%, les 4 samedis à 56% et les 4 dimanches à 55%. Étrangement, bien qu'un nombre appréciable de tournées ait été identifié pendant les 2 jours fériés, très peu d'entre elles sont associées à un horaire valide.

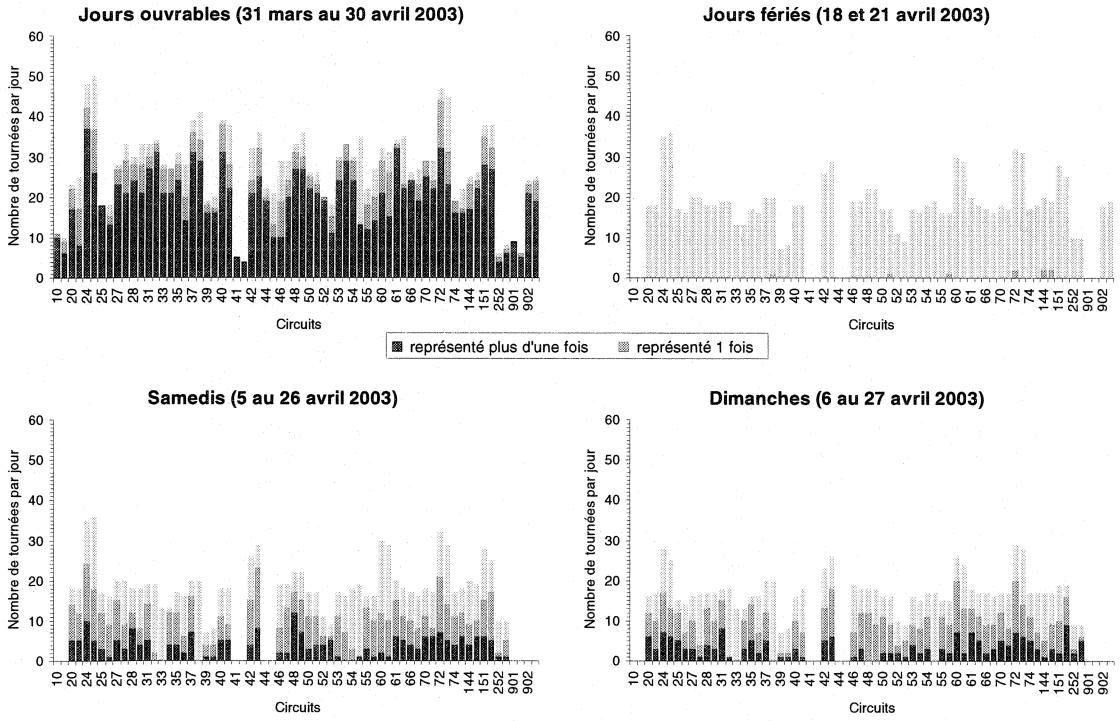


Figure 4-13 : Représentativité des données collectées pendant avril 2003

Comme dans le cas précédent, on observe une diminution de l'utilisation des systèmes Infodev pendant les fins de semaine. Rappelons toutefois que 19 des autobus représentés ici sont du modèle Classic, ce qui semble expliquer cette variation (laquelle est semblable à celle de la Figure 4-6). Les 23 autres sont du modèle Nova LFS, et cela pourrait bien expliquer l'augmentation notable de la couverture pendant les fins de semaine.

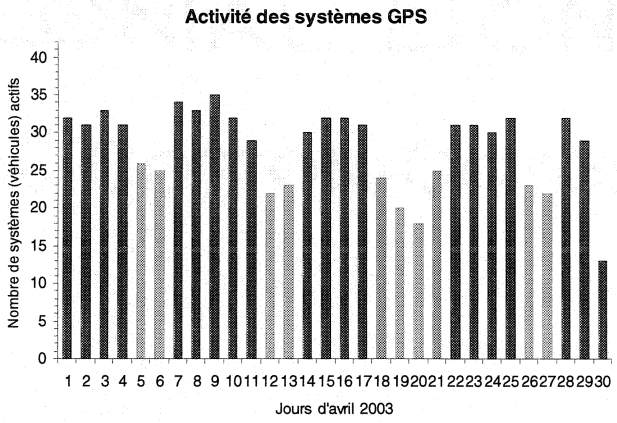


Figure 4-14 : Utilisation des systèmes Infodev pendant avril 2003

4.6 Balises virtuelles

Le réseau de la STL est constitué de 2309 arrêts d'autobus (selon les données d'avril 2004) répartis sur l'île Jésus et aux environs. À noter que certains arrêts partagent les mêmes coordonnées géographiques. De fait, les arrêts d'autobus qui se font face (i.e. situés de chaque côté de la rue) sont souvent décrits par les mêmes coordonnées. Par conséquent, on ne retrouve en réalité que 1537 localisations différentes.

Des repères géographiques sont stockés dans la mémoire des ordinateurs embarqués afin que, lorsqu'un autobus muni d'un système GPS passe à proximité de l'un de ces repères, un enregistrement soit déclenché dans l'ordinateur. Cela fait en sorte que des données seront collectées à des endroits importants même si aucun autre événement ne survient. Ces localisations stockées en mémoire sont appelées *balises virtuelles*.

On retrouve au total 236 balises virtuelles réparties sur le réseau (Figure 4-15).

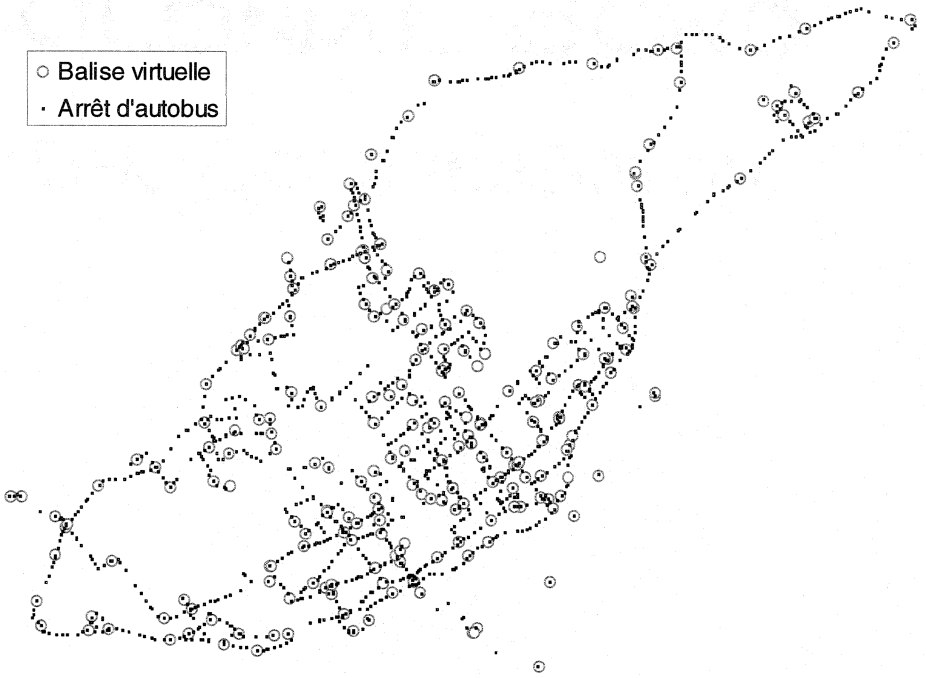


Figure 4-15 : Localisation des arrêts d'autobus et des balises virtuelles de la STL

4.6.1 Détection des balises virtuelles

En moyenne, les balises virtuelles sont détectées à 96 %. La détection de balises sur un parcours s'avère donc un outil très intéressant pour identifier les circuits qui sont parcourus.

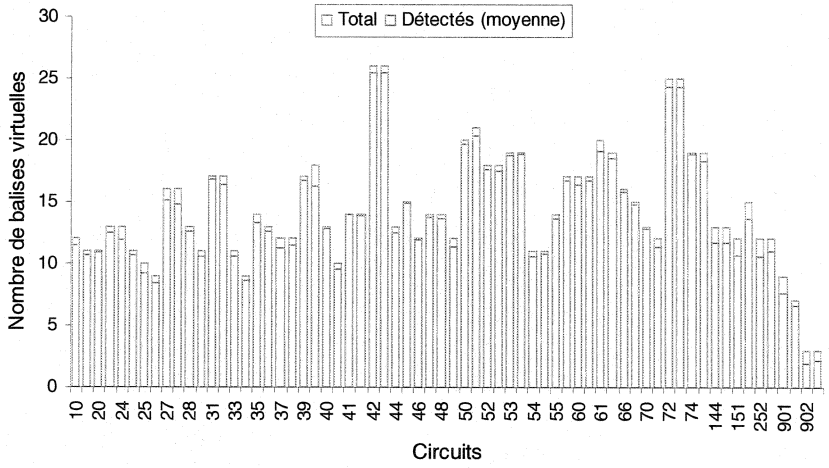


Figure 4-16 : Nombre moyen de balises détectées par circuit

La détection de balise semble s'effectuer avec la même efficacité pour tous les ordinateurs embarqués, sauf peut-être pour celui de l'autobus SL0101. Cependant, cette différence est probablement reliée au fait que très peu de données proviennent de ce véhicule (Figure 4-12).

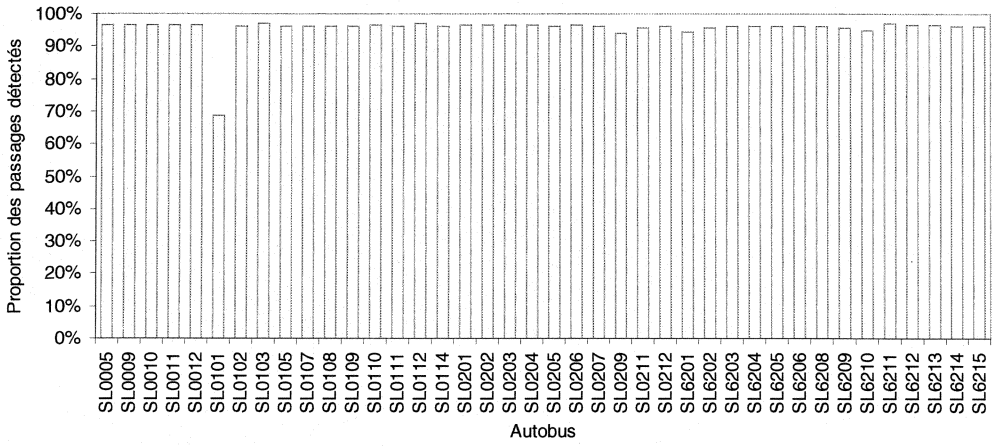


Figure 4-17 : Taux de détection de balise pour chaque autobus

La grande majorité des balises sont détectées plus de 95 % du temps. Néanmoins, un nombre significatif de balises ont un taux de détection de moins de 80 %. Le positionnement de certaines balises est sans doute la principale raison de ce problème.

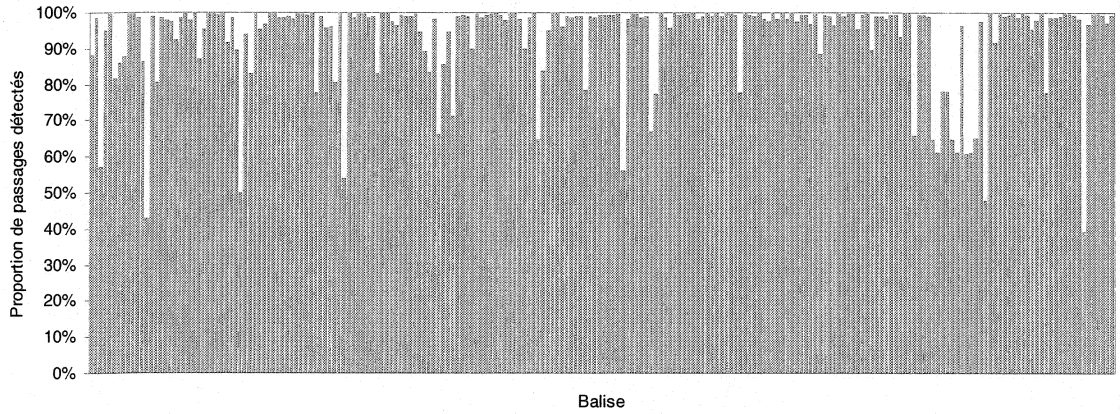


Figure 4-18 : Taux de détection de chaque balise virtuelle

L'approche des véhicules peut aussi être un facteur important, étant donné que certaines balises sont probablement moins bien détectées dans certaines directions que dans d'autres. Cette problématique a été abordée par JACQUOT (2001), lequel illustre dans son rapport la variation de la position des véhicules à chaque fois que l'un d'eux détectait une balise (occurrence).

Sur un tronçon isolé, chaque occurrence sera située près de l'endroit théorique et l'ensemble des occurrences formera un amas à peu près elliptique. Mais près d'une intersection, la forme peut varier grandement et les occurrences peuvent former plusieurs amas, comme dans les exemples ci-dessous (les lignes indiquent la position approximative des routes mais non leur largeur).

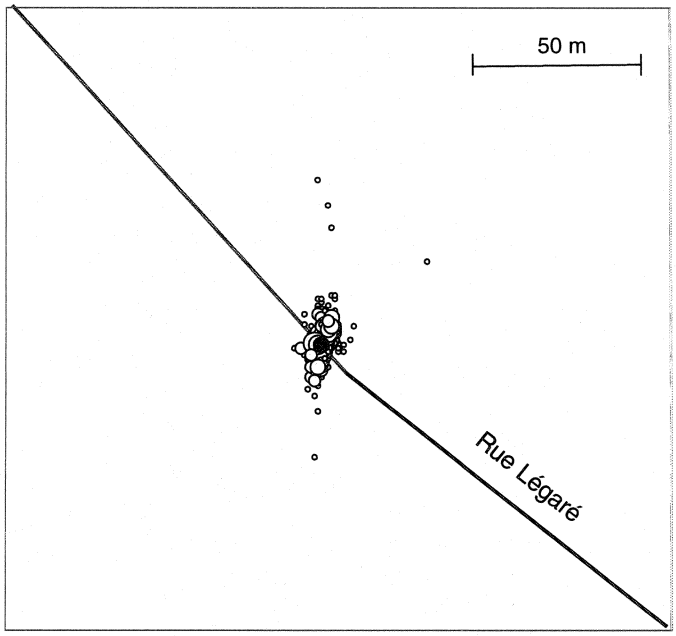


Figure 4-19 : Occurrences de la balise 5500 (circuit 55, près du parc Légaré)

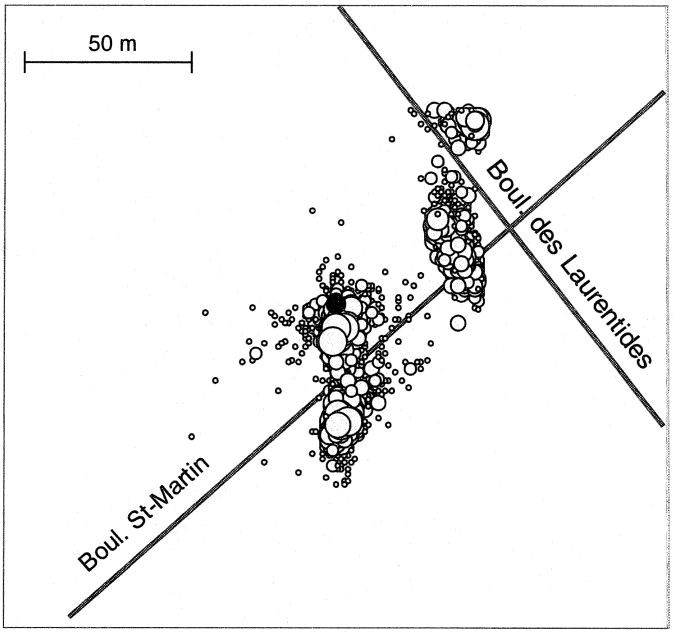


Figure 4-20 : Occurrences de la balise 236

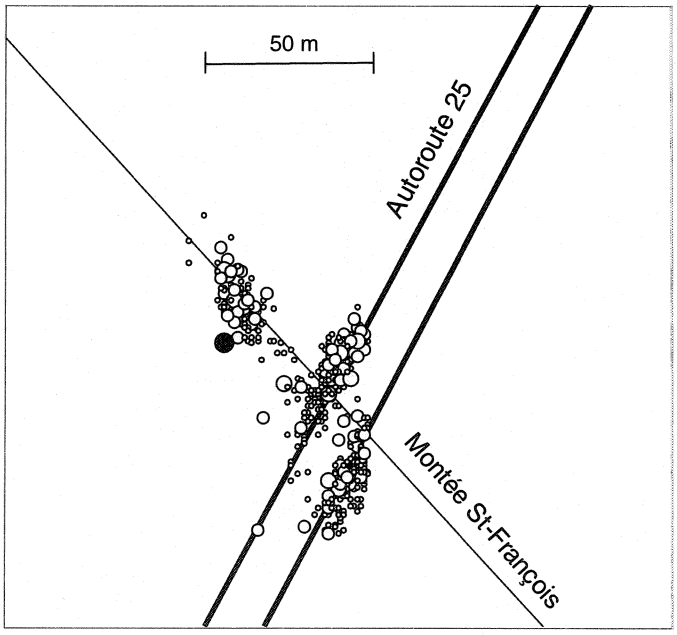


Figure 4-21 : Occurrences de la balise 9105

CHAPITRE 5 : Expériences sur le terrain

Quelques expériences ont été effectuées à bord d'autobus de la STL. Ces expériences avaient comme but principal de se familiariser avec la réalité que les données GPS tentent de décrire de façon plus ou moins précise. De plus, ils ont permis d'obtenir une trace GPS continue grâce à un appareil GPS indépendant et, à deux occasions, un comptage manuel fut effectué afin de vérifier la précision des détecteurs de mouvement.

5.1 Première expérience

Une première expérience sur le terrain a été conduite à l'aide d'un appareil GPS portatif. Ces appareils à usage personnel permettent d'accumuler une quantité appréciable de coordonnées et d'obtenir un tracé très précis.

La trace d'un autobus de la STL équipé d'un système GPS a été obtenue à l'aide d'un tel dispositif de poche. Cette expérience permet de comparer les informations géographiques collectées par deux types de systèmes.

Pendant le déroulement de cette première expérience, l'autobus SL0011 a effectué les voyages suivants (Figure 5-1) :

- 7h02 : départ du terminus Laval vers le point d'injection du circuit 54 sud (centre Duvernay);
- 7h17 à 7h48 : circuit 54 sud (vers le terminus Laval);
- 8h00 : départ du terminus Laval vers le point d'injection du circuit 72 sud (gare Sainte-Dorothée);
- 8h42 à 9h45 : circuit 72 sud (vers le terminus Laval);
- 9h52 à 10h40 : circuit 60 nord (vers Chomedey);
- 10h44 à 11h25 : circuit 60 sud (vers le terminus Laval).

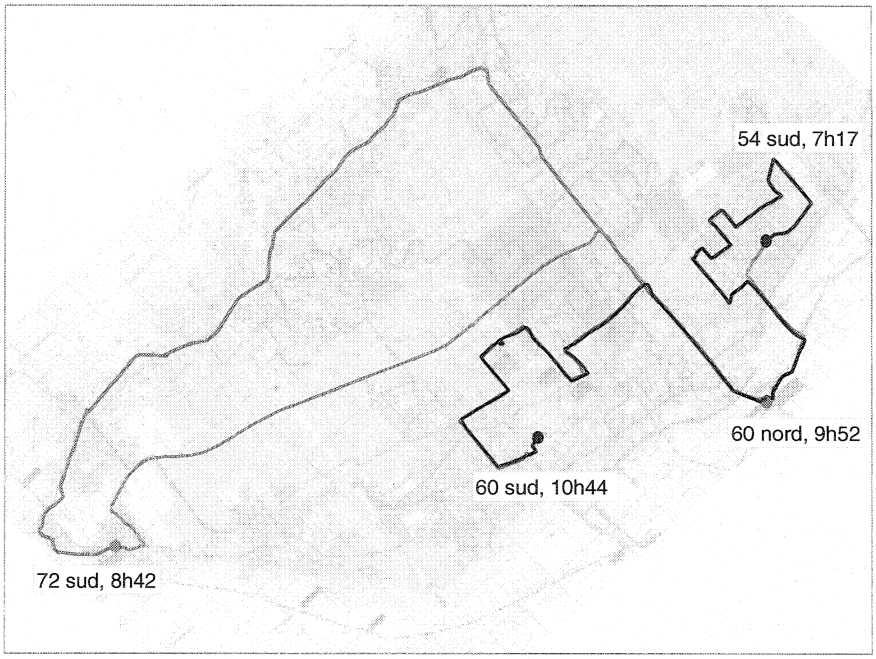


Figure 5-1 : Parcours effectué pendant l'expérience 1 réalisée le 18 juillet 2003

Tel qu'expliqué à la section 3.2, l'ordinateur embarqué effectue la collecte des données de façon événementielle. Ainsi, les événements « Ouverture de portes », « Halte de plus de 15 s » et « Détection de balise virtuelle » ont provoqué une prise de données aux endroits suivants :

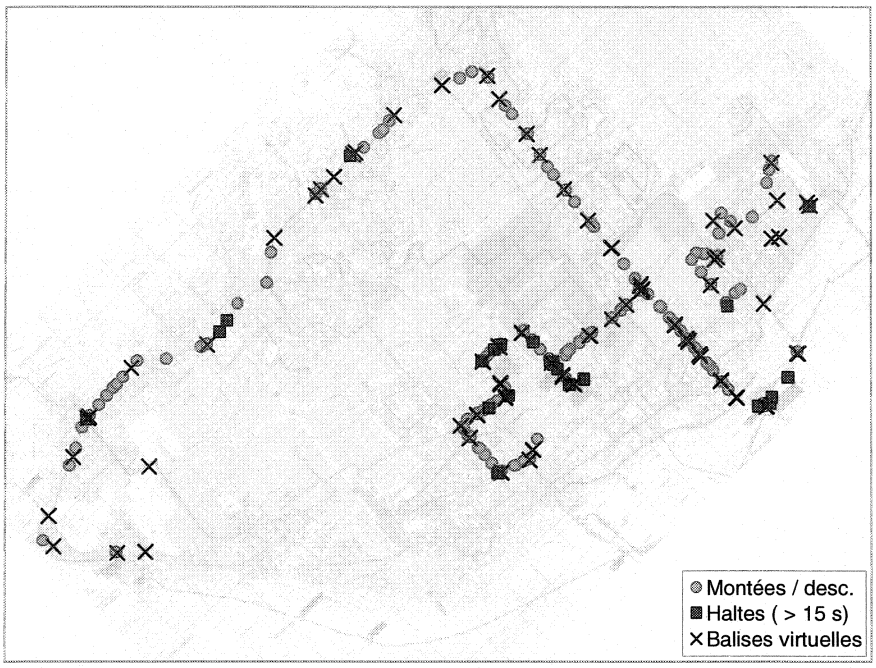


Figure 5-2 : Expérience 1 – données collectées par l'ordinateur embarqué

À première vue, la Figure 5-2 permet de déterminer avec une certitude presque absolue la séquence des circuits qui ont été effectués par l'autobus SL0011 le matin du 11 juillet 2003. Elle permet même une analyse minutieuse du comportement du véhicule dans les zones où la circulation et l'achalandage sont les plus élevés.

Néanmoins, la nature événementielle de la collecte des données révèle sa grande faiblesse : aux endroits où les arrêts du véhicule sont moins fréquents, le *niveau de résolution* de la trace GPS diminue dramatiquement et ne permet qu'une analyse plutôt vague du mouvement du véhicule. L'exemple le plus révélateur est cette absence d'information sur le voyage effectuée entre le terminus Laval et la gare Sainte-Dorothée. Seules quelques rares balises témoignent du passage de l'autobus sur l'autoroute 440.

La Figure 5-3 montre le résultat obtenu avec l'appareil GPS portatif utilisé dans le cadre de l'expérience.

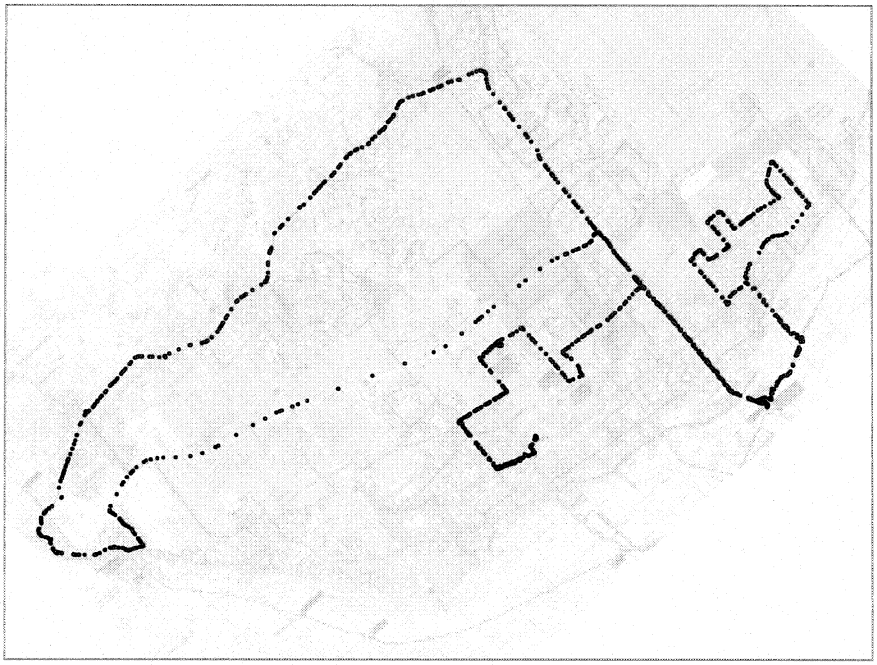


Figure 5-3 : Expérience 1 – données collectées par un appareil GPS portatif

La trace GPS obtenue ici décrit beaucoup plus précisément le parcours effectué. Combiné à une base de données GIS, une telle résolution pourrait s'avérer un outil puissant. Il permet aussi, et surtout, une analyse désagrégée des mouvements de véhicules sur tout le réseau.

> *Les temps de parcours et la ponctualité observés sur le terrain doivent être analysés non pas de manière générale et avec des moyennes de temps de parcours par tronçon, mais bien plutôt avec une approche désagrégée qui permettrait de saisir toutes les facettes des comportements des chauffeurs et des conditions auxquelles ils sont soumis.* (LAVIGUEUR, 2001)

5.1.1 Distribution des données

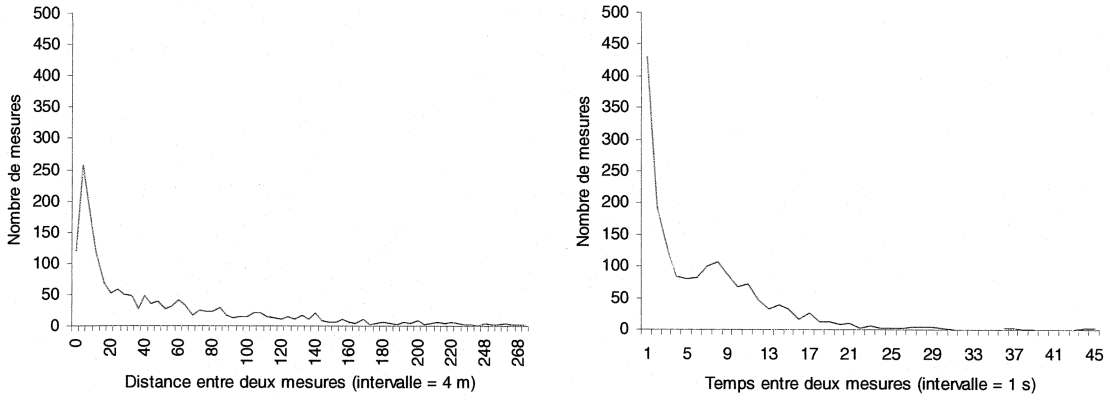


Figure 5-4 : Distribution spatiale et temporelle des données prises par le GPS portatif

Pendant l'expérience 1, l'appareil GPS portatif a accumulé 1770 points en mémoire pour former la trace de la Figure 5-3. Plus de 80 % de ces points sont distancés de moins de 150 mètres entre eux et ont été collectés à moins de 15 secondes d'intervalle (plus de 90 % sont distancés de moins de 150 mètres ou ont été collectés à moins de 15 secondes d'intervalle).

Tableau 5-1 : Intervalles entre chaque collecte de données par un GPS portatif

	Spatial	Temporel
Classe contenant le plus grand nombre d'incidence	4 à 8 m (22 %)	0 à 1 s (24 %)
Intervalle maximum pour 50 % des données	0 à 28 m	0 à 5 s
Intervalle maximum pour 90 % des données	0 à 144 m	0 à 15 s

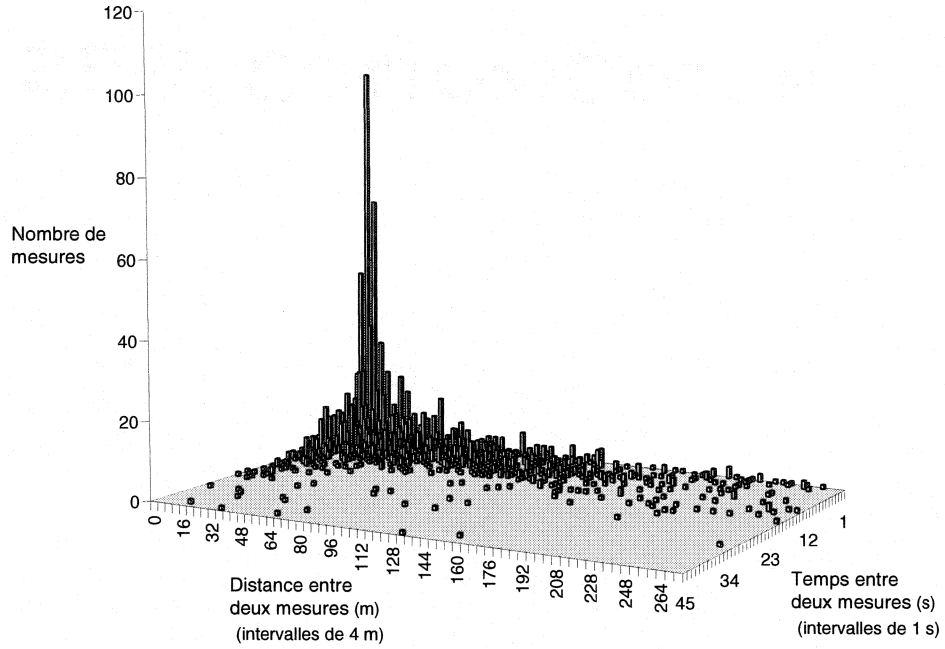


Figure 5-5 : Distribution spatio-temporelle des données prises par le GPS portatif

5.1.2 Profils de charge

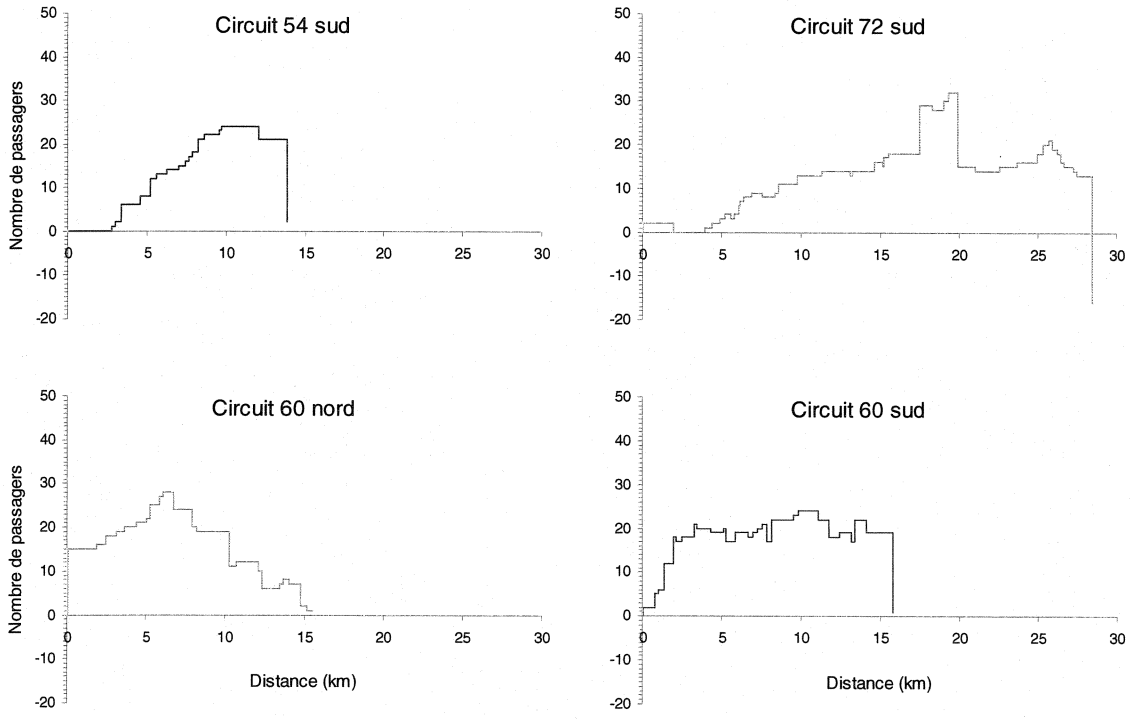


Figure 5-6 : Expérience 1 - profils de charge (détectés par l'ordinateur embarqué)

La charge « négative » qui termine le profil de charge du circuit 72 sud est le résultat d'une accumulation d'erreurs causées par les détecteurs de mouvement. Le comptage des passagers est parfois à ce point erroné et remet en question l'efficacité théorique des capteurs directionnels.

Les deux prochaines expériences furent donc consacrées au comptage manuel des passagers et à l'évaluation de l'efficacité des systèmes de comptage.

5.2 Calcul des erreurs de comptage

Dans le texte qui suit, les valeurs *M* et *D* sont respectivement le nombre total de montées et de descentes pour tout le parcours, soient :

$$M = \sum_{i=1}^{n}\sum_{j=1}^{p} m_{ij} \qquad D = \sum_{i=1}^{n}\sum_{j=1}^{p} d_{ij}$$

avec m = nombre de montées, d = nombre de descentes, n = nombre d'arrêts et p = nombre de portes. Les valeurs primées (M', D', m', d') représentent des montées et des descentes détectées par l'ordinateur.

Avec des comptes manuels, il est possible de calculer l'erreur causée par le système de comptage. On considère quatre types d'erreurs (IRIS, 2004) :

Erreur de comptage équilibrée (*passenger error*) :

$$\varepsilon_e = \frac{(M'+D')-(M+D)}{M+D}$$

Erreur de comptage équilibrée en montée et en descente (*balanced entering/exiting passenger error*) :

$$\varepsilon_e^m = \frac{M'-M}{M} \qquad \varepsilon_e^d = \frac{D'-D}{D}$$

Erreur de comptage non-équilibrée (*unbalanced error*) :

$$\varepsilon = \frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\left|m'_{ij} - m_{ij}\right| + \sum_{i=1}^{n}\sum_{j=1}^{p}\left|d'_{ij} - d_{ij}\right|}{M + D}$$

Le calcul de l'erreur équilibrée sert à déterminer la précision dans le calcul de la charge et du nombre total d'usagers. L'erreur non-équilibrée sert à mesurer la précision dans le comptage absolu des montants et des descendants et détermine l'efficacité réelle des capteurs.

Lorsqu'on ne dispose pas de compte manuel, il demeure possible de mesurer l'*écart moyen* entre les montées et les descentes détectées (IRIS, 2004) :

$$EM_{\text{Iris}} = \frac{\left|M' - D'\right|}{M' + D'}$$

Cette équation est toutefois incomplète puisqu'elle ne tient pas compte des charges résiduelles à la fin de chaque circuit. C'est pourquoi elle est remplacée par celle-ci :

$$EM = \frac{\sum_{k=1}^{v}\left|M'_k - D'_k\right|}{M' + D'}$$

avec v = nombre de voyages ou de circuits. Avec les données de l'expérience précédente, on obtient EM = 5,5 %.

5.3 Deuxième expérience

Cette seconde expérience sur le terrain est semblable à la première : un appareil GPS portatif fut utilisé pour obtenir un tracé continu du voyage d'un autobus sur

une assignation donnée. Cette fois-ci, en plus, un comptage manuel des passagers fut effectué afin de vérifier l'exactitude des capteurs de mouvement directionnels.



Figure 5-7 : Parcours effectué pendant l'expérience 2 réalisée le 14 avril 2004

L'autobus a effectué l'assignation suivante :

- 13h36 : départ du garage vers le point d'injection du circuit 48 sud (pénitencier de Saint-Vincent-de-Paul);
- 13h50 à 14h15 : circuit 48 sud (vers le terminus Laval);
- 14h25 à 15h29 : circuit 55 nord (vers Laval-Ouest);
- 15h59 à 16h59 : circuit 55 sud (vers le terminus Laval);
- 17h15 à 17h43 : circuit 24 nord (vers Chomedey);
- 17h43 : retour au garage.

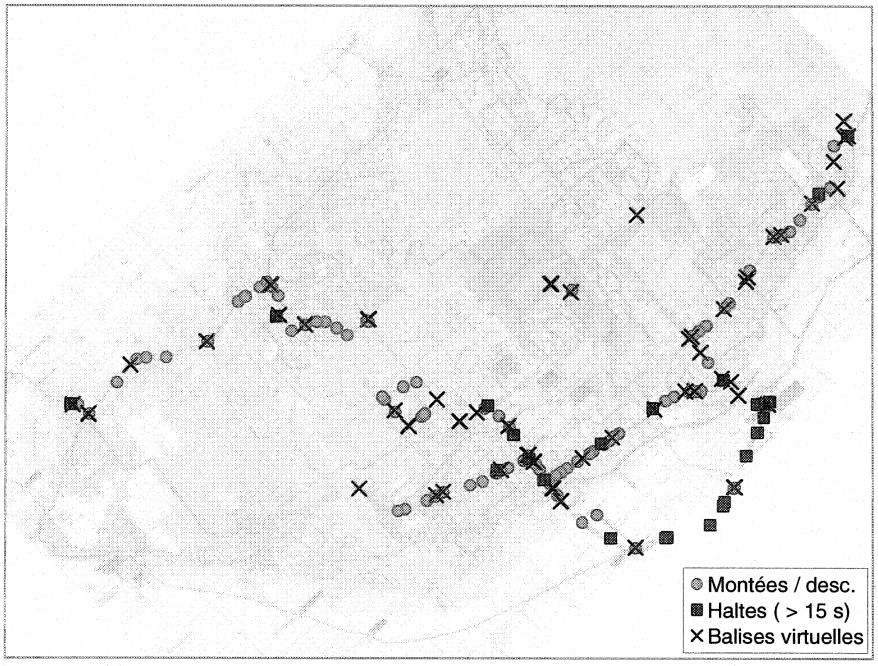


Figure 5-8 : Expérience 2 – données collectées par l'ordinateur embarqué

5.3.1 Résultats

La Figure 5-11 montre la charge en fonction de la distance parcourue pour chacun des circuits effectués. La Figure 5-9 et la Figure 5-10 ne tiennent pas compte de la distance parcourue mais servent à mettre en évidence les montées et les descentes.

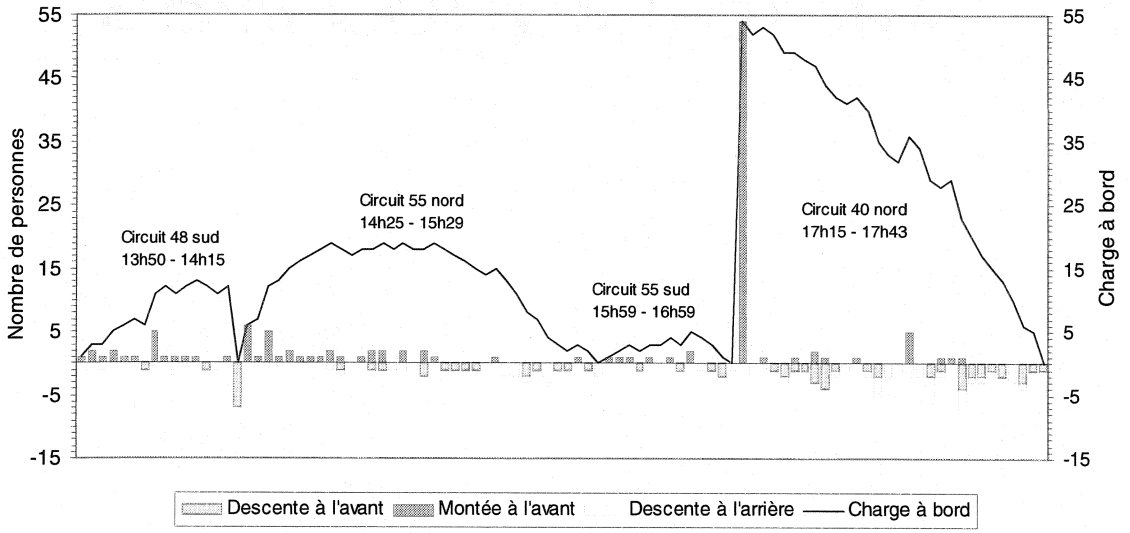


Figure 5-9 : Expérience 2 - montées, descentes et charge réelles à chaque arrêt (comptage manuel)

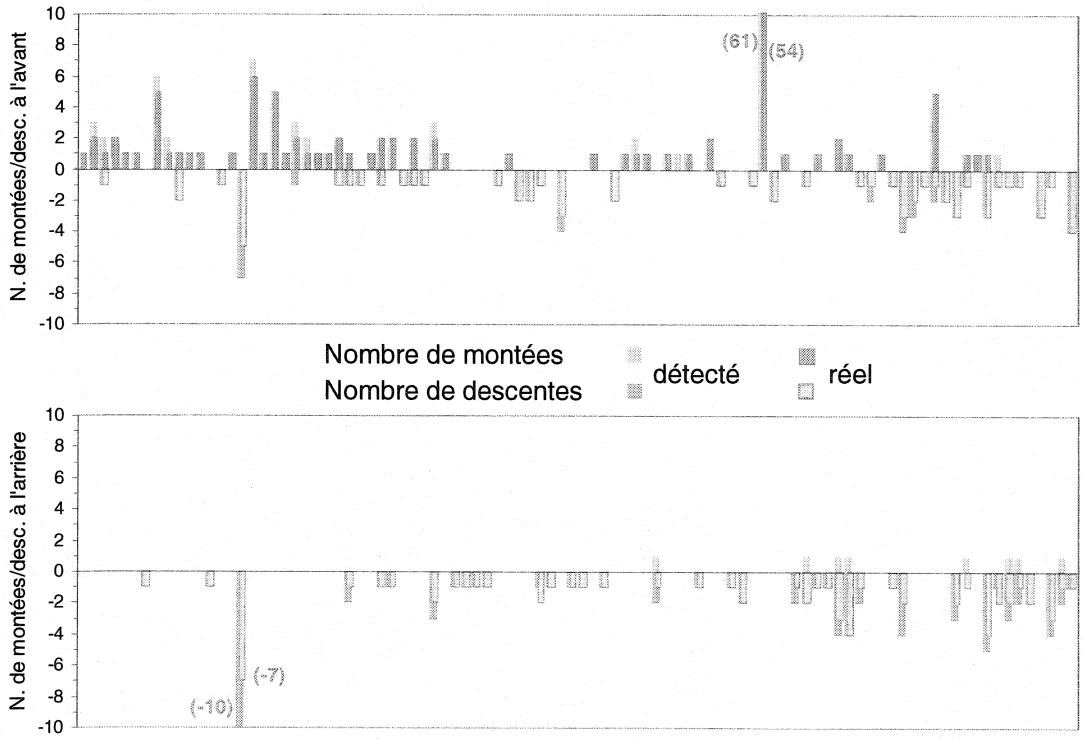


Figure 5-10 : Expérience 2 - montées et de descentes détecté et réel à chaque arrêt

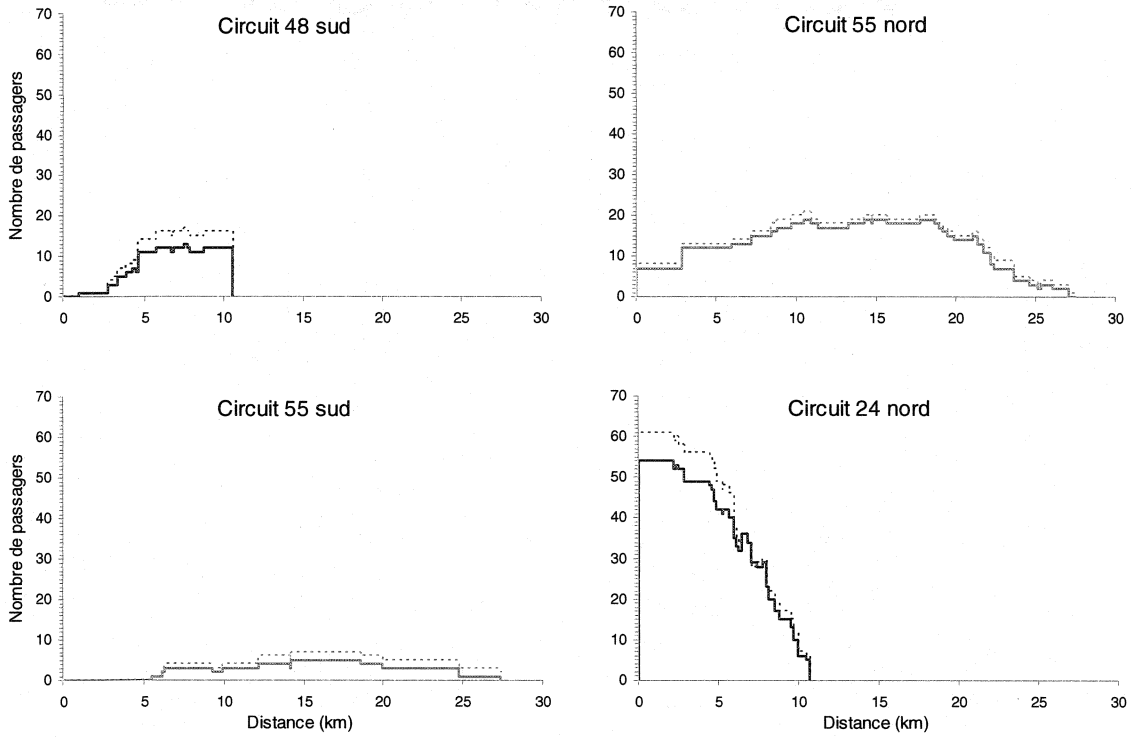


Figure 5-11 : Expérience 2 - profils de charge (pointillé = détecté, continu = réel)

Les erreurs obtenues sont les suivantes :

Erreur équilibrée : $\varepsilon_e = 18{,}7\ \%$

Erreur équilibrée en montée : $\varepsilon_e^m = 19{,}8\ \%$

Erreur équilibrée en descente : $\varepsilon_e^d = 17{,}5\ \%$

Erreur non-équilibrée : $\varepsilon = 21{,}8\ \%$

Écart moyen : $EM = 1{,}7\ \%$

D'après les données collectées par l'ordinateur embarqué, les profils de charge des quatre circuits contiennent une charge résiduelle très faible, soit de -1 à 2. Toutefois, lorsqu'on compare avec le comptage manuel, on découvre que les erreurs de comptage sont plus importantes. L'écart entre la charge calculée en parcours et la charge réelle mesurée atteint 8 personnes (circuit 24 nord à 4,7 km du point de départ).

Lorsqu'elle n'est pas nulle, la différence entre le comptage automatique et le comptage manuel est presque toujours positif; autrement dit, l'ordinateur tend à compter des personnes (montantes ou descendantes) en trop. Dans quelques rares cas le comptage automatique est inférieur au comptage manuel.

Autre détail à noter : lorsque l'autobus est bondé (ex. circuit 24 nord), le capteur situé à la porte arrière de l'autobus se met à détecter des personnes *montantes*. Puisque aucune montée n'a été observée à l'arrière, on peut donc admettre que la présence de passagers près de la porte déclenche un comptage à certains moments.

Évidemment, d'autres expériences semblables doivent être effectuées si l'on veut tirer des conclusions générales.

5.4 Troisième expérience

Une troisième expérience, semblable à la deuxième, a été effectuée afin d'accumuler plus d'informations sur la fiabilité des capteurs de mouvement directionnels.

L'autobus a effectué le voyage suivant :

- 7h16 : départ du garage vers le point d'injection du circuit 901 sud (Saint-François);
- 7h35 à 8h10 : circuit 901 sud (vers le terminus Laval);
- 8h20 à 9h04 : circuit 31 nord (vers Auteuil);
- 9h09 à 9h42 : circuit 31 sud (vers le terminus Laval);
- 9h57 : retour au garage

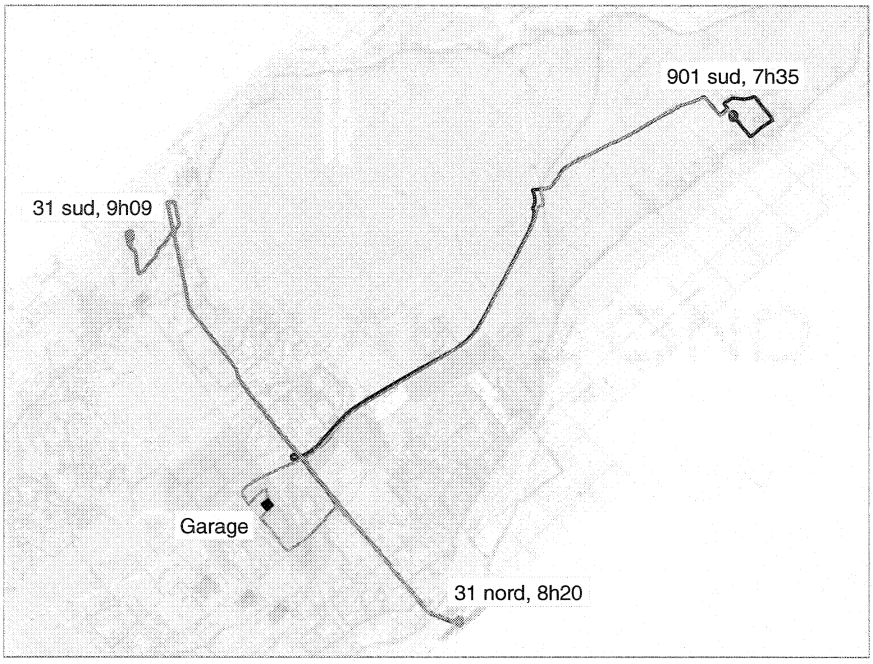


Figure 5-12 : Parcours effectué pendant l'expérience 3 réalisée le 13 mai 2004

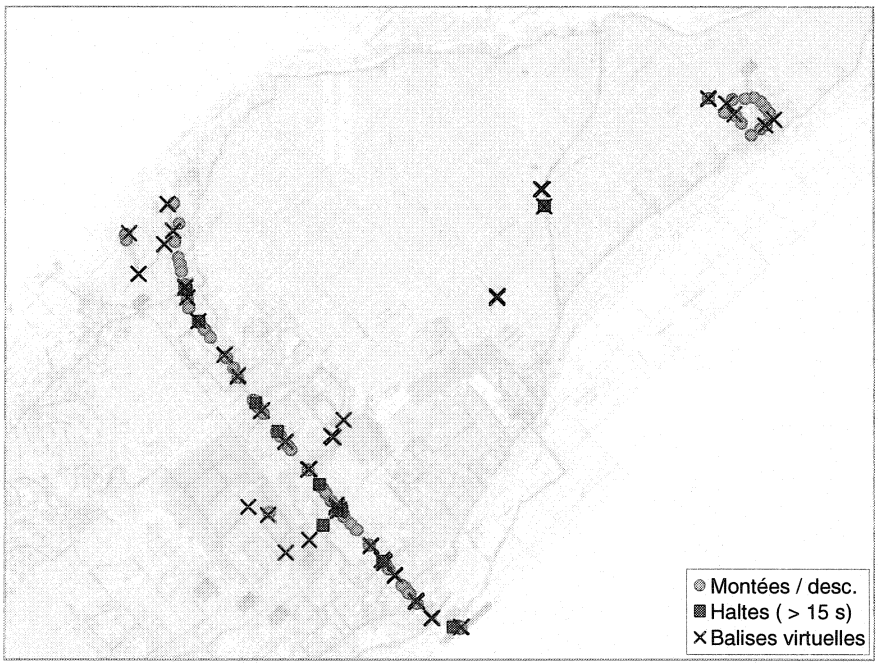


Figure 5-13 : Expérience 3 – données collectées par l'ordinateur embarqué

5.4.1 Résultats

Ici, la courbe de charge a été inversée pour éviter que le graphique ne prenne trop de place. Mais il est intéressant de constater le phénomène *inverse* à celui de l'expérience 2 : les gens se dirigent en grande majorité vers la station Henri-Bourassa alors que précédemment les usagers en majorité rentraient chez eux.

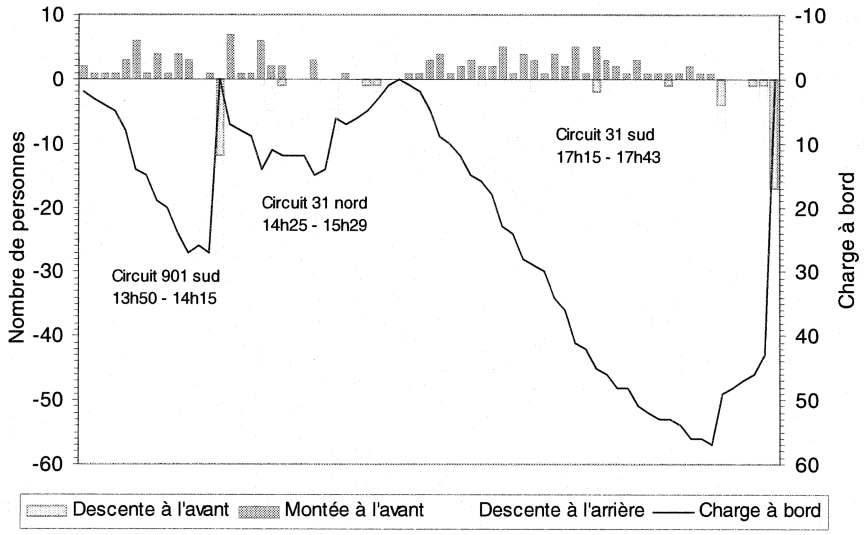


Figure 5-14 : Expérience 3 - montées, descentes et charge réelles à chaque arrêt (comptage manuel)

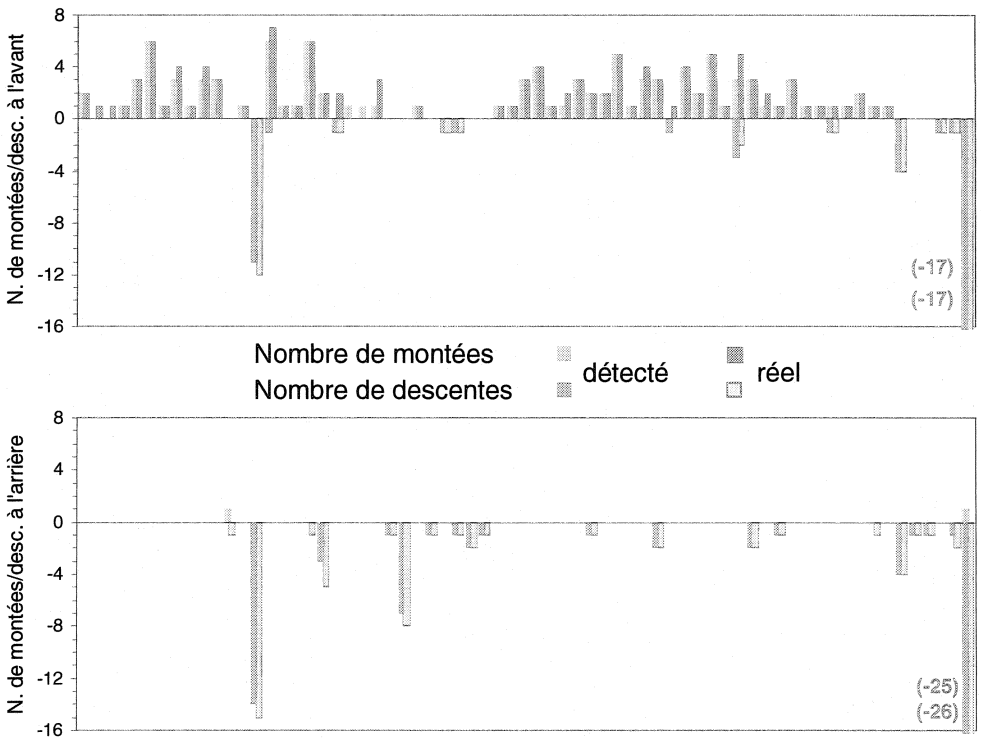


Figure 5-15 : Expérience 3 - montées et de descentes détecté et réel à chaque arrêt

Les erreurs obtenues sont les suivantes :

Erreur équilibrée : $\varepsilon_e = -7{,}6\ \%$

Erreur équilibrée en montée : $\varepsilon_e^m = -9{,}3\ \%$

Erreur équilibrée en descente : $\varepsilon_e^d = -5{,}9\ \%$

Erreur non-équilibrée : $\varepsilon = 13{,}6\ \%$

Écart moyen : $EM = 1{,}8\ \%$

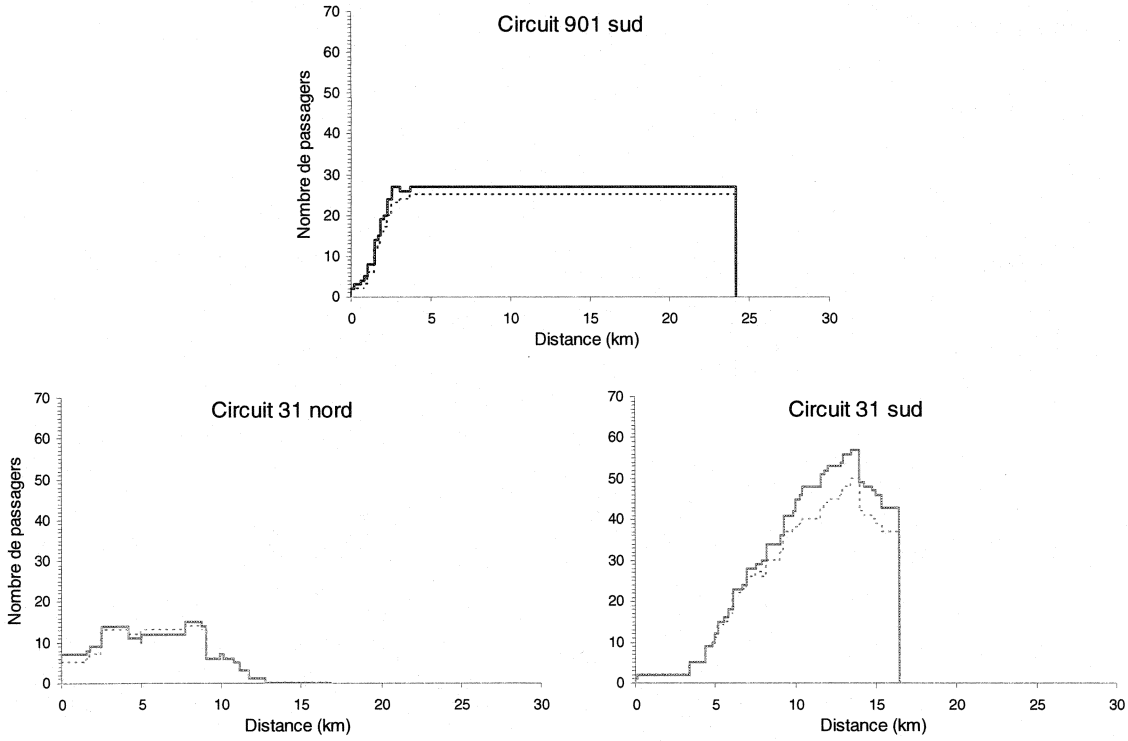


Figure 5-16 : Expérience 3 - profils de charge (pointillé = détecté, continu = réel)

Dans ce cas-ci, les profils de charge des trois circuits contiennent eux aussi une charge résiduelle nulle ou très faible, soit de 0 à -4. Et encore une fois, en comparant avec le comptage manuel on découvre que les erreurs de comptage sont un peu plus importantes. L'écart entre la charge calculée en parcours et la charge réelle mesurée atteint encore une fois 8 personnes (circuit 31 sud de 10,5 à 13,5 km du point de départ).

Mais cette fois, la différence entre le comptage automatique et le comptage manuel est presque toujours négatif; autrement dit, l'ordinateur tend à sous-estimer le nombre de personnes montantes ou descendantes. Dans quelques rares cas seulement le comptage automatique est supérieur au comptage manuel.

5.5 Conclusion

Les données obtenues avec un appareil GPS portatif mettent en évidence l'avantage d'une collecte continue de données par rapport à la collecte événementielle. Une résolution élevée rend possible une analyse plus désagrégée des mouvements de véhicules et facilite l'identification des circuits.

Il faut néanmoins s'interroger sur l'utilité réelle d'avoir des données plus riches, plus denses et plus détaillées, surtout lorsque l'on sait que cette solution entraîne des coûts et qu'il en résulte une base de données beaucoup plus volumineuse susceptible de causer des problèmes au niveau du traitement et de l'archivage. « *Deux fois plus n'est pas nécessairement deux fois mieux.* » (WATZLAWICK, 1986).

Est-ce que l'on veut assurer une meilleure reconnaissance des circuits? Une méthode de reconnaissance des circuits est proposée au chapitre 7 et les résultats obtenus expérimentalement montrent que la presque totalité des circuits peuvent être adéquatement identifiés seulement avec des données événementielles.

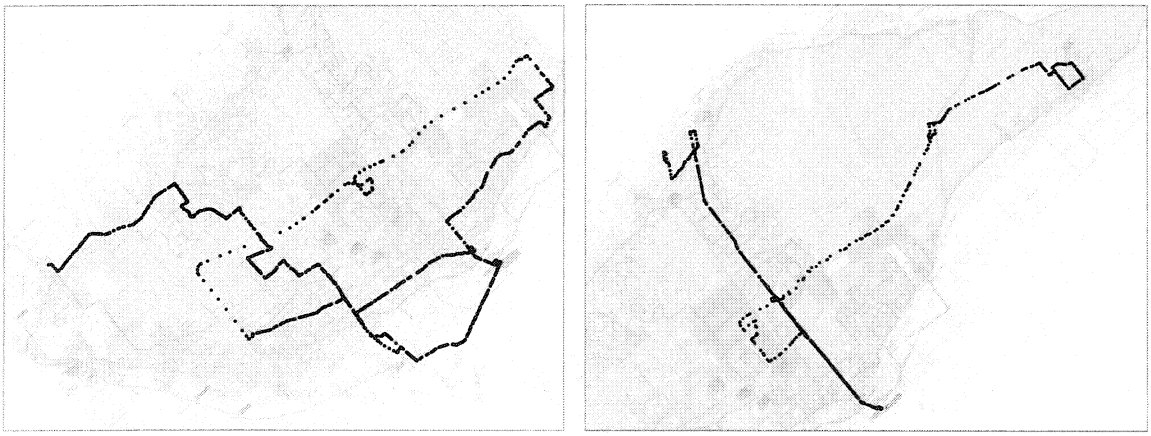


Figure 5-17 : Expériences 2 et 3 – données collectées par un appareil GPS portatif

D'un autre côté, des données d'une résolution très élevée permettront certes une meilleure analyse du mouvement des véhicules sur le réseau. Mais encore faut-il déterminer ce que l'on veut *réellement* et définir la façon dont on devra s'y prendre pour obtenir des résultats pertinents.

Tel que déjà mentionné à la section 3.6.1, il serait possible d'améliorer les systèmes actuels afin d'obtenir un compromis entre la résolution actuelle des données événementielles et celle obtenue ici avec l'appareil GPS portatif. D'un autre côté, il se pourrait qu'une compression adéquate des données dans les ordinateurs embarqués puisse bien un jour permettre un plus grand échantillonnage, voire même la collecte continue des données sur tout un réseau...

Le comptage manuel quant à lui met en lumière certaines erreurs cachées qui influencent les analyses sur l'achalandage. Les expériences ne sont pas assez nombreuses pour permettre une description générale du comportement des capteurs, mais elles donnent une idée de leur efficacité.

La STL effectue périodiquement un comptage manuel afin de vérifier la précision de leurs systèmes et Infodev intervient lorsqu'un capteur s'avère vraiment défectueux. L'ennui est que le comportement des capteurs est parfois imprévisible et leur dégradation est difficile à mesurer. Il est possible cependant de détecter certains types de mauvais fonctionnement à partir des données qu'ils produisent. Le chapitre 8 aborde ce sujet.

La réalisation de ces expériences fut prétexte à quelques représentations graphiques dont plusieurs se trouvent dans ce chapitre. D'autres représentations sont montrées au chapitre suivant.

CHAPITRE 6 : Visualisation des données pour fins d'interprétation

6.1 Graphiques à bulles

6.1.1 Achalandage sur un parcours

L'achalandage d'une ligne de transport est maintenant examiné dans un contexte géographique. Cette réalité est représentée ici par des graphiques à bulles. Les graphiques à bulles (graphiques « 2½D ») exploitent trois variables à la fois et s'approchent des graphiques en trois dimensions (3D), la troisième dimension étant représentée ici par la taille des bulles.



Figure 6-1 : Achalandage sur les circuits effectués pendant l'expérience 1

Les exemples de la Figure 6-1 sont des parcours effectués séquentiellement par le même autobus un même avant-midi. Deux informations sont représentées en même temps : les montées (bulles vert foncé) et les descentes (bulles vert clair).

Un graphique à bulles nécessite trois séries de données : coordonnées en X, coordonnées en Y et aire des bulles.

Jc071931a.xls

	A	B	C	D	E	F	G	O	P
1	INFODEV	DL10B	1.1	SL0011	31				
2	Type	Heure D	Heure F	Distance	Parcours	Long	Lat	Montées	Descentes
3	V	2003-07-18 23:53	2003-07-18 23:53	0		287082	5049006		
4	B	2003-07-18 04:06	2003-07-18 04:06	4		287101	5049067		
5	B	2003-07-18 04:15	2003-07-18 04:15	0		287101	5049067		
6	P	2003-07-18 04:05	2003-07-18 04:15	2		287101	5049067	5	3
7	E	2003-07-18 04:17	2003-07-18 04:17	166		0	0		
8	B	2003-07-18 06:13	2003-07-18 06:13	0		0	0		
9	P	2003-07-18 04:17	2003-07-18 06:15	3		0	0	2	3
10	P	2003-07-18 06:15	2003-07-18 06:15	28		0	0	2	2
11	V	2003-07-18 06:17	2003-07-18 06:17	819		286617	5049197		
12	H	2003-07-18 06:27	2003-07-18 06:27	12088		282877	5043689		
13	V	2003-07-18 06:27	2003-07-18 06:27	20	24 1 (2)	282861	5043678		
14	P	2003-07-18 06:28	2003-07-18 06:28	50	24 1 (2)	282858	5043628	1	0
15	P	2003-07-18 06:28	2003-07-18 06:29	199	24 1 (2)	283030	5043617	1	0
16	P	2003-07-18 06:29	2003-07-18 06:29	235	24 1 (2)	283264	5043676	1	0
17	P	2003-07-18 06:30	2003-07-18 06:30	194	24 1 (2)	283457	5043724	1	0
18	P	2003-07-18 06:30	2003-07-18 06:30	301	24 1 (2)	283727	5043871	2	0
19	P	2003-07-18 06:31	2003-07-18 06:31	211	24 1 (2)	283918	5043976	2	0
20	V	2003-07-18 06:31	2003-07-18 06:31	35	24 1 (2)	283946	5043995		
21	P	2003-07-18 06:32	2003-07-18 06:32	149	24 1 (2)	284083	5044067	2	0
22	V	2003-07-18 06:32	2003-07-18 06:32	41	24 1 (2)	284121	5044084		
23	P	2003-07-18 06:32	2003-07-18 06:32	167	24 1 (2)	284288	5044128	1	0
24	P	2003-07-18 06:33	2003-07-18 06:33	439	24 1 (2)	284721	5044253	3	0
25	P	2003-07-18 06:34	2003-07-18 06:34	350	24 1 (2)	285068	5044354	2	0
26	P	2003-07-18 06:35	2003-07-18 06:35	449	24 1 (2)	285439	5044614	1	0
27	V	2003-07-18 06:35	2003-07-18 06:35	24	24 1 (2)	285461	5044623		
28	P	2003-07-18 06:36	2003-07-18 06:36	141	24 1 (2)	285593	5044682	1	0
29	P	2003-07-18 06:36	2003-07-18 06:36	203	24 1 (2)	285780	5044772	2	0
30	P	2003-07-18 06:37	2003-07-18 06:37	132	24 1 (2)	285902	5044828	2	0
31	P	2003-07-18 06:37	2003-07-18 06:38	204	24 1 (2)	286093	5044915	3	5
32	V	2003-07-18 06:38	2003-07-18 06:38	157	24 1 (2)	286200	5044812		

Figure 6-2 : Extrait modifié d'un fichier de données LADRA ouvert dans Excel

6.1.2 Circonstances temporelles d'un parcours

Chaque arrêt de l'autobus constitue un événement et provoque la prise de données par le système GPS à bord. Ces données permettent de connaître la durée de chacun de ces arrêts.

Tel qu'expliqué au chapitre 3, il existe deux sortes d'événements reliés aux arrêts : l'ouverture des portes et la halte de plus de 15 s sans ouverture de porte.

La durée de chacun de ces événements est représentée le long du voyage par des bulles bleues et rouges respectivement.

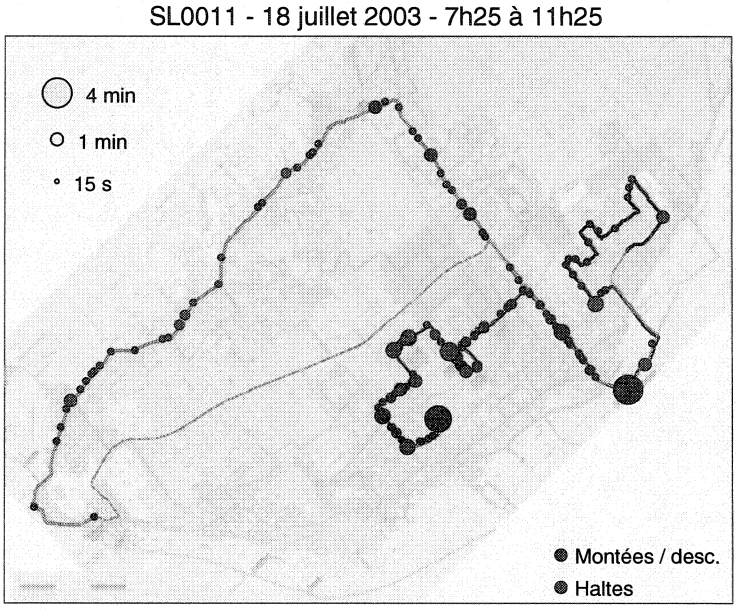


Figure 6-3 : Durée des arrêts sur le parcours effectué pendant l'expérience 1

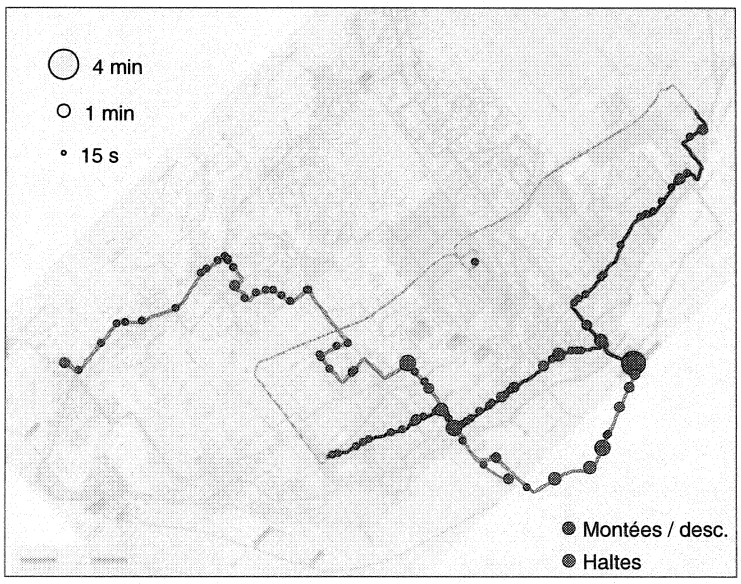


Figure 6-4 : Durée des arrêts sur le parcours effectué pendant l'expérience 2

6.2 Diagrammes espace-temps

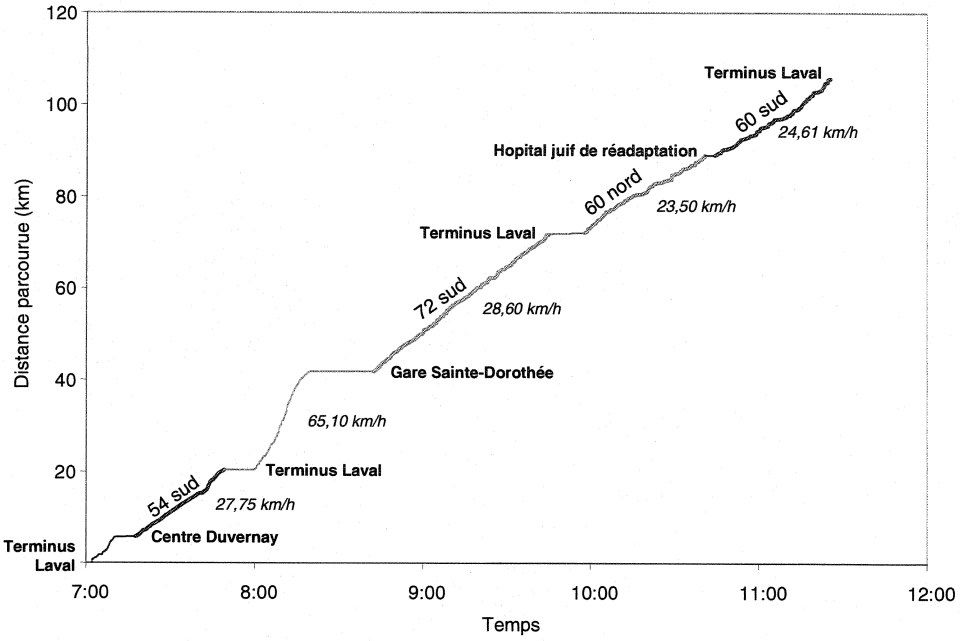


Figure 6-5 : Expérience 1 - diagramme espace-temps

La Figure 6-5 représente le déplacement de l'autobus SL0011 lors de l'expérience 1. La pente illustre la vitesse du véhicule et les nombres indiquent la vitesse commerciale (i.e. la distance parcourue divisée par le temps entre le terminus de départ et celui d'arrivée) pour chaque circuit. À noter la pente prononcée entre le circuit 54 et le circuit 72 qui correspond bien sûr au voyage à vide qui a été effectué en partie sur l'autoroute.

6.3 Réalité virtuelle

À quoi aurait l'air une trace GPS avec une coordonnée *z* égale au temps écoulé? Comme les outils graphiques de Microsoft Excel ne permettent pas de répondre directement à cette question (il n'existe pas de graphique « nuage de points 3D »), le langage VRML fut utilisé. Le VRML (*Virtual Reality Modeling Language*) est un langage public conçu pour la modélisation en 3D et la description de scènes virtuelles.

Les deux schémas ci-dessous furent réalisés avec le VRML. Le premier schéma est une représentation 3D du parcours effectué pendant l'expérience 1; la coordonnée verticale correspond au temps écoulé pendant l'expérience. Le second schéma représente le même parcours, mais cette fois-ci la coordonnée verticale correspond à l'heure de départ de chaque circuit effectué.

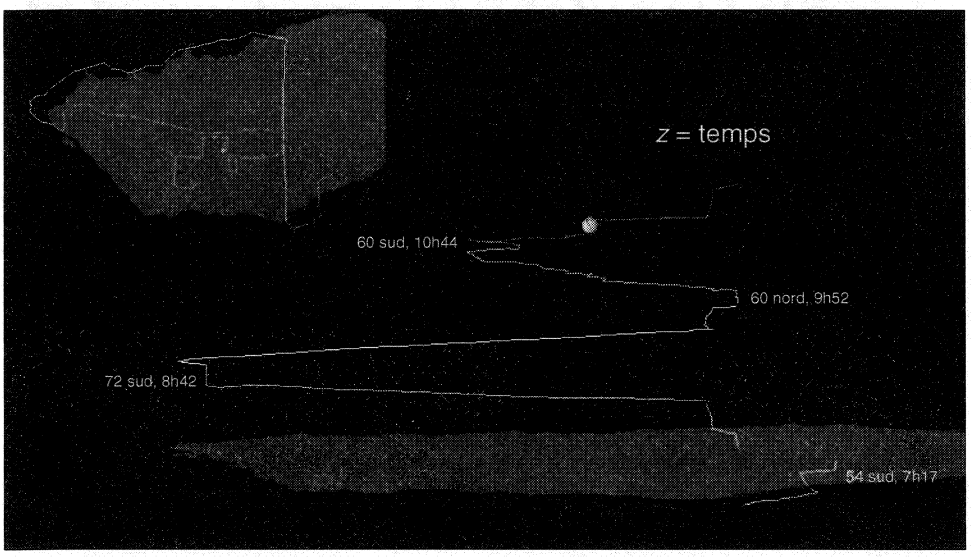


Figure 6-6 : Représentation en VRML du parcours de l'expérience 1 (*z* = temps)

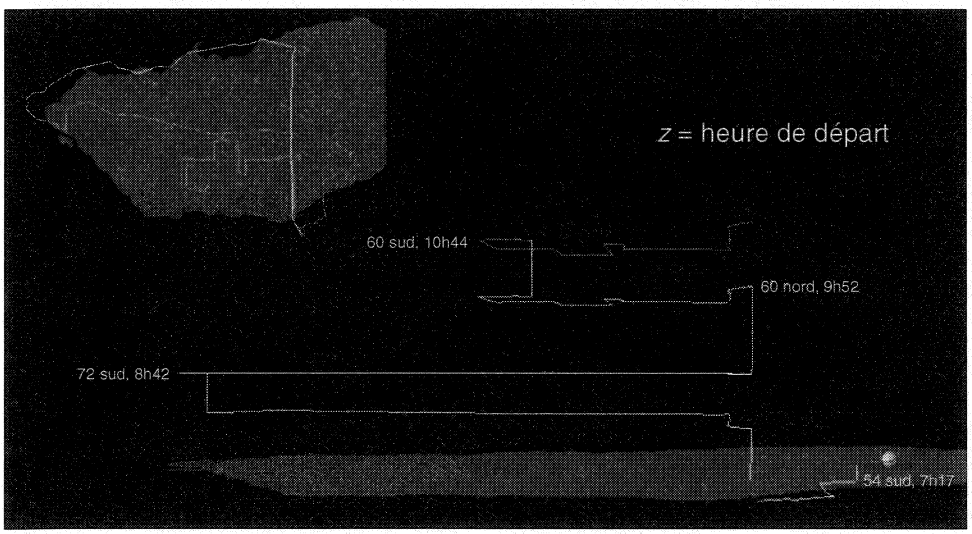


Figure 6-7 : Représentation en VRML du parcours de l'expérience 1 (*z* = heure de départ)

Il est possible de modéliser des objets mobiles en VRML. Ici, la bille grise présente dans les deux figures se déplace selon une matrice décrivant son emplacement à chaque instant. Cette technique ne permet malheureusement pas une animation en temps réel, du moins pas de façon directe. Un langage plus récent et plus sophistiqué, le X3D, permet l'animation en temps réel et est compatible XML (www.web3d.org).

Sa compatibilité avec Internet et le HTML, ainsi que sa structure modulaire orientée-objet, fait du VRML un outil extrêmement intéressant dans le monde de la technologie informationnelle, surtout lorsque combiné aux systèmes d'information géographiques (GIS). Mentionnons en particulier les travaux de KWAN (2003) qui illustre avec beauté les déplacements d'individus dans plusieurs villes américaines à l'aide de courbes espace-temps et de données GIS.

Mentionnons aussi les travaux de codification et de modélisation des stations de métro de la Société de transport de Montréal (STM) effectués par MADITUC (CHAPLEAU et al, 2003) qui pourraient éventuellement servir à informer la clientèle de la STM, comme McDONOUGH (2002) le démontre en prenant l'exemple de la station Henri-Bourassa.

CHAPITRE 7 : Reconnaissance des tournées

7.1 Introduction

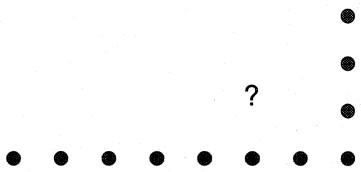


Figure 7-1 : Exemple de trace GPS

L'un des problèmes les plus importants rencontrés dans l'analyse des données collectées par les autobus est l'absence d'information sur les circuits d'autobus qui ont été parcourus. Le GPS localise le véhicule sans égard à l'infrastructure urbaine ou à l'organisation d'un quelconque réseau de transport et aucune contribution extérieure ne permet de compléter l'information de façon directe.

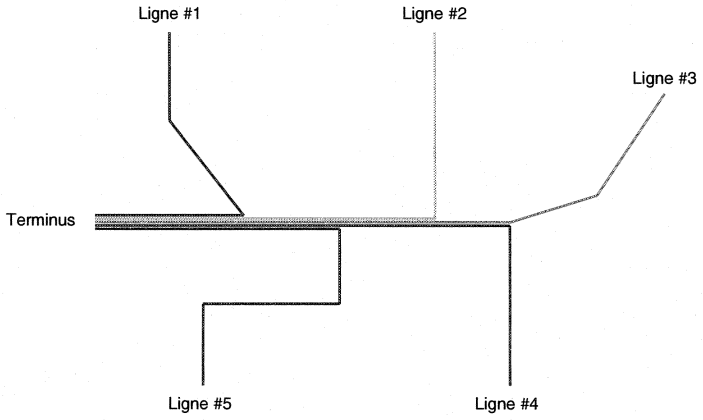


Figure 7-2 : Exemple de réseau de transport

Pour pallier cette lacune, les données collectées sont couplées avec les informations sur le réseau d'autobus. Cette technique s'apparente au *map matching*, soit la correspondance entre la trace GPS d'une automobile et le réseau routier sur lequel elle circule. Par contre, dans le cas d'un autobus, ce ne

sont pas des tronçons de routes isolés que l'on doit identifier mais des circuits entiers.

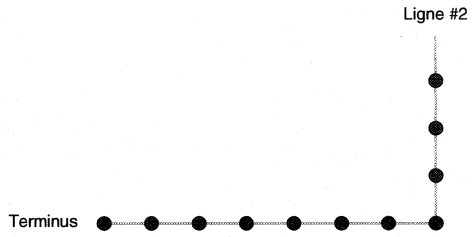


Figure 7-3 : Exemple de fusion des données GPS avec les données planifiées

7.2 Problématique

Comme il a été expliqué au chapitre 3, la reconnaissance (l'identification) des circuits et des tournées à la STL se fait respectivement par BusStops et LADRA. BusStops compare les données GPS avec la description des circuits composant le réseau, puis LADRA effectue un couplage avec les horaires planifiés. À première vue, la majorité des tournées effectuées sont reconnues et bien que BusStops semble commettre un certain nombre d'erreurs, très peu d'entre elles passent inaperçues avec LADRA.

Toutefois, on remarque dans cette méthode une faiblesse importante : *les informations spatiales sont séparées des informations temporelles*. En effet, lorsque BusStops analyse une portion des données GPS et identifie un circuit potentiel, il ne peut vérifier s'il y a concordance avec les horaires planifiés. Bien entendu, c'est le travail de LADRA mais lorsque plusieurs circuits semblent concorder avec le tracé GPS ou que très peu d'informations permettent d'identifier un circuit avec certitude, BusStops est forcé de faire un choix qui ne sera pas forcément le bon et LADRA ne peut corriger ce choix. Il peut seulement l'invalider.

Il est cependant curieux de constater jusqu'à quel point le logiciel BusStops peut se tromper (Figure 4-3). Même sans un accès direct aux horaires planifiés, on ne peut s'empêcher de considérer que l'algorithme utilisé ici, quel qu'il soit, est *améliorable* (à moins qu'il ait été induit en erreur par les données planifiées qu'on lui a fournies, ce qui est une possibilité).

Ce chapitre propose une méthode de reconnaissance de circuits quelque peu différente de celle impliquée ici. Cette méthode a l'avantage d'être plus robuste, de causer moins d'erreurs et d'avoir un meilleur rendement, particulièrement au niveau de l'identification des tournées.

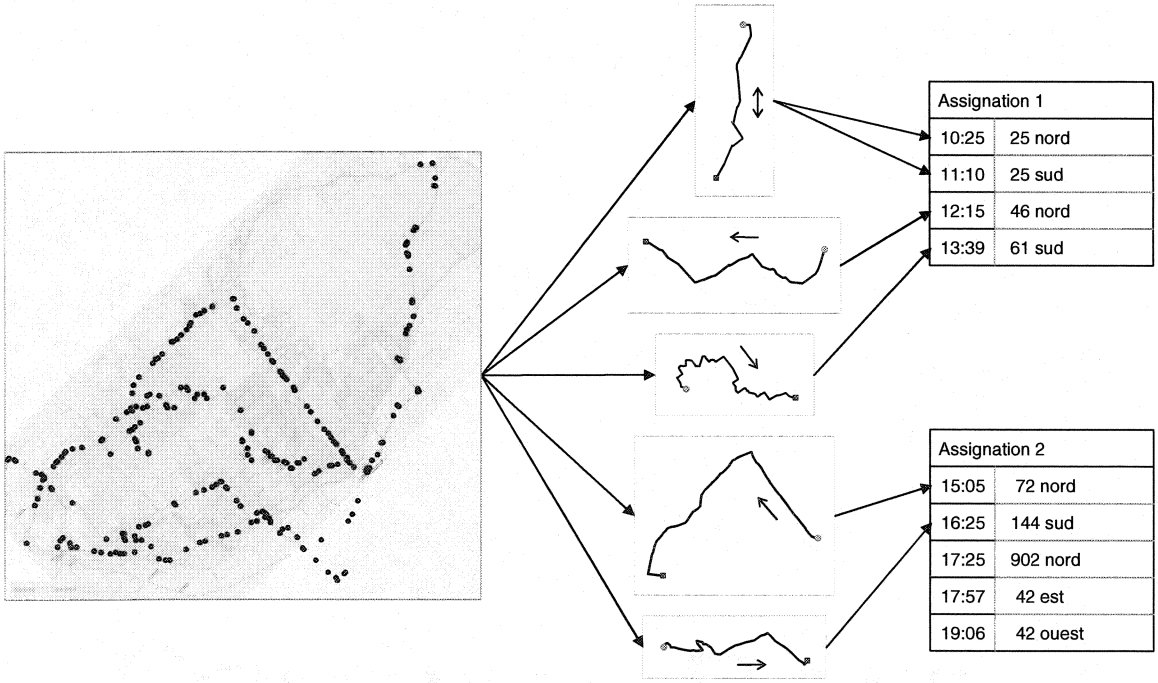


Figure 7-4 : Procédure de reconnaissance des tournées

Un logiciel expérimental fut conçu selon cette méthode et ses résultats sont comparés avec ceux de BusStops/LADRA plus loin dans ce chapitre. Le but de ce chapitre n'est pas d'expliquer en détail comment programmer un tel logiciel mais seulement de démontrer que l'interprétation des données GPS qui fut

effectuée jusqu'ici peut être notablement améliorée, au point de permettre certains résultats graphiques (ex. : profils de charges 3D au chapitre 9).

À noter que la version de BusStops concernée ici et actuellement utilisée par la STL est une *ancienne* version. Une nouvelle version de BusStops a été développée par Infodev et est présentement utilisée par le RTC et le RTL. Cette nouvelle version exploite simultanément les informations spatiales et temporelles planifiées (Figure 3-7). Par contre, jusqu'à la fin de ce chapitre, il sera question de la version *antérieure* de BusStops, qui n'exploite que les informations spatiales (circuits planifiés).

7.3 Principes

Du point de vue des systèmes d'information géographiques, un réseau routier est un ensemble de nœuds (intersections) reliés entre eux par des liens (tronçons de rue). Une rue est donc un assemblage de tronçons de rues entre lesquelles on retrouve des intersections.

Similairement, un réseau de transport collectif (TC) est un ensemble de nœuds reliés par des lignes. Chaque nœud représente un arrêt, une station, un terminus ou une gare, en bref un point de service (aire d'embarquement et de débarquement) pour les usagers. Chaque ligne relie séquentiellement un ensemble de nœuds dans l'ordre où il doit être desservi.

Les dualités lien-nœuds routiers et ligne-nœuds TC sont représentées à la Figure 7-5.



Figure 7-5 : Relations lien-noeuds routiers et ligne-noeuds TC

À noter que la symbolique orientée-objet utilisée ici est simplifiée. Certains documents (ex. TRÉPANIER, 1999) considèrent les liens et les nœuds routiers comme étant des objets cinétiques et distinguent les liens routiers (objets cinétiques) des tronçons de rues (objets statiques). Ici, aucune distinction n'est faite pour fins de simplicité et parce qu'on fait abstraction de l'automobile (objet dynamique).

Dans le cas d'un réseau d'autobus, toutes les lignes de transport passent par des liens routiers. Par contre, pour d'autres types de véhicules (ex. trains, métro), la ligne correspondante passe ailleurs. Inversement, un lien routier peut contenir une, plusieurs ou aucune ligne de transport.

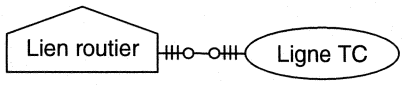


Figure 7-6 : Relation ligne TC - lien routier



Figure 7-7 : Relation ligne TC - lien routier (variante)

Les systèmes de collecte de données Infodev sont événementiels et la plupart des événements qui surviennent lors d'un voyage sont reliés à des points de service (en particulier les ouvertures de portes et les balises virtuelles). On peut donc associer la plupart des coordonnées GPS (localisations) stockées à bord des véhicules à des nœuds TC.

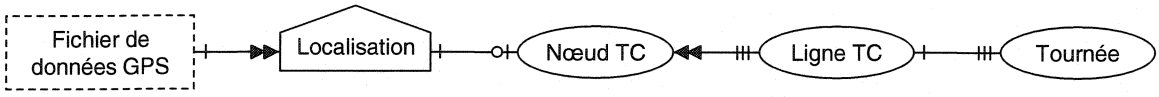


Figure 7-8 : Reconnaissance des tournées avec les nœuds TC

Si les données étaient collectées en continu et qu'il était possible de les comparer à une base de données décrivant le réseau routier, nous aurions alors la relation suivante :

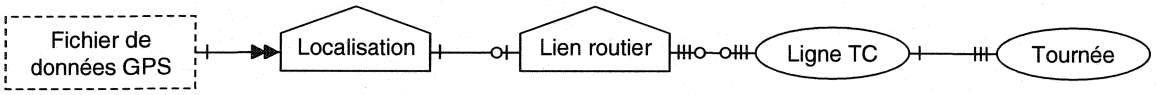


Figure 7-9 : Reconnaissance des tournées avec les liens routiers

Par ailleurs, il existe aussi un rapport entre les tournées effectuées et l'autobus produisant les données GPS. Par contre, ce rapport ne peut être exploité que si l'on sait exactement quel chauffeur utilise quel autobus et sous quelle assignation. Autrement, il est néanmoins utile de retenir que, sauf exception, un véhicule effectuera des séquences de tournées qui correspondront aux assignations des chauffeurs qui le conduisent. En d'autres termes, tout comme la reconnaissance de nœuds TC permet de deviner les lignes TC effectuées, la reconnaissance des tournées permet de deviner les assignations effectuées. La reconnaissance des assignations, elle, permet de mieux identifier les tournées.

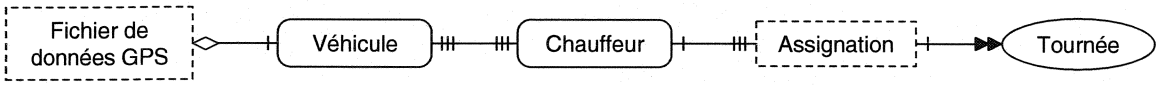


Figure 7-10 : Reconnaissance des tournées par les assignations

Finalement, on peut imaginer un ordinateur embarqué capable de communiquer avec le système d'affichage du véhicule (le dispositif qui affiche sur l'autobus le numéro et le nom du circuit effectué). Un tel système faciliterait grandement la reconnaissance des tournées.

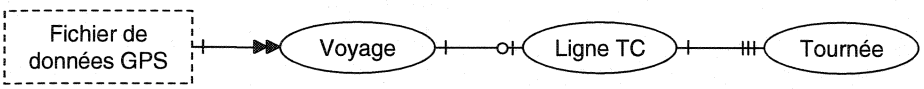


Figure 7-11 : Reconnaissance des tournées par un lien avec le système d'affichage

Si l'on réunit tous les concepts abordés ici, notre système de reconnaissance des tournées pourrait ressembler à ceci :

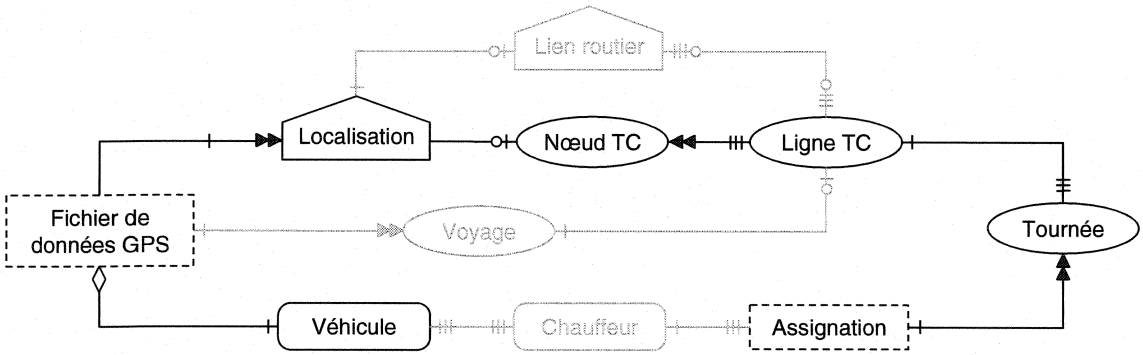


Figure 7-12 : Reconnaissance des tournées - vue d'ensemble

Les données en trait foncé sont celles qui sont utilisées dans les pages qui suivent et dans l'élaboration du logiciel expérimental.

7.4 Sources d'information

Plusieurs sources d'information permettant la reconnaissance des circuits. Citons les plus importants :

7.4.1 Liste des nœuds

Une ligne de transport est habituellement une entité discrète (par opposition à continue) puisque le véhicule s'arrête à des endroits quasi ponctuels définis à l'avance, généralement des points de service (cette règle est confirmée par quelques exceptions; par exemple, le service « Entre deux arrêts » de la STM et de la STL). Une liste de ces points accompagnés de leur localisation géographique permet de connaître avec précision les endroits géographiques où les embarquements et les débarquements ont lieu théoriquement et de deviner le circuit qui est effectué.

7.4.2 Liste des lignes de transport

Une base de donnés décrivant les lignes de transport d'un réseau devrait inclure la liste des points de service desservis par chaque ligne ainsi que leur séquence et divers renseignements particuliers tels les points de contrôle (*checkpoints*), l'accessibilité (montée/descente permise?), les boucles et les balises virtuelles.

7.4.3 Liste des horaires

Une telle liste décrit toutes les tournées qui doivent être effectuées par les chauffeurs sur chacune des lignes composant le réseau de transport. Ces tournées sont regroupées en *assignations*, soient des séquences de tournées que les chauffeurs doivent respecter.

7.4.4 Liste des jours fériés

Un jour férié est très différent d'un jour ouvrable en terme d'achalandage et, conséquemment, la liste des horaires sera différente – tout comme elle l'est pour les samedis et les dimanches. La détermination du jour est donc une étape nécessaire dans la reconnaissance des tournées.

7.5 Méthode

Les étapes suivantes sont proposées pour la reconnaissance des circuits :

1. Identification des nœuds;
2. Identification des circuits associés aux nœuds;
3. Vérification des circuits et élimination;
4. Identification des tournées potentielles;
5. Vérification des tournées à l'aide des assignations.

7.5.1 Identification des nœuds

Dans la procédure événementielle de collecte des données, telle que décrite dans ce document, les points qui sont enregistrés en mémoire correspondent la plupart du temps à des arrêts d'autobus (ouverture des portes et balises virtuelles), des intersections (haltes de plus de 15 secondes), des terminus (coupure ou démarrage du moteur) ou le garage (téléchargement des données).

Il est alors possible d'effectuer une concordance entre les données collectées et les nœuds qui composent chacune des lignes de transport. Ainsi, pour chaque point enregistré dans l'ordinateur embarqué, on cherche le nœud qui est le plus près dans un certain rayon. Ce rayon correspond à l'écart entre la position théorique des nœuds et la position des points de tracé, écart attribuable au fait que le véhicule ne se trouve pas exactement aux endroits théoriques (i.e. arrêt décalé par rapport à l'arrêt d'autobus), mais aussi aux erreurs de positionnement et aux erreurs dans les données théoriques. Dans notre cas, le rayon acceptable est de l'ordre d'une centaine de mètres.

7.5.2 Identification des circuits associés aux noeuds

L'identification des circuits parcourus est simplifiée lorsqu'on connaît les arrêts d'autobus desservis. L'identification de chaque circuit doit cependant respecter certaines conditions :

- Le véhicule doit passer par plus d'un nœud appartenant au circuit;
- Les nœuds parcourus doivent être dans le bon ordre; cette dernière condition permet de déterminer la direction du circuit (la plupart des circuits sont bidirectionnels).

7.5.3 Vérification des circuits et élimination

À partir des circuits potentiels identifiés, il est possible dans la plupart des cas de filtrer les circuits erronés et de ne retenir que les circuits qui ont véritablement été effectués.

Un circuit authentique répond aux critères suivants :

- Le véhicule passe par le terminus de départ;
- Le véhicule se rend jusqu'au terminus d'arrivée – ou du moins à proximité;
- Entre le départ et l'arrivée, le véhicule ne s'éloigne pas plus d'une certaine distance (de l'ordre de plusieurs centaines de mètres ou une fraction de la distance entre les arrêts, le plus élevé des deux valeurs) du circuit;
- Un départ est prévu au moment ou le véhicule quitte le terminus de départ;
- Le véhicule croise la plupart des balises virtuelles situées sur le circuit.

Si, sur une certaine distance, tous les circuits potentiels sont éliminés, alors la distance en question correspond probablement à un voyage à vide ou un voyage particulier (ex. scolaire, dépannage).

7.5.3.1 Découpage

Avant de passer à l'étape suivante, un petit détail doit être éclairci : le découpage des circuits identifiés.

Cas 1 : Un autobus arrivant à un terminus entre deux circuits peut effectuer les actions suivantes : débarquement des passagers, stationnement, coupure du moteur, pause, démarrage du moteur, positionnement à l'aire d'embarquement, embarquement des passagers et départ. Les actions détectées par l'ordinateur embarqué seront habituellement les suivantes : ouverture et fermeture des portes, coupure du moteur, ouverture des portes, démarrage du moteur,

fermeture des portes, ouverture et fermeture des portes (une ou plusieurs fois, ou pendant de longues minutes), détection de balise.

À partir d'informations aussi limitées, trois questions se posent : Quand se termine le circuit précédent? Quand commence le circuit suivant? Quelle est l'heure réelle de départ à partir du terminus? Avant de répondre, examinons deux autres cas typiques.

Cas 2 : Arrivée à un terminus, débarquement des passagers, coupure du moteur, pause (à l'aire d'embarquement), embarquement des passagers, démarrage du moteur et départ. Ici, les actions enregistrées sont : ouverture des portes, coupure du moteur, démarrage du moteur, fermeture des portes et détection de balise.

Cas 3 : Il peut arriver aussi que l'autobus n'effectue qu'un simple arrêt puis reparte immédiatement pour effectuer son circuit suivant. Dans ce cas, seules les actions d'ouverture et de fermeture des portes seront enregistrées avec la détection de balise.

Le découpage des circuits doit permettre le comptage de tous les descendants du circuit précédent et de tous les montants du circuit suivant. Le circuit suivant ne peut être considéré comme commencé après la dernière ouverture des portes puisque des passagers peuvent être montés avant (cas 1). Il ne peut non plus commencer au démarrage du moteur puisque encore une fois des montées ne seront pas comptées (cas 2). Et dans certains cas, les deux circuits devront bien se chevaucher puisqu'une seule ouverture de porte se produit dans le cas 3.

Pour ce qui est de l'heure réelle de départ, ce ne peut être lorsque le chauffeur démarre son véhicule ou lorsqu'il ouvre ses portes puisqu'il peut être en avance

à ce moment-là. Ce ne peut être non plus lors de la détection de balise puisqu'il a déjà quitté le terminus lorsque cet événement se produit (par contre, cette solution sera retenue s'il n'y a aucune ouverture de portes).

La méthode proposée ici ne cherche pas à faire une analyse laborieuse des événements aux terminus mais avance une solution simple et directe : le circuit précédent se termine au dernier événement se produisant au terminus, le circuit suivant commence au premier événement se produisant au terminus et l'heure de départ réelle correspond au moment où le chauffeur referme les portes *pour la dernière fois* au terminus.

A priori, on peut craindre que ce genre de découpage entraîne des dédoublements de comptage aux terminus mais ce problème est facilement résolu en ne comptant que les descentes pour le circuit précédent et les montées pour le circuit suivant (voir section 8.1). Le seul problème qui n'est pas résolu est la montée et la descente du chauffeur (le système de comptage ne fait pas de discrimination). Ce problème est plus difficile à résoudre mais n'est pas majeur puisque dans la pire des situations – et en supposant que le comptage est parfait –, la différence de charge ne sera que de ±1 personne.

7.5.4 Identification des tournées potentielles

L'identification des tournées ressemble quelque peu à l'identification des nœuds. Pour chaque circuit identifié, on cherche la tournée qui correspond le plus possible à l'heure réelle de départ du terminus. Cette identification peut être plus difficile lorsque le véhicule effectue sa tournée pendant une période où la fréquence de passage est plus élevée, sachant qu'un autobus peut être en retard ou en avance et que l'heure réelle de départ est parfois difficile à déterminer à partir des données collectées (voir section 7.5.3.1).

On retient alors toutes les tournées ayant une heure de départ à l'intérieur d'une marge d'incertitude avant et après l'heure réelle. Si aucune tournée n'est retenue pour un circuit donné, alors le circuit en question est rejeté. Cela se produit régulièrement lorsqu'un voyage à vide est confondu avec un voyage en circuit (ex. le voyage à vide du terminus Laval au point d'injection du circuit 10 sud est identique à un voyage du circuit 10 nord).

7.5.5 Sélection des tournées par élimination

Des conflits d'horaire peuvent survenir pour deux raisons :

- Plusieurs circuits ont été identifiés pour un trajet donné et tous correspondent à un horaire planifié.
- Pour un circuit donné, plusieurs tournées ont été identifiées.

Ce problème peut être résolu en comparant les résultats avec les assignations des chauffeurs. Une assignation est une séquence de circuits qu'un chauffeur doit effectuer lorsqu'il prend possession de son véhicule (Figure 7-13).

L I S T E D U M A S T E R 2 (MARS 2004)

DJ = Debut Journee GV = Garer Vehicule E = Ecole
FJ = Fin Journee PP = Placer autobus & DH = Dead head
RP = Repas Prendre passagers S ou SP = Voyage special
DC = Dhoc D = Voyage double

CIRC.	VOY.	TYPE	DEPART	POUR	ARRIVE	KILOMETRAGES REG.	KILOMETRAGES DHOC
				(ME.)	ASS: L 1 4 3		
		DJ	06h43	GARAGE /			
		DH	06h53	BOUL. LE CORBUSIER /CENTRE LAVAL	07h00	3.2	
42	3		07h00	CHARTRAND /DERY	07h55	23.3	
42	18		08h09	Terminus Carrefour /	09h15	29.2	
35	18		09h22	TERMINUS LAVAL /MONTREAL	10h01	15.2	
35	13		10h20	Terminus Carrefour /	10h59	15.3	
		DH	10h59	GARAGE /	11h07	3.2	
					11h12		
		RP					
			13h21				
		DH	13h36	Saint-François /Pénitencier	13h50	10.2	
48	40		13h50	TERMINUS LAVAL /MONTREAL	14h15	10.3	
55	19		14h25	12 IEME RUE /BOUL. SAUVE	15h29	28.2	
55	38		15h59	TERMINUS LAVAL /MONTREAL	16h59	27.8	
24	57		17h15	MONTGOLFIER /NOTRE-DAME	17h43	10.8	
		DH	17h43	GARAGE /	17h56	10.5	
		FJ			18h01		

********** Duree **** T.C. **** L.H. ****** Km ***** Dhoc ********** Salaire *********
* JOUR ----> 11h18 7h13 2h18 187.2 (approx.) $239.51*
****

Figure 7-13 : Exemple d'assignations à la STL pour un chauffeur

Chaque tournée est associée à un seul numéro d'assignation. Une tournée qui aurait un numéro différent des autres tournées identifiées pendant la même période d'assignation est automatiquement rejetée. En cas de conflit, la tournée « gagnante » est celle dont le numéro d'assignation compte le plus grand nombre d'occurrences.

Tableau 7-1 : Établissement des horaires et assignations à partir des circuits parcourus

Heure de départ	Circuit	Direction	Assignations (horaire planifié)				
			450873	450942	450759	~~450775~~	450760
2002-04-02 - 05:25	31	sud	5:22				
2002-04-02 - 06:15	31	nord	6:10				
2002-04-02 - 06:51	31	sud	6:52				
2002-04-02 - 07:45	31	nord	7:40				
2002-04-02 - 08:21	31	sud	8:22				
2002-04-02 - 10:11	50	ouest	10:12				
2002-04-02 - 11:43	50	est	11:44				
2002-04-02 - 13:18	66	ouest		13:17			
2002-04-02 - 14:39	66	est		14:40			
2002-04-02 - 16:42	35	nord			16:35	~~16:55~~	
2002-04-02 - 17:36	35	sud			17:35		
2002-04-02 - 18:26	61	nord			18:20		
2002-04-02 - 19:25	61	sud			19:25		
2002-04-02 - 21:57	61	nord					21:50
2002-04-02 - 22:50	61	sud					22:50
2002-04-02 - 23:57	61	nord					23:50

7.6 Résultats

La méthode explicitée dans ce chapitre a donné naissance à un logiciel expérimental baptisé pour la circonstance « Identificateur de trajets en transport collectif » (ITTC).

ITTC est un ensemble d'algorithmes conçus avec l'interface VBA du logiciel Microsoft Excel. Il fait la lecture de chaque fichier de données brutes provenant des autobus de la STL et effectue l'identification des tournées à l'aide des données sur les circuits d'autobus et les tournées planifiées. Le résultat final est sauvegardé dans un fichier de même format.

7.6.1 Comparaisons

Une première comparaison entre les données d'avril 2002 obtenues avec BusStops et celles obtenues avec ITTC avait mené à la conclusion qu'environ 19 % des circuits identifiés par BusStops sont incorrectement identifiés ou carrément faux (Figure 7-14), et que ITTC permet de détecter 33 % plus de circuits que BusStops (Figure 7-15). Toutefois, cette analyse est basée sur l'hypothèse prétentieuse que les résultats obtenus avec ITTC ne contiennent aucune erreur.

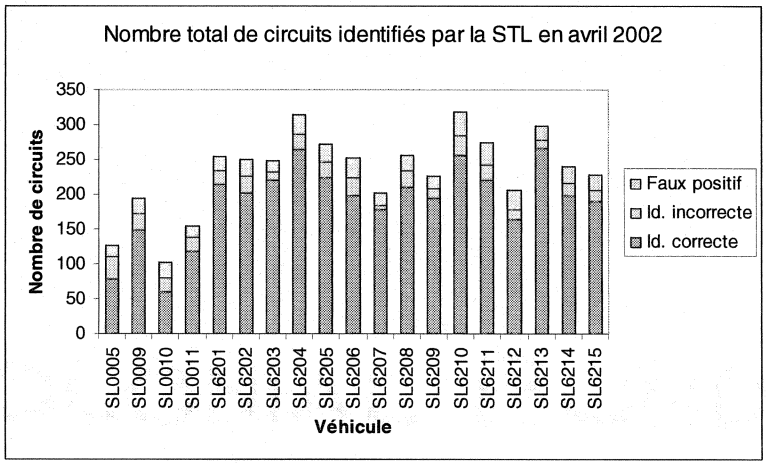


Figure 7-14 : Nombre total de circuits identifiés par la STL en avril 2002 (selon ITTC)

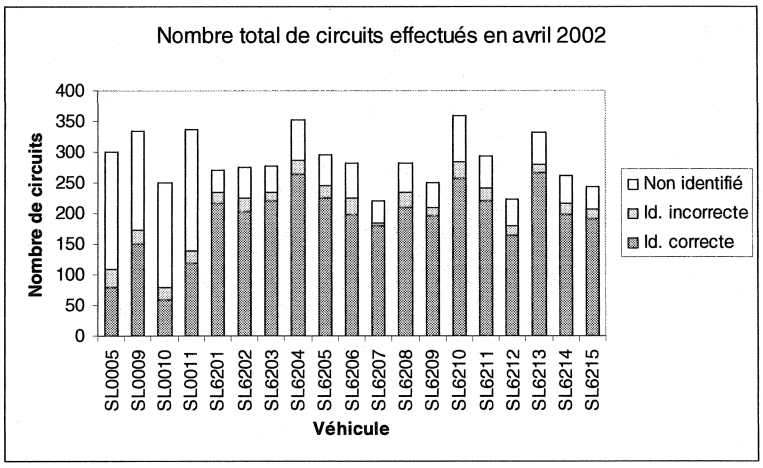


Figure 7-15 : Nombre total de circuits effectués en avril 2002 (selon ITTC)

Pour vérifier l'efficacité réelle du logiciel ITTC – et par la même occasion, celle de BusStops[3] et de LADRA – une expérience fut effectuée sur un échantillon de 527 tournées d'avril 2003 (représentant environ 5 % du nombre total de tournées de cette période) et provenant de 39 fichiers (sur 725) sélectionnés au hasard. Ces tournées furent observées manuellement et comparées avec les tournées identifiées de ITTC, BusStops et LADRA.

Tableau 7-2 : Tournées d'avril 2003 observées manuellement

Observations	BusStops	LADRA	ITTC
N. de tournées identifiées correctement	481	456	516
N. de tournées mal identifiées	8	2	1
N. de fausses tournées (« faux-positifs »)	24	0	0
N. de tournées non-identifiées (« faux-négatifs »)	38	61	10

Comme il a été expliqué précédemment, BusStops et LADRA sont utilisés conjointement dans l'identification des circuits effectués. Cela explique (en partie du moins) le nombre élevé d'erreurs de la part de BusStops et le nombre élevé de circuits non-identifiés (i.e. sans horaire valide) de LADRA.

---

[3] N'oublions pas qu'il est question ici de la version de BusStops utilisée présentement par la STL et non de la nouvelle version présentement utilisée au RTC et au RTL.

Le nombre de tournées « ignorées » par LADRA est élevé (11,6 %) car des 38 tournées non-identifiées de BusStops s'ajoutent 23 tournées identifiées correctement mais sans horaire valide. En tant que filtre, LADRA effectue très bien son travail : toutes les tournées erronées de BusStops ont été rejetées et seulement deux erreurs d'identification (0,4 %) ont été commises.

Le logiciel ITTC donne des résultats beaucoup plus intéressants : il reconnaît 13 % de tournées de plus que LADRA, une seule erreur est commise et seulement 10 tournées (1,9 %) sont ignorées.

Tous les graphiques qui suivent ce chapitre sont conçus avec les données obtenues avec le logiciel expérimental ITTC.

7.6.2 Représentativité des données

En avril 2002, le nombre de tournées qui furent identifiées par BusStops/LADRA pendant les jours ouvrables s'élève à 2441. À partir des données brutes produites pendant cette période, ITTC reconnaît 4038 tournées, soit 65 % de plus. BusStops semble avoir eu de la difficulté à identifier les tournées effectuées par les autobus SL0005, SL0009, SL0010 et SL0011 (Figure 7-15 et Figure 7-16).

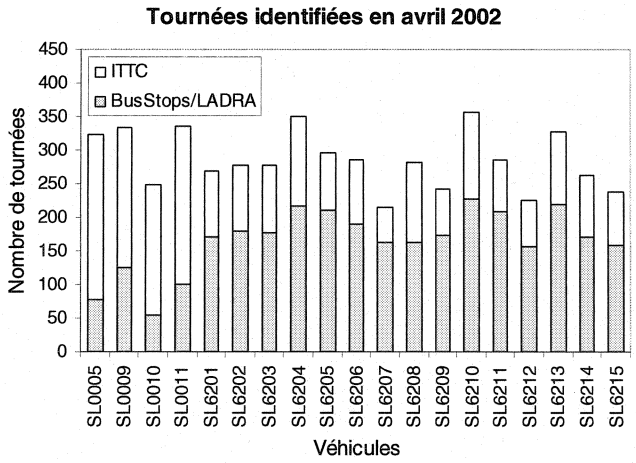


Figure 7-16 : Nombre de tournées identifiés par LADRA et ITTC en avril 2002

Avec BusStops/LADRA, la couverture du réseau pendant les jours ouvrables était de 65 %. Avec ITTC, elle monte à 85 %.

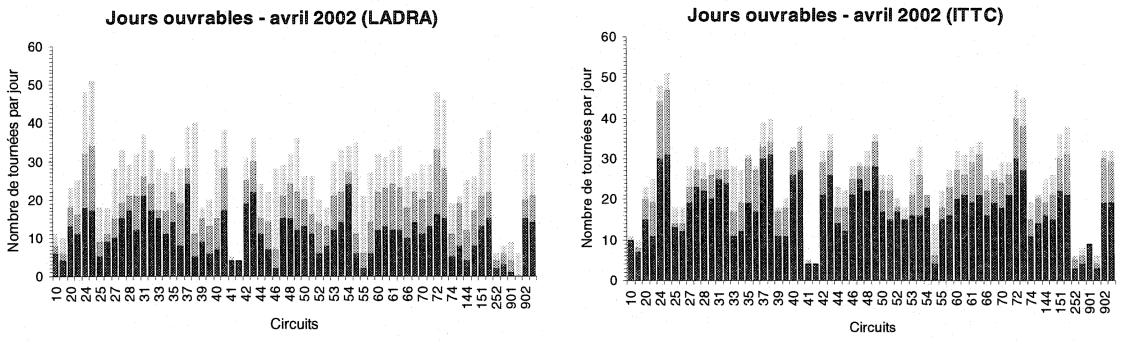


Figure 7-17 : Représentativité des données traitées avec LADRA et ITTC en avril 2002

En avril 2003, avec 23 autobus supplémentaires, BusStops et LADRA ont reconnu 6117 tournées pendant les jours ouvrables. ITTC en a reconnu 6992, soit 14 % de plus (incidemment, on remarque que le couple BusStops/LADRA semble plus productif en avril 2003).

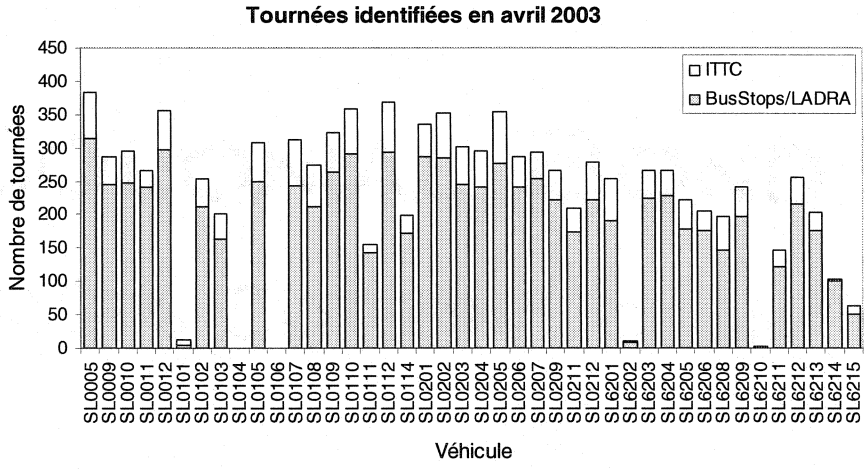


Figure 7-18 : Nombre de tournées identifiés par LADRA et ITTC en avril 2002

Avec BusStops/LADRA, la couverture du réseau pendant les jours ouvrables était de 86 %. Avec ITTC, elle monte à 96 %.

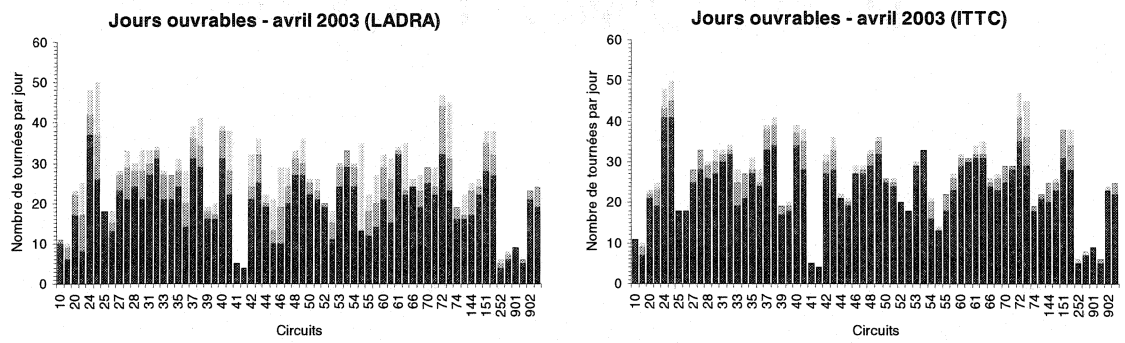


Figure 7-19 : Représentativité des données traitées avec LADRA et ITTC en avril 2003

Il y a cependant quelque chose d'étrange dans ces chiffres : comment se fait-il qu'une augmentation d'à peine 14 % du nombre de tournées identifiées se traduise par une amélioration de 10 % de la couverture du réseau? L'hypothèse la plus plausible est que BusStops n'arrive pas à détecter toutes les tournées de façon égale (ITTC non plus d'ailleurs) et que certaines tournées particulières ne sont pas du tout détectées. En particulier, BusStops semble avoir de la difficulté à reconnaître les *variantes* de certains circuits. Comme ITTC arrive à identifier des tournées que BusStops ne détecte pas, la couverture du réseau est améliorée en conséquence.

On doit s'assurer que l'algorithme utilisé pour la reconnaissance des tournées (BusStops, ITTC ou autre) soit capable de reconnaître *toutes* les tournées au moins une fois pendant la période de temps choisie. Sinon, une couverture de 100 % serait impossible même avec une flotte entière d'autobus équipée d'appareils GPS.

7.7 Discussions

7.7.1 Améliorations possibles

Contrairement à ce que laisse sous-entendre la section 7.5.3, les balises virtuelles ne furent pas utilisées pour confirmer ou infirmer l'identification d'un circuit. Cela a comme conséquence qu'un voyage à vide pourrait éventuellement être considéré comme une tournée possédant les mêmes points de départ et d'arrivée et la même heure de départ, puisqu'il n'y a souvent aucune donnée intermédiaire (aucun *événement*) permettant de confirmer que le véhicule s'est effectivement éloigné du parcours identifié. Ceci dit, aucune erreur de ce genre n'apparaît dans le Tableau 7-2 mais cela ne veut pas dire qu'il n'y en a pas dans l'ensemble des données.

En outre, ITTC ne se sert que des heures de départ pour identifier les tournées. Il semble que cette information soit suffisante pour identifier les tournées dans la plupart des cas, mais cela peut soulever des doutes à certaines occasions.

L'utilisation d'autres informations pourrait améliorer le logiciel et le rendre plus robuste; par exemple l'exploitation des heures de passage aux points de contrôle et des distances théoriques entre les points de contrôle. Sans compter que les heures de passage aux points de contrôle permettent de mesurer directement la ponctualité des véhicules tandis que les distances entre les points de contrôle pourraient éventuellement servir à calibrer les odomètres.

7.7.2 Objectif (rappel)

À la base, l'objectif derrière la conception d'ITTC est de souligner qu'une partie des données collectées par les véhicules n'est pas utilisée à cause de problèmes au niveau de la fusion avec les données planifiées, et qu'il est possible d'obtenir un résultat plus satisfaisant en appliquant quelques règles relativement *simples*.

D'ailleurs, une nouvelle version de BusStops existe déjà et exploite à la fois les informations sur les circuits, les horaires et les opérations journalières (Figure 3-7). Le RTC travaille en collaboration avec Infodev afin de rendre le logiciel le plus efficace et robuste possible. Il serait intéressant de voir comment la nouvelle version de BusStops se comparerait à ITTC...

CHAPITRE 8 : Détection de problèmes de comptage

> *By their nature, quantitative descriptions are so one-dimensional and featureless that they cannot help but conceal the structures that give rise to them. This is inescapable, since any act that makes two different things comparable must do it by deflecting our attention from their differences. Numbers themselves are the greatest masters of disguise because they perfectly conceal all traces of their origins.*
>
> – Marvin Minsky, *The Society of Mind* (1986)

Bien que l'achalandage d'un système de transport fluctue avec une certaine prévisibilité surtout en période de pointe, sa composante stochastique rend les erreurs difficiles à éliminer et facile à confondre avec une variation brutale – mais parfois réelle – de l'achalandage.

Les risques de mauvaise interprétation des données sont nombreux à cause des erreurs causées par les systèmes de collecte de données. On parvient à éliminer certaines de ces erreurs grâce à quelques techniques qui sont abordées ici, notamment l'élimination des charges résiduelles trop grandes et l'utilisation de ce qu'on appellera ici *indices de fiabilité*.

En outre, une méthode permettant d'avoir une vue d'ensemble des données collectées par les ordinateurs embarqués est présentée à la fin de ce chapitre.

8.1 Charges résiduelles

Comme il a été mentionné à la section 4.4, des erreurs s'introduisent dans le comptage des passagers et ces erreurs se traduisent souvent par une différence entre le nombre de montées et le nombre de descentes sur une ligne de transport, et donc une charge résiduelle en fin de parcours.

Dans les cas où le nombre de montées et le nombre de descentes ne s'annulent pas, on s'intéresse à l'ampleur de la charge résiduelle. Si ce déséquilibre est jugé trop important, alors on rejette les données concernées.

Dans le cas où la charge résiduelle est faible sans être nulle, on applique une correction afin de la rendre nulle pour que le profil de charge soit plausible. Mais où cette correction doit-elle être appliquée? On fait l'hypothèse que les erreurs sont réparties à peu près également à mesure que le comptage est effectué et que l'écart entre le nombre de personnes mesuré et réel à un arrêt est proportionnel au nombre de personnes mesuré. On répartit alors l'erreur sur le profil de charge en conséquence tout en respectant quelques règles simples, semblables à celles décrites sur le site web de la compagnie Iris (IRIS, 2004) :

- Le nombre de personnes à bord du véhicule est égal à 0 (chauffeur non inclus) ou 1 (chauffeur inclus) au début et à la fin d'un parcours.
- Le nombre de personnes est toujours positif ou nul (chauffeur non inclus) pendant l'exécution du parcours.
- Le nombre de descendants est nul au terminus de début de parcours et le nombre de montants est nul au terminus de fin de parcours.
- Le nombre de montants et de descendants est toujours entier positif.
- Pour chaque arrêt, lorsque le nombre de montants ou de descendants est très petit (1 ou 2), on suppose qu'il n'y a pas d'erreur de comptage.

Odo. (m)	Montants	Descendants	Charge
0	1	0	1
157	11	0	12
2615	1	0	13
3311	1	0	14
4390	3	2	15
4811	0	1	14
6246	0	1	13
6443	0	1	12
6764	0	4	8
7721	0	1	7
8259	2	1	8
8853	1	2	7
9128	0	1	6
10346	3	0	9
12941	0	1	8
13754	0	3	5
14918	0	1	4
15421	0	1	3
16268	0	0	3

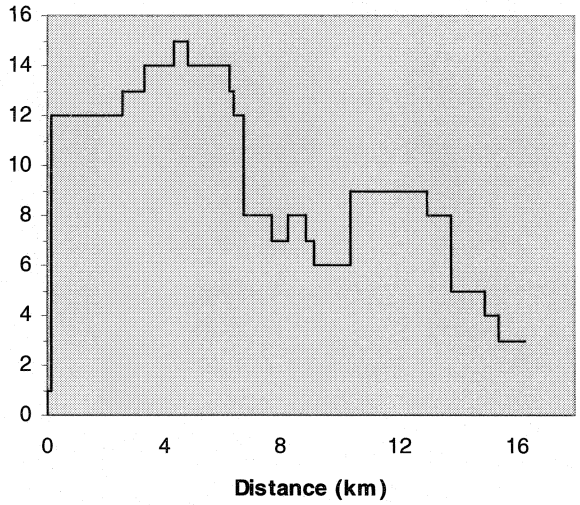


Figure 8-1 : Exemple de profil de charge non-équilibré

La Figure 8-1 montre un profil de charge contenant une charge résiduelle faible. Pour éliminer cette charge résiduelle, on fait l'addition de tous les montants et descendants qui peuvent être modifiés en vertu des règles citées ci-haut. On exclut donc les valeurs égales ou inférieures à deux.

Odo. (m)	Montants ajustables	Desc. ajustables
0		
157	9	
2615		
3311		
4390	1	
4811		
6246		
6443		
6764		2
7721		
8259		
8853		
9128		
10346	1	
12941		
13754		1
14918		
15421		
16268		
Total ajustable	14	
Facteur d'ajustement	3 / 14 = 0,214	

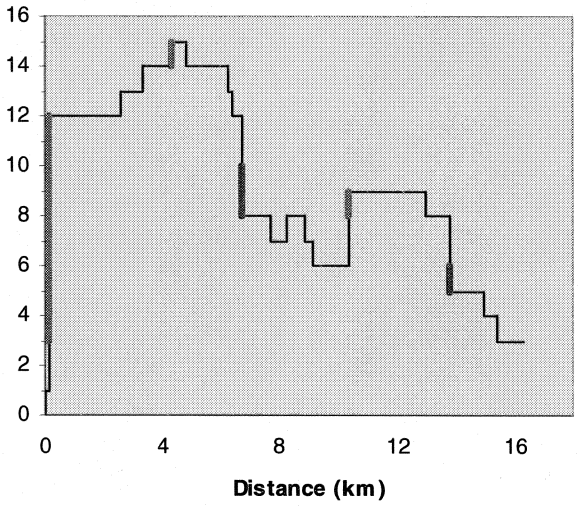


Figure 8-2 : Exemple de correction de profil de charge

En multipliant les valeurs ajustables par le facteur d'ajustement, on obtient les valeurs à retrancher (bleu) ou à ajouter (rouge). Le résultat final est le suivant :

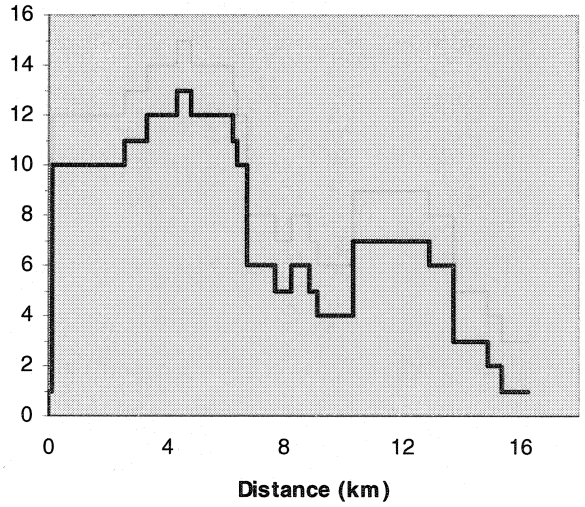


Figure 8-3 : Exemple de profil de charge corrigé

La charge résiduelle est ici égale à un. La raison pour laquelle elle n'a pas été entièrement éliminée est que l'on traite ici avec des valeurs entières. Le calcul pourrait être amélioré en transférant la partie décimale à chaque calcul.

8.1.1 Boucles

Une attention particulière doit être portée aux lignes de transport finissant ou commençant dans une boucle (exemples : circuits 33, 37 et 54 de la STL) où les véhicules peuvent se retrouver avec une charge « résiduelle » au « terminus ». Une façon de contourner le problème est de traiter l'aller et le retour comme une seule ligne de transport aller-retour. Le problème devient plus complexe si une boucle se trouve à chaque extrémité ou si la ligne constitue elle-même une boucle. Une autre solution est de traiter chaque parcours séparément mais en considérant les montées à *x* arrêts avant le terminus de départ ou les descentes à *x* arrêts après le terminus de fin, ou les deux.

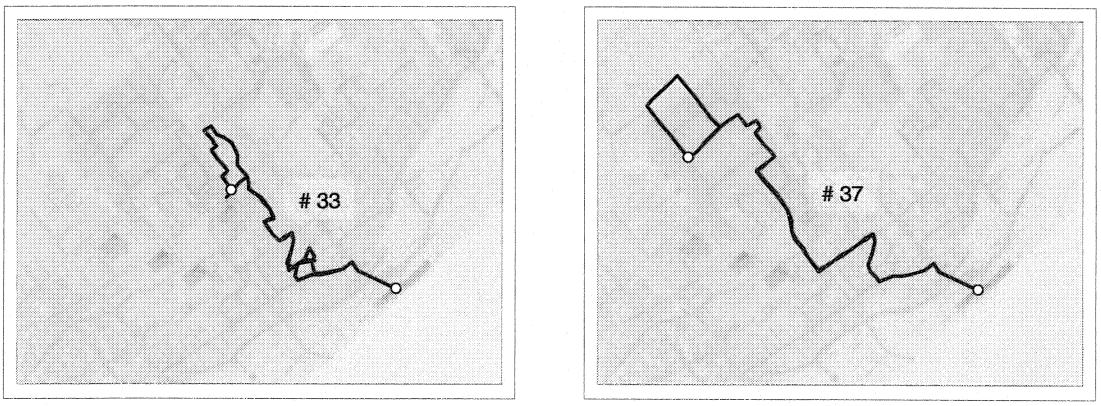


Figure 8-4 : Circuits aller-retour contenant une boucle à une extrémité

8.2 Indices de fiabilité

On se sert de la charge résiduelle pour éliminer les données douteuses, mais cela ne signifie pas que les données conservées sont nécessairement correctes.

Les expériences 2 et 3 du chapitre 5 montrent en effet que même avec un nombre presque égal de montées et de descentes, les erreurs peuvent être significatives.

Un cas extrême est le système d'acquisition de données dont les capteurs sont peu sensibles et ne détectent presque rien, donnant des résultats semblables à celui du graphique central de la Figure 8-5.

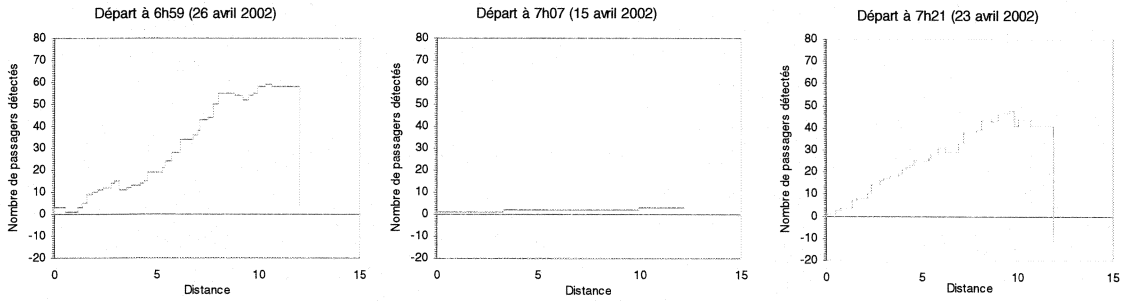


Figure 8-5 : Profils de charge mesurés de la ligne 24 sud à 6h59, 7h07 et 7h21

Selon les critères établis précédemment, le profil du centre est considéré acceptable puisqu'il possède une charge résiduelle très faible (alors que le profil de droite doit être éliminé avec $C_R = -11$). Pourtant, ce profil de charge semble bien douteux. Il serait intéressant de vérifier les autres données provenant du même autobus qui a produit cette courbe.

Les graphiques suivants représentent l'achalandage à chaque ouverture de porte tel que mesuré par deux autobus pendant une journée complète. La Figure 8-6 est le profil de charge d'un autobus dont le système de comptage fonctionne à peu près correctement; on distingue d'ailleurs facilement chacun des circuits effectués par cet autobus. On observe une dérive vers la fin de la journée et la charge est négative à la fin, mais cette erreur disparaît lorsqu'on applique une correction sur chaque circuit.

La Figure 8-7 quant à elle est manifestement problématique : des ouvertures de portes sans montée ni descente, quelques variations isolées, une charge résiduelle considérable, tantôt négative, tantôt positive.

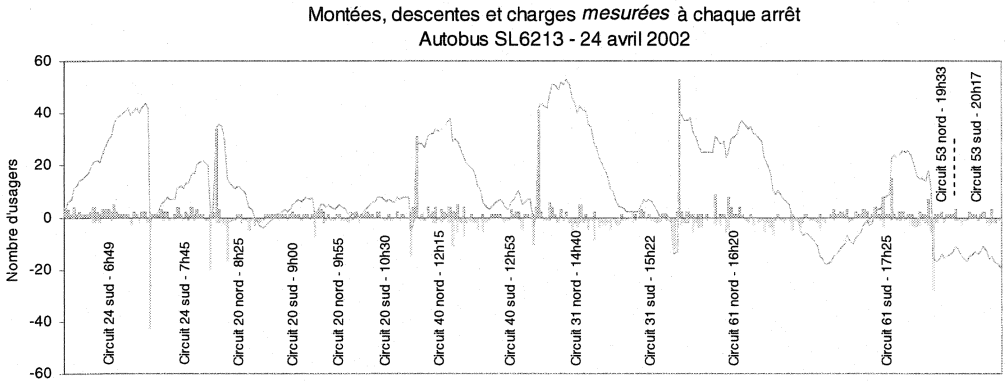


Figure 8-6 : Achalandage mesuré à bord de l'autobus SL6213 pendant un jour

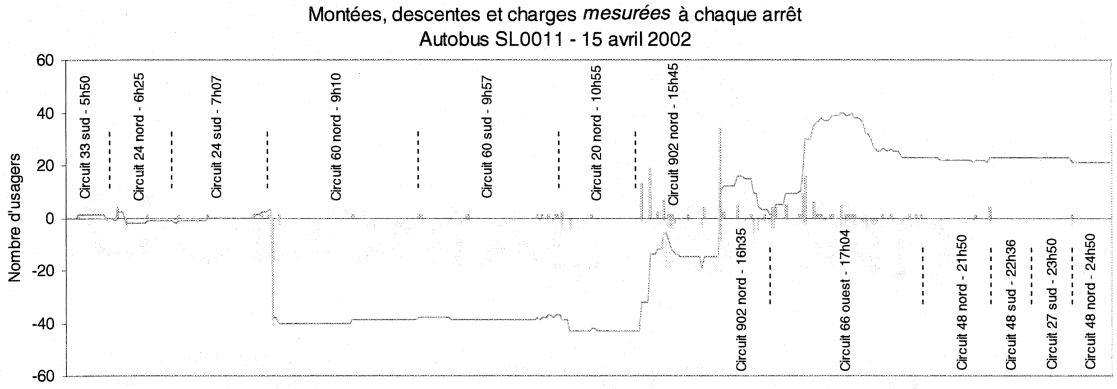


Figure 8-7 : Achalandage mesuré à bord de l'autobus SL0011 pendant un jour

Nul doute qu'un tel résultat de la part de l'autobus SL0011 indique une défectuosité du système et que toute donnée provenant de cet autobus doit être écartée, du moins pour cette journée-là.

Du point de vue analytique (Tableau 8-1 et Tableau 8-2), on observe un nombre de montées et de descentes plus bas à bord de l'autobus SL0011; les valeurs

erratiques des capteurs sont également mises en évidence dans le calcul de la moyenne et de l'écart-type des charges résiduelles.

Tableau 8-1 : Achalandage mesuré à bord de l'autobus SL6213 le 24 avril 2002

	Heure	Ligne	Montées	Descentes	Différence	EM (%)
SL6213 – 24 avril 2002	6h49	24 sud	55	56	1	1
	7h45	24 sud	29	30	1	2
	8h25	20 nord	42	44	2	2
	9h00	20 sud	16	12	4	14
	9h55	20 nord	12	12	0	0
	10h30	20 sud	12	20	8	25
	12h15	40 nord	74	63	11	8
	12h53	40 sud	14	21	7	20
	14h40	31 nord	73	67	6	4
	15h22	31 sud	11	28	17	44
	16h20	61 nord	98	102	4	2
	17h25	61 sud	72	71	1	1
	19h33	53 nord	9	9	0	0
	20h17	53 sud	10	12	2	9
	Somme		527	547	64	6
	Moyenne				4,6	
	Moyenne + écart-type				9,5	

Tableau 8-2 : Achalandage mesuré à bord de l'autobus SL0012 le 15 avril 2002

	Heure	Ligne	Montées	Descentes	Différence	EM (%)
SL0012 – 15 avril 2002	5h50	33 sud	1	7	6	75
	6h25	24 nord	5	1	4	67
	7h07	24 sud	3	44	41	87
	9h10	60 nord	5	1	4	67
	9h57	60 sud	4	4	0	0
	10h55	20 nord	5	9	4	29
	15h45	902 nord	44	17	27	44
	16h35	902 nord	43	22	21	32
	17h04	66 ouest	64	38	26	25
	21h50	48 nord	1	4	3	60
	22h36	48 sud	4	0	4	100
	23h50	27 sud	0	3	3	100
	24h50	48 nord	1	0	1	100
	Somme		180	150	144	87
	Moyenne				11,1	
	Moyenne + écart-type				24,2	

Un autobus qui n'est en opération que pendant les périodes de pointe n'aura pas le même achalandage que celui qui fonctionne en période hors-pointe. En cas

d'incertitude, on doit se fier à des variables plus indépendantes, par exemple le rapport *Différence absolue* : *Nombre de montées & descentes* que l'on appelle aussi *écart moyen* (IRIS, 2004):

$$EM = \frac{\sum_{k=1}^{v} |M'_k - D'_k|}{M' + D'}$$

avec v le nombre de lignes (voyages) et M'_k et D'_k respectivement le nombre total de montées et de descentes détectées pour chaque parcours k. L'écart moyen permet de mesurer l'équilibre des capteurs directionnels. Un capteur *déséquilibré* favorisera une direction aux dépens d'une autre (ex. plusieurs montées seront comptées comme des descentes ou vice versa). Le Tableau 8-2 indique un écart de 87 % pour l'ensemble des circuits effectués le 15 avril par l'autobus SL0011, ce qui est énorme et discrédite encore une fois les données obtenues.

Par contre, une valeur aussi élevée a été obtenue parce que les capteurs ont fonctionné de façon intermittente. S'ils avaient été complètement amorphes, nous aurions obtenu une déviation presque nulle et nous n'aurions pas conclu à un mauvais fonctionnement.

Dans les deux cas précédents, les autobus effectuent des assignations différentes mais comparables. À preuve, le nombre d'ouvertures de portes (événement P) est semblable : 346 contre 380. Cela nous permet de comparer les sommes des montées et des descentes comme ça a été fait précédemment. Mais ça aurait pu être différent; les deux exemples choisis auraient pu illustrer des réalités différentes – l'un des deux autobus aurait pu n'être utilisé qu'en pointe ou hors-pointe seulement ou seulement pendant une partie de la journée.

Par conséquent, une autre variable qui devient intéressante ici est le nombre moyen de montées et descentes à chaque ouverture de portes :

$$\overline{(M'+D')} = \frac{M'+D'}{n}$$ avec n = nombre d'ouvertures de portes

Tableau 8-3 : Nombre moyen de montées et de descentes par ouverture de portes

Autobus – jour	$M'+D'$	n	$\overline{(M'+D')}$
SL6213 – 24 avril 2002	1074	346	3,10
SL0012 – 15 avril 2002	330	380	0,87

En principe, on s'attend à ce qu'au moins une personne monte ou descende de l'autobus. Avec l'autobus SL0011, cette valeur est inférieure à 1, ce qui démontre qu'il y a un problème.

Les figures suivantes montrent les différents indices de fiabilité calculés sur un mois pour les autobus munis de systèmes de collecte de données.

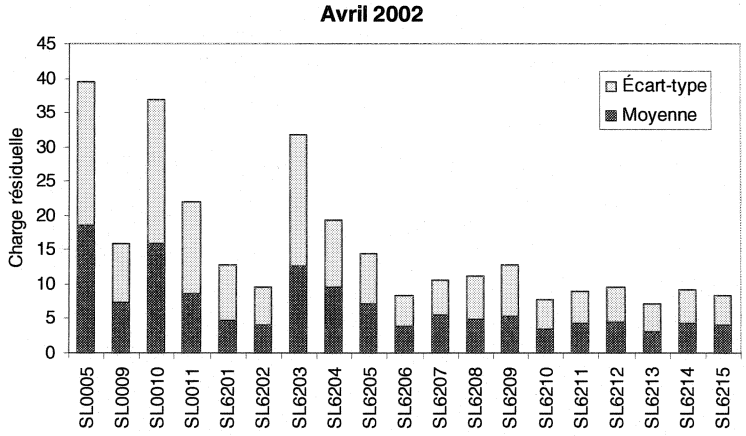


Figure 8-8 : Moyenne et écart-type des charges résiduelles des 19 autobus en avril 2002

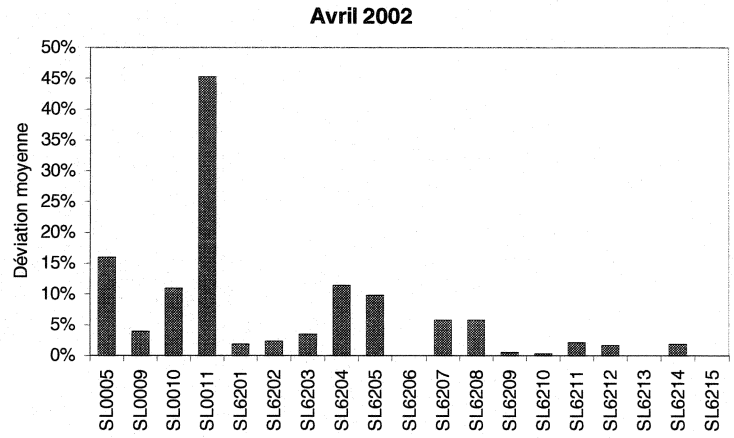


Figure 8-9 : Écart moyen des 19 autobus en avril 2002

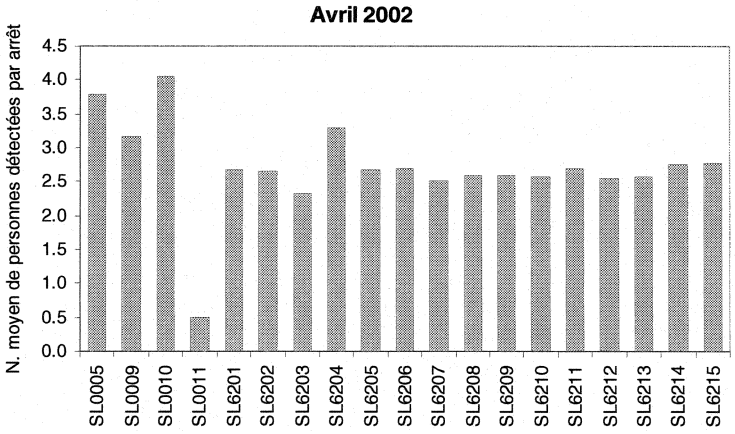


Figure 8-10 : Achalandage moyen détecté à chaque arrêt par les 19 autobus en avril 2002

Tous ces graphiques mettent en évidence le mauvais fonctionnement du système de comptage de l'autobus SL0011. Mais ils attirent aussi notre attention sur les autobus SL0005, SL0010, SL6203 et SL6204. En particulier, l'autobus SL0005 présente des valeurs élevées dans les trois cas, ce qui semble indiquer une trop grande sensibilité des capteurs. Cette hypothèse est prouvée lorsqu'on examine visuellement l'évolution de la charge mesurée dans cet autobus.

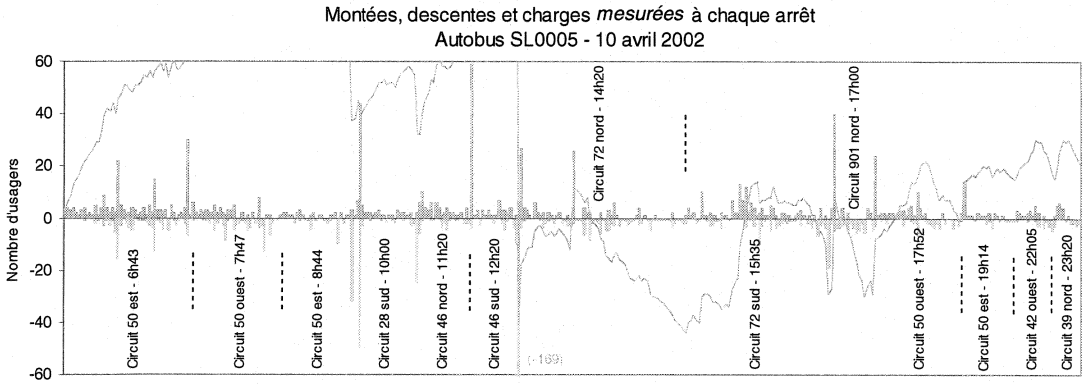


Figure 8-11 : Achalandage mesuré à bord de l'autobus SL0005 pendant un jour

La Figure 8-11 montre l'achalandage mesuré à bord de l'autobus SL0005 pendant la journée du 10 avril 2002. La plupart des données prises circuit par circuit seraient rejetées de toute façon à cause du déséquilibre entre les montants et les descendants (contrairement aux données de l'autobus SL0011) mais les données de circuits dont les montants et les descendants sont équilibrés (charge résiduelle faible) n'en sont pas moins suspectes.

8.2.1 Données d'avril 2003

En avril 2003, quarante-deux autobus étaient équipés de systèmes de comptage et de GPS. La performance de 39 d'entre eux est illustrée ci-dessous (les données des autres autobus sont rejetées à cause de problèmes dus aux systèmes embarqués – voir section 4.5).

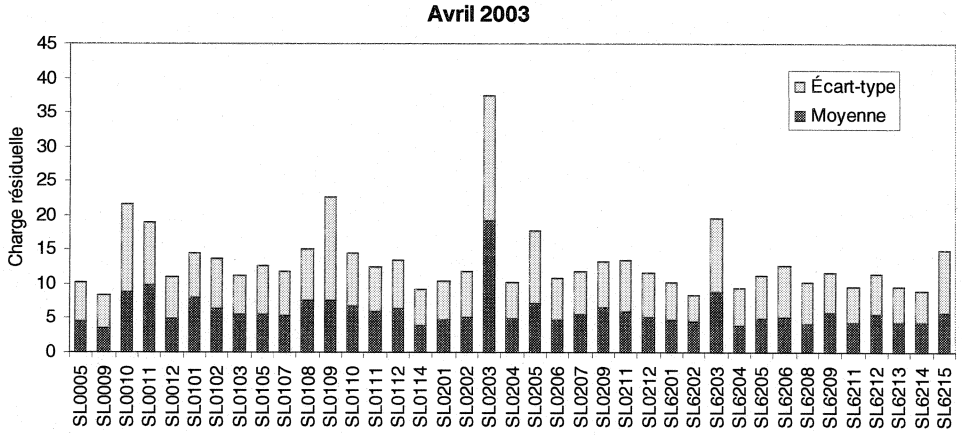


Figure 8-12 : Moyenne et écart-type des charges résiduelles des 39 autobus en avril 2003

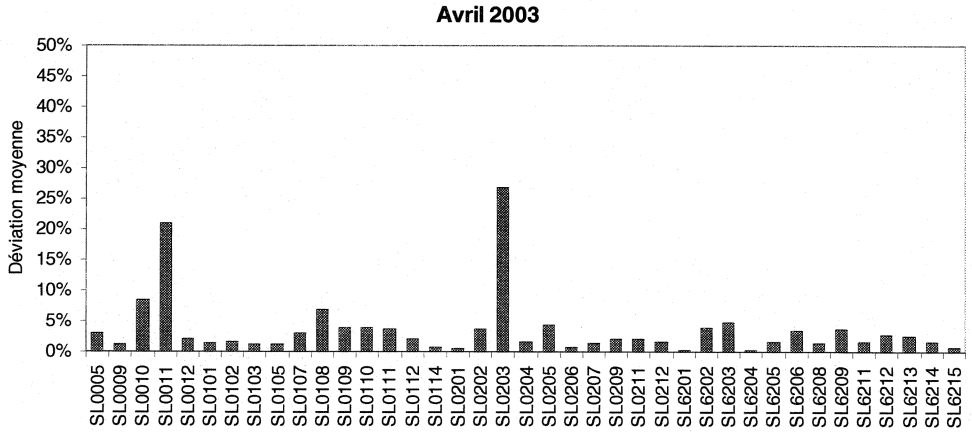


Figure 8-13 : Écart moyen des 39 autobus en avril 2003

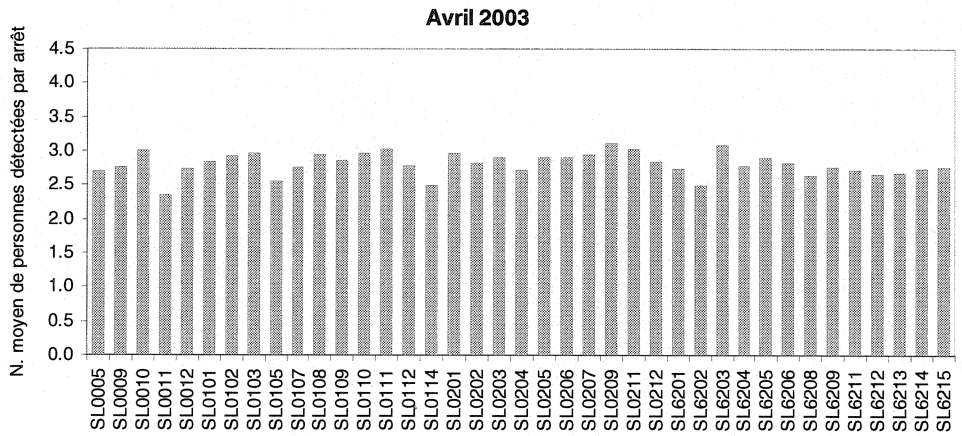


Figure 8-14 : Achalandage moyen détecté à chaque arrêt par les 39 autobus en avril 2003

On remarque sans peine dans la Figure 8-12 et la Figure 8-13 l'autobus SL0203 qui semble avoir un problème de capteurs ainsi que, dans une moindre mesure, les autobus SL0010 et SL0011. Un analyste devra examiner sérieusement les systèmes de comptage de ces véhicules et effectuer les ajustements ou réparations nécessaires.

Dans l'ensemble, les systèmes de comptage semblent fonctionner mieux que ceux de l'année précédente.

Il est à noter que les calculs et les graphiques présentés ici ne montrent que quelques aspects de la collecte de données. Un analyste pourra développer d'autres outils qui lui permettront de régler d'autres problèmes ou de mieux régler ceux décrits dans ce chapitre.

8.3 Vue d'ensemble

À de nombreuses reprises, les données collectées par les ordinateurs embarqués sont regroupées et classées afin d'illustrer certaines problématiques, que ce soit la distance parcourue par les véhicules (chapitre 4), le nombre de circuits ou tournées détectés par véhicule (chapitres 4 et 7), ou le nombre de véhicules utilisés par jour (chapitre 4).

Tous ces graphiques ne présentent qu'un élément à la fois. Le graphique qui suit montre tous les voyages effectués par tous les véhicules pendant tout le mois d'avril 2003. Un code de couleur permet de déterminer la nature du voyage ainsi que la validité des données collectées.

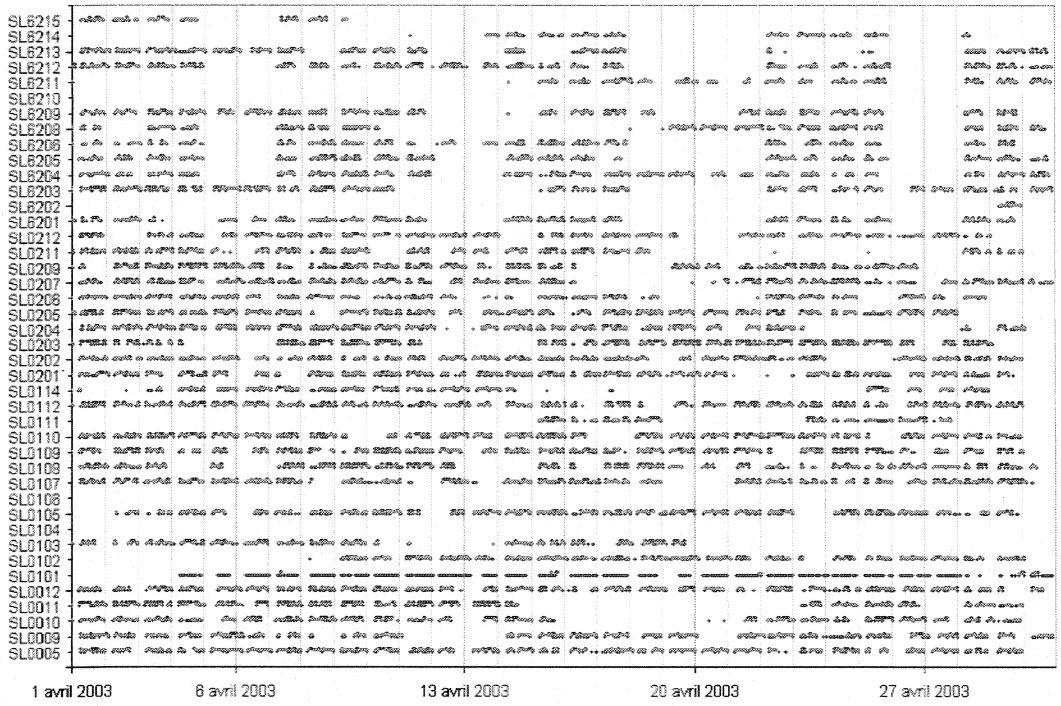


Figure 8-15 : Vue d'ensemble des déplacements effectués pendant avril 2003

La Figure 8-15 montre en vert les voyages ayant une charge résiduelle faible ($|C_R| \leq 5$), en rouge ceux ayant une charge résiduelle élevée ($|C_R| > 5$) et en bleu les voyages qui ne sont pas identifiés comme circuits.

On y voit entre autres les périodes d'utilisation de chaque véhicule et, par exemple, ceux qui ne sont pas utilisés les fins de semaine (un fin observateur remarquera que ces véhicules sont pratiquement tous des autobus Classic). On y voit aussi par la couleur ceux dont le système de comptage fournit des résultats questionnables (SL0010, SL0011 et SL0203). On remarque finalement l'autobus SL0101 dont aucun des voyages n'est identifié – à cause de son GPS défectueux.

Un analyste pourra toujours changer l'échelle horizontale afin de visionner les données sur une base quotidienne ou hebdomadaire (Figure 8-16, Figure 8-17).

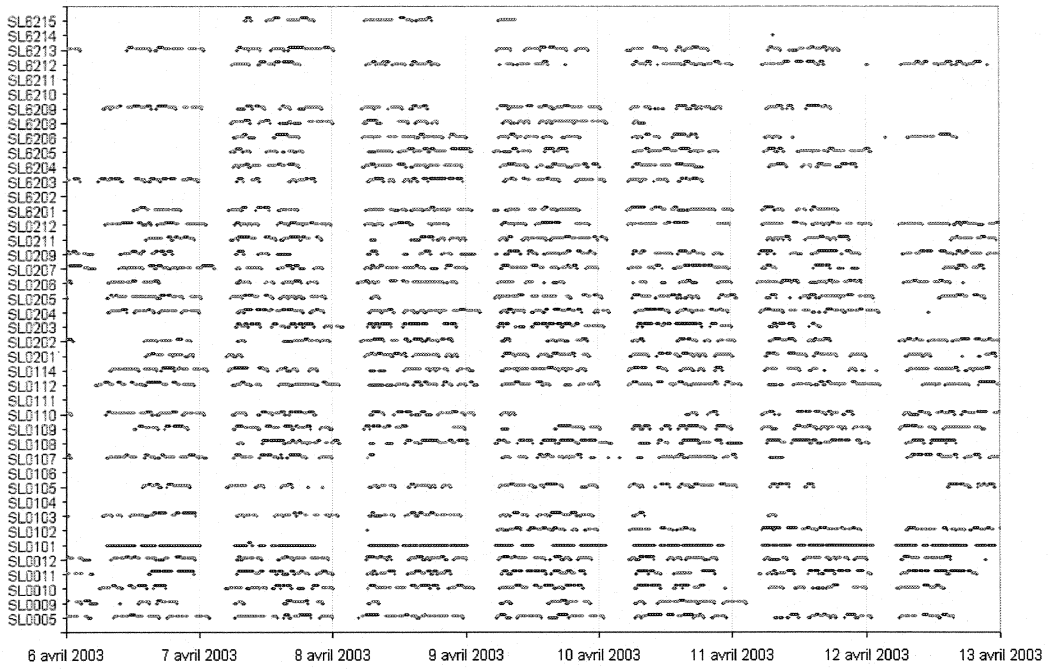


Figure 8-16 : Vue d'ensemble des déplacements effectués du 6 au 12 avril 2003

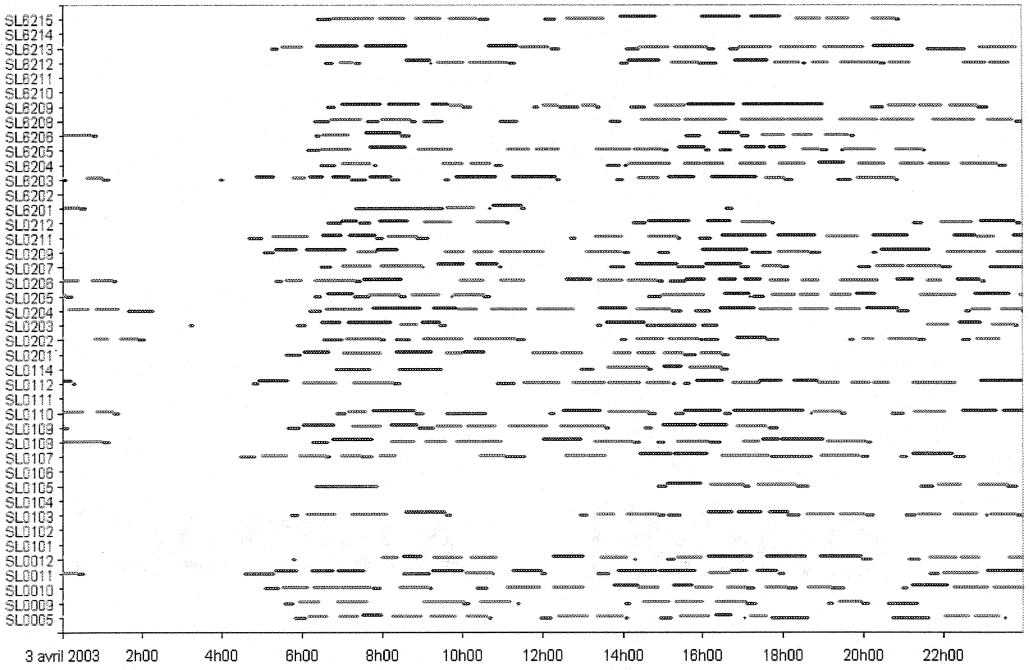


Figure 8-17 : Vue d'ensemble des déplacements effectués le 3 avril 2003

CHAPITRE 9 : Modélisation saisonnalisée de la demande

L'analyse d'une base de données issue d'un système de comptage de passagers et de localisation géographique devrait, idéalement, nous permettre de quantifier précisément l'achalandage sur l'ensemble du réseau. Toutefois, avant d'y parvenir, plusieurs difficultés doivent être surmontées.

Agrégation. Pour des raisons d'économie et aussi par prudence, les sociétés de transport hésitent à équiper la totalité de leur flotte d'autobus de systèmes de collecte de données. Dans des cas comme celui de la STL, qui, depuis 2003, dispose de 42 systèmes sur un total d'un peu plus de 220 autobus (19 %), seulement une petite partie du réseau sera couverte à chaque jour. Pour obtenir une couverture plus grande, on doit rassembler des données collectées sur plusieurs jours; c'est l'agrégation temporelle.

Validation. L'imperfection des systèmes embarqués fait en sorte que des erreurs se produisent fréquemment durant le processus de collecte des données. Dans certains cas, les erreurs introduites sont mineures et ne nécessitent qu'une simple correction. Dans d'autres cas, les erreurs sont importantes et les données concernées doivent être rejetées.

Synthèse. Sur une période donnée, certaines tournées peuvent être couvertes plusieurs fois. Lorsqu'on effectue une agrégation temporelle, on doit alors faire un choix sur l'interprétation de ces données. Bien que le calcul d'une valeur moyenne soit une solution naturelle dans la plupart des cas, d'autres solutions pourraient s'avérer plus adéquates.

Interpolation. Dans le cas où certaines tournées ne sont pas couvertes sur la période agrégée – ou qu'elles ont été rejetées –, on doit compléter en procédant à une estimation.

9.1 Agrégation temporelle

Comme il en a été discuté à la section 4.3, une agrégation temporelle permet de réunir des informations qui ont été collectées pendant une certaine période de temps. Le principal avantage de cette méthode est qu'elle permet d'obtenir des données fiables en éliminant les variations aléatoires et les erreurs tout en travaillant avec des ressources limitées. Son principal défaut est que les variations de l'achalandage qui surviennent à l'intérieur de la période d'agrégation sont effacées.

Pour obtenir une couverture complète et fiable du réseau, une agrégation temporelle devra s'étendre sur plusieurs semaines, voire plusieurs mois. Néanmoins, on ne doit pas négliger les variations saisonnières, c'est pourquoi en général la période d'agrégation ne dépassera pas trois mois (ce qui a l'avantage de correspondre à la période normale d'habillage des horaires).

On reconnaît actuellement quatre types de jours : ouvrable, samedi, dimanche et férié (les dates dans le Tableau 9-1 sont tous des jours ouvrables). Toutefois, on ne doit pas écarter la possibilité de classer les jours autrement, surtout si l'on dispose de suffisamment de données pour le faire.

Tableau 9-1 : Exemple de liste des échantillons classés par tournées

Circuit (var)	Départ	Date	No bus
10 0 (1)	15:50	2002-04-03	SL0012
10 0 (1)	15:50	2002-04-22	SL6212
10 0 (1)	16:19	2002-04-03	SL0010
10 0 (1)	16:19	2002-04-05	SL0010
10 0 (1)	16:34	2002-04-03	SL6210
10 0 (1)	16:34	2002-04-15	SL6215
10 0 (1)	16:50	2002-04-03	SL0012
10 0 (1)	16:50	2002-04-22	SL6212
10 0 (1)	16:58	2002-04-09	SL6206
10 0 (1)	16:58	2002-04-11	SL6203
10 0 (1)	16:58	2002-04-16	SL6208
10 0 (1)	16:58	2002-04-26	SL6211
10 0 (1)	17:20	2002-04-03	SL0010
10 0 (1)	17:20	2002-04-05	SL0010
10 0 (1)	17:30	2002-04-10	SL0009
10 0 (1)	17:30	2002-04-19	SL6215
10 0 (1)	17:30	2002-04-22	SL6204
10 0 (1)	17:30	2002-04-24	SL6212
10 0 (1)	17:35	2002-04-03	SL6210

(...)

(...)

Par exemple, de l'avis de certains chauffeurs d'autobus, les lundis et les vendredis ont des achalandages différents des mardis, mercredis et jeudi; l'hypothèse émise à ce sujet est qu'une proportion non négligeable de la clientèle travaille quatre jours par semaine au lieu de cinq. Aussi, certaines sociétés de transports reconnaissent une augmentation de leur clientèle les jeudis et vendredis soirs et offrent un service accru en conséquence.

Bien sûr, dans un contexte où les données sont très abondantes, on pourrait éventuellement se permettre le luxe d'agréger sur chaque jour de la semaine au lieu de les regrouper en « jours ouvrables ».

9.2 Validation des données

La validation des données consiste entre autres à conserver les tournées dont le profil de charge se termine par une charge résiduelle faible ou nulle et à exclure les autres données.

De plus, on doit s'assurer que les données agrégées contiennent le moins possible d'erreurs causées par des systèmes défectueux. Parmi les défectuosités possibles, notons :

- Système GPS en panne (aucune donnée de localisation);
- Capteurs de mouvement amorphes, trop sensibles ou déséquilibrés (favorisant une direction aux dépens de l'autre);
- Absence de données secondaires (odomètre, compas, certains types d'événements);
- Horloge mal ajustée ou défectueuse.

Les données produites par un véhicule dont le GPS est en panne ne permettent pas la reconnaissance des tournées et elles seront automatiquement rejetées.

Les problèmes de comptage sont moins évidents à détecter et peuvent être confondus avec des fluctuations irrégulières de l'achalandage. Néanmoins, certains indices permettent de cerner un système de comptage défectueux. Le chapitre précédent est entièrement consacré à cette problématique.

L'absence de certaines données secondaires peut avoir des conséquences importantes. Par exemple, l'absence de données d'odomètre n'empêche pas la reconnaissance des tournées ou l'analyse de l'achalandage sur le territoire mais rend difficile certaines analyses basées sur la distance parcourue, tels les profils de charge et les diagrammes espace-temps.

Les données associées à de mauvaises dates (ex. 1er décembre 2022) sont faciles à éliminer, mais celles associées à de mauvaises *heures* sont carrément vicieuses puisqu'il n'y a pratiquement aucun moyen de savoir si l'heure est juste ou non. Seuls certains indices mettent la puce à l'oreille : séquences de tournées ne correspondant à aucune assignation, tournées impossibles (ex. en pleine nuit).

Aux problèmes techniques s'ajoutent les problèmes logiciels, soient les erreurs de découpage et la mauvaise identification des tournées. Cette problématique est abordée au chapitre 7.

9.3 Synthèse des données

Les graphiques montrant la représentativité des données, comme ceux de la Figure 4-5, identifient en bleu foncé la proportion des tournées qui sont identifiées plus d'une fois. La figure ci-dessous illustre l'une de ces tournées.

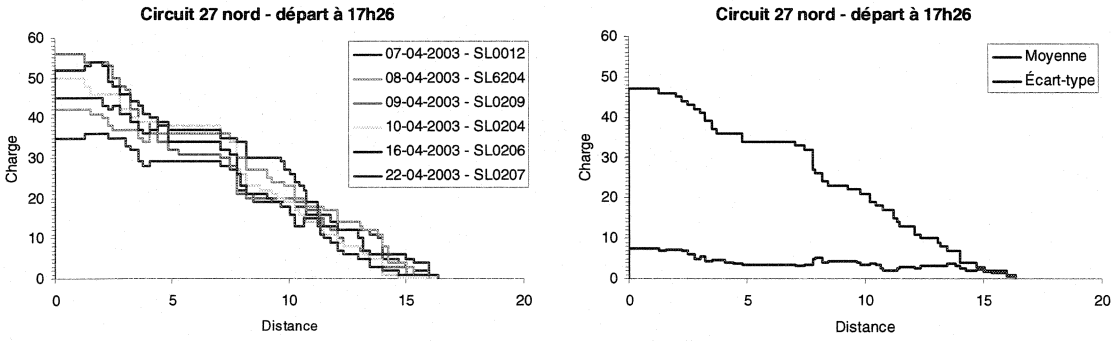


Figure 9-1 : Profils de charge du circuit 27 nord à 17h26

L'intégration de ces données en une seule courbe peut se faire en calculant la moyenne des charges le long du parcours. Dans ce cas-ci, cette technique est tout à fait naturelle car les courbes sont regroupées en peloton et aucune d'entre elles ne se détache du groupe de façon notable.

Dans le cas contraire, il est utile de chercher à connaître les raisons pour lesquelles les courbes diffèrent.

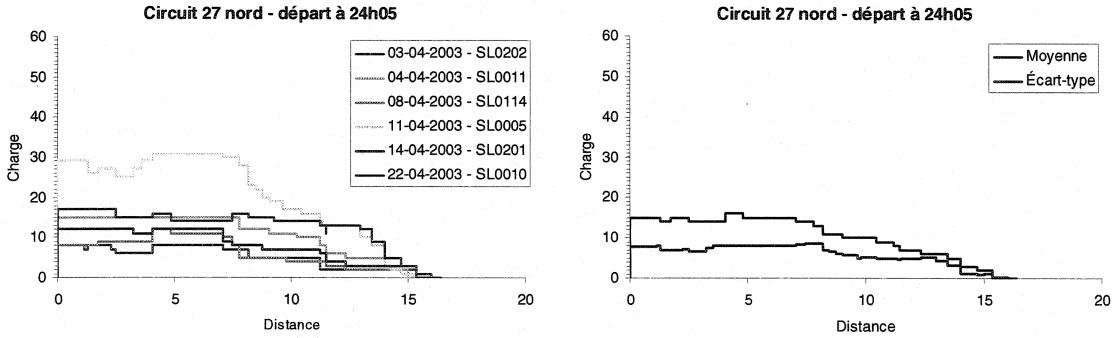


Figure 9-2 : Profils de charge du circuit 27 nord à 24h05

Ces différences peuvent être causées par des erreurs de comptage ou par un achalandage irrégulier. Mais elles peuvent aussi être la conséquence du retard ou de l'avance du véhicule. En effet, un véhicule en avance aura généralement une clientèle plus faible tandis qu'un autobus en retard ramassera un plus grand nombre de clients (chez lesquels on percevra un certain mécontentement!). Aussi, il ne faut pas négliger les effets que provoquent les problèmes sur d'autres circuits. Un autobus en retard au terminus peut forcer les clients à utiliser un circuit alternatif.

Il serait intéressant de connaître les effets d'un retard ou d'une avance sur le profil de charge d'un circuit donné.

Pour l'instant, la solution la plus appropriée consiste vraisemblablement à calculer la moyenne de toutes les courbes obtenues tout en écartant celles trop éloignées de la moyenne (en comparant avec l'écart-type, par exemple).

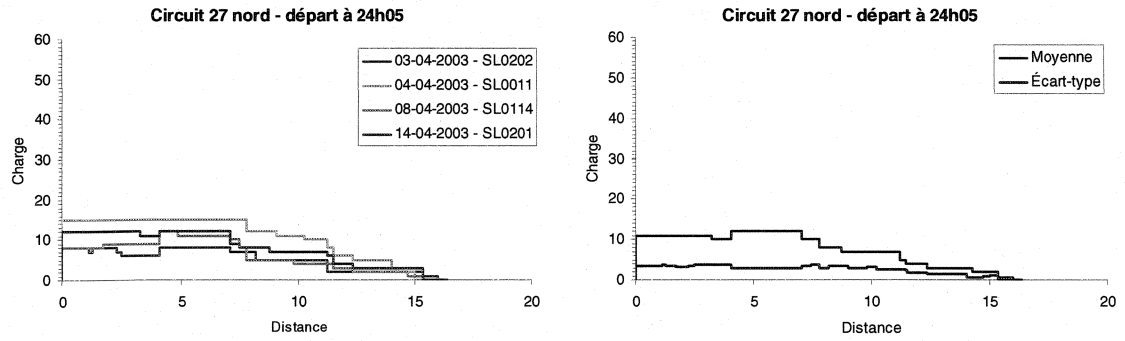


Figure 9-3 : Profils de charge filtrées du circuit 27 nord à 24h05

Malheureusement, cette solution est très limitée et inefficace lorsque les données sont peu nombreuses ou même inexistantes pour certaines tournées.

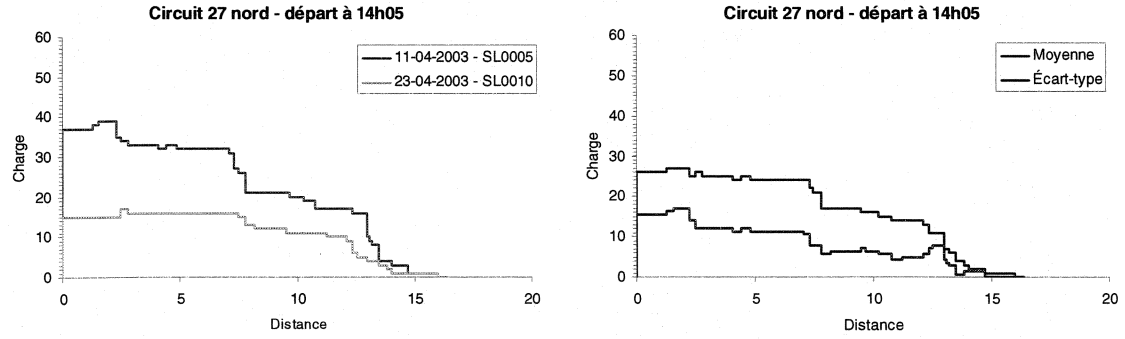


Figure 9-4 : Profils de charge du circuit 27 nord à 14h05

9.4 Interpolations

La Figure 9-5 montre les profils de charge disponibles pour toutes les tournées de la ligne 39 sud (les profils de charge avec une charge résiduelle supérieure à 5 ont été ignorés). On remarque qu'il n'y a aucune donnée disponible pour la tournée de 7h15.

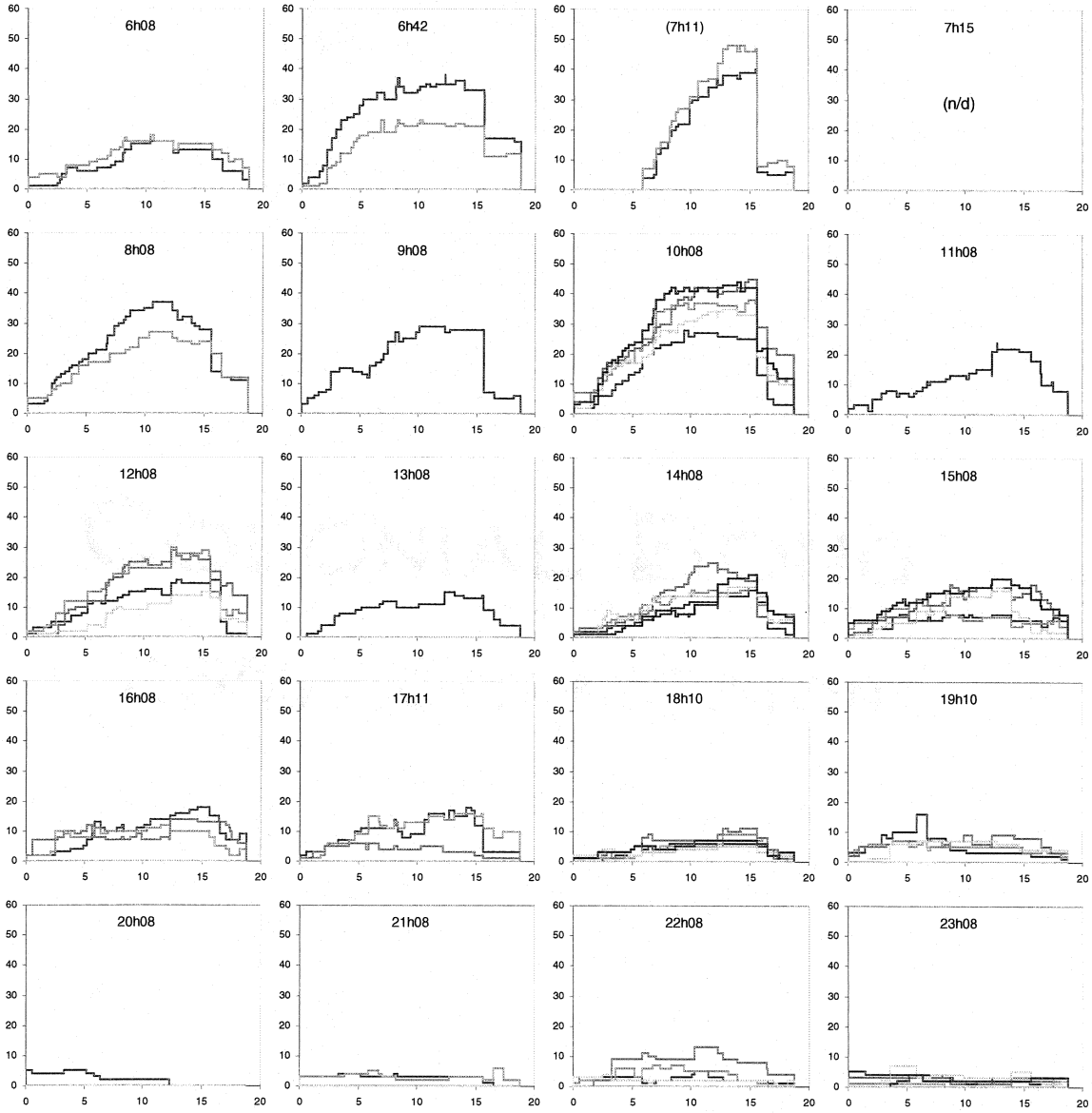


Figure 9-5 : Variations du profil de charge de la ligne 39 sud

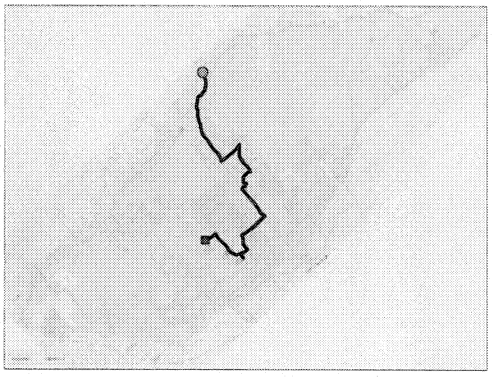


Figure 9-6 : Tracé du circuit 39 sud de la STL

La figure suivante montre la variation de l'achalandage à différents endroits sur le circuit. Une interpolation linéaire est représentée par les 'x'.

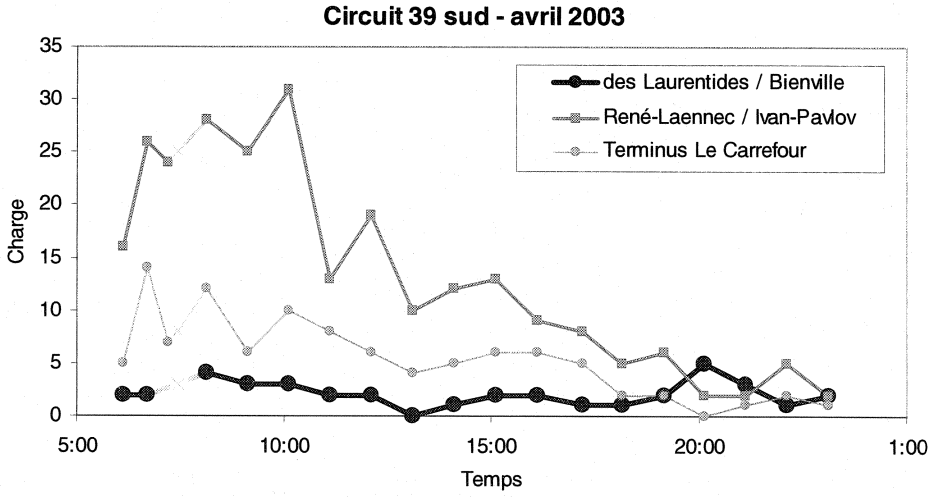


Figure 9-7 : Variation de la charge à différents endroits sur le circuit 39 sud

Cette méthode, à défaut de pouvoir utiliser une autre solution, doit être utilisée avec réserve. Elle relève d'une certaine impertinence puisque la variation de l'achalandage est toute sauf linéaire. De plus, elle s'avère inutilisable lorsqu'il manque beaucoup de données.

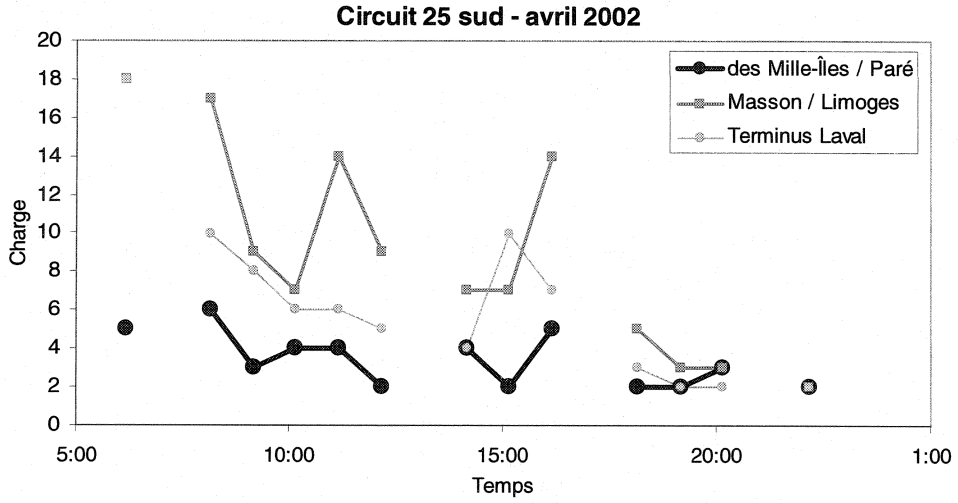


Figure 9-8 : Variation de l'achalandage à différents endroits sur le circuit 25 sud (2002)

Une méthode plus sûre (qui n'est malheureusement pas abordée ici) est l'utilisation d'un modèle basé sur des données historiques, c'est-à-dire des informations accumulées pendant plusieurs mois, et qui tiendrait compte des variations saisonnières.

9.5 Regroupement des tournées en intervalles

Les figures suivantes illustrent la variation du profil de charge en fonction de l'heure de départ pour le circuit 27 nord.

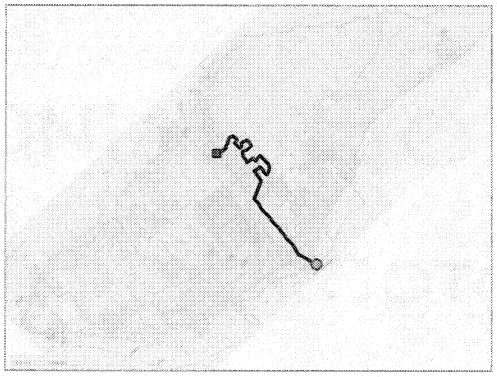


Figure 9-9 : Tracé du circuit 27 nord de la STL

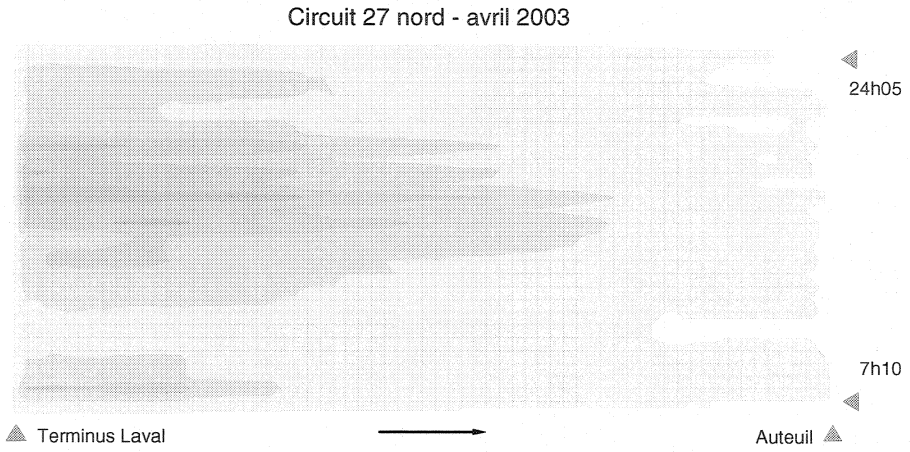


Figure 9-10 : Profil de charge 3D du circuit 27 en fonction des tournées – vue de dessus

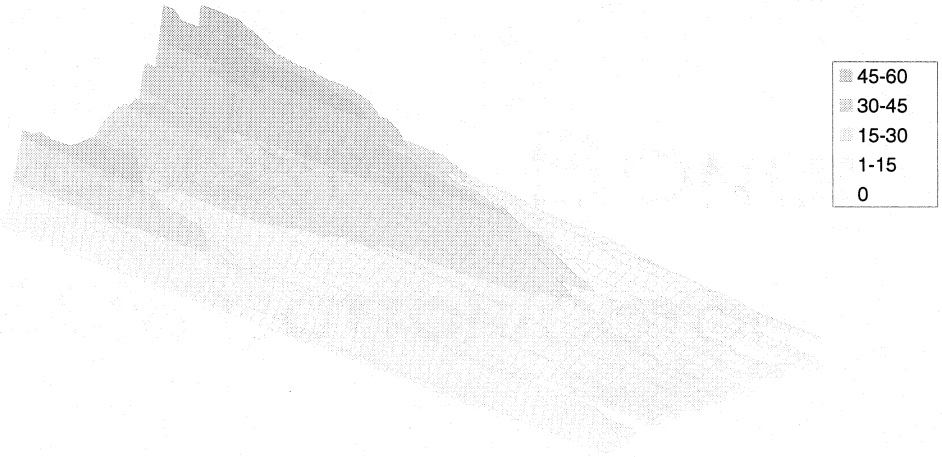


Figure 9-11 : Profil de charge 3D du circuit 27 en fonction des tournées – vue en plongée

Chaque unité en *x* représente un arrêt d'autobus et chaque unité en *y*, un passage d'autobus. Nous savons donc comment évolue la charge à chaque arrêt et à chaque passage d'autobus. Mais, sachant que la fréquence des passages varie en fonction du moment de la journée, comment la charge évolue-t-elle *en fonction du temps*?

Pour le savoir, on regroupe chaque passage en différentes classes, chaque classe représentant un intervalle fixe de temps. Le tableau suivant montre la variation de la charge au terminus Laval après la fermeture des portes.

Tableau 9-2 : Variation de la charge du circuit 27 nord au terminus Laval (avril 2003)

Départs	Charge par départ	Charge par intervalle	Intervalle
7:10	13	46	7h00 à 7h59
7:40	33		
8:05	18	35	8h00 à 8h59
8:35	17		
9:05	14	25	9h00 à 9h59
9:35	11		
10:05	7	7	10h00 à 10h59
11:05	15	15	11h00 à 11h59
12:05	16	16	12h00 à 12h59
13:05	22	22	13h00 à 13h59
14:05	26	54	14h00 à 14h59
14:35	28		
15:05	28	45	15h00 à 15h59
15:35	17		
16:03	40	74	16h00 à 16h59
16:35	34		
17:00	58	177	17h00 à 17h59
17:15	42		
17:26	47		
17:37	30		
18:08	53	74	18h00 à 18h59
18:35	21		
19:05	17	17	19h00 à 19h59
20:05	18	18	20h00 à 20h59
21:05	17	17	21h00 à 21h59
22:05	21	21	22h00 à 22h59
23:05	18	18	23h00 à 23h59
0:05	12	12	24h00 à 24h59

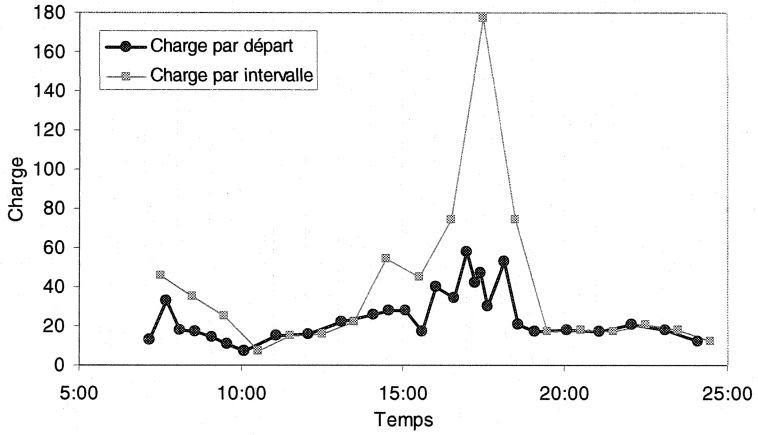


Figure 9-12 : Variation de la charge du circuit 27 nord au terminus Laval (avril 2003)

Le regroupement des charges à tous les points de service donne un profil de charge 3D en fonction du temps, tel qu'illustré ci-dessous.

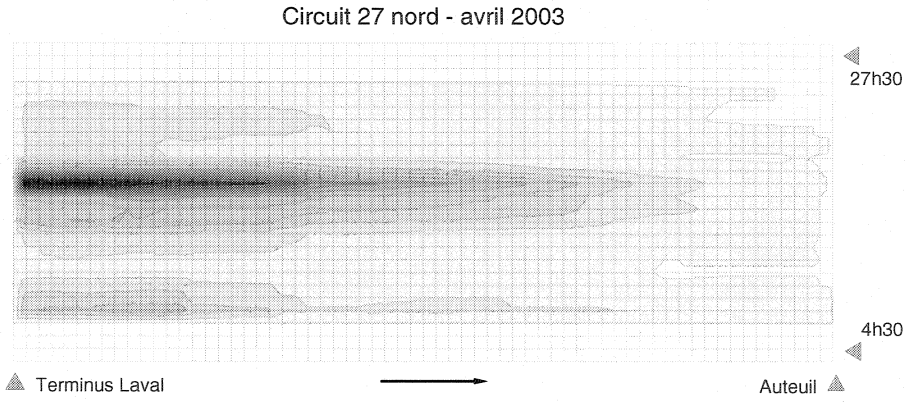


Figure 9-13 : Profil de charge 3D du circuit 27 nord en fonction du temps – vue de dessus

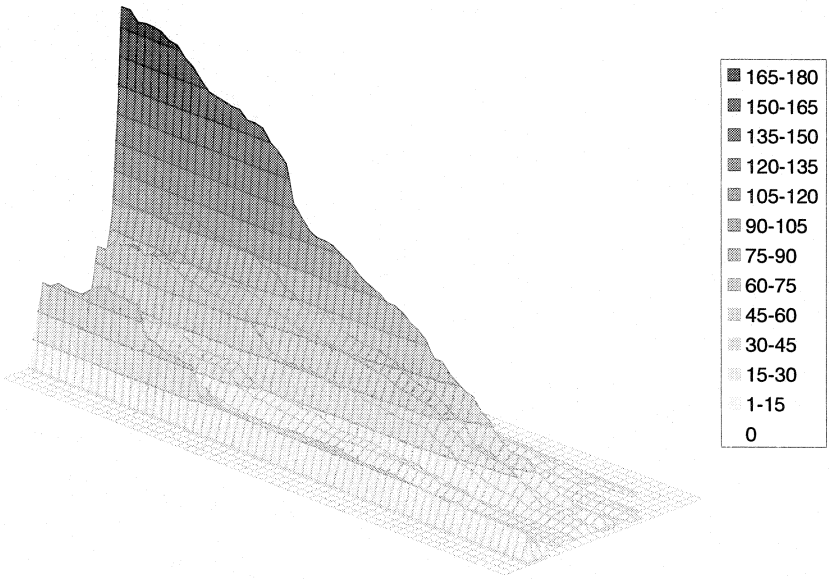


Figure 9-14 : Profil de charge 3D du circuit 27 nord en fonction du temps – vue en plongée

Cette nouvelle organisation des données de charge nous permet d'effectuer des comparaisons avec d'autres parcours. Par exemple, la ligne 27 *dans l'autre direction* :

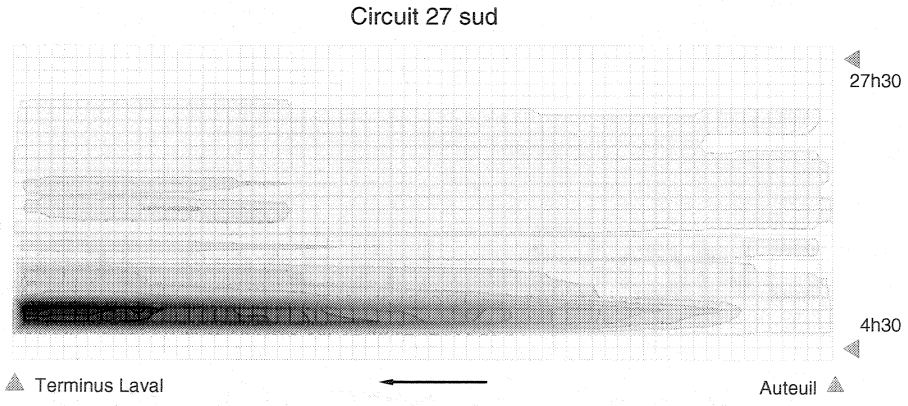


Figure 9-15 : Profil de charge 3D du circuit 27 sud – vue de dessus

9.6 Quelques profils de charge 3D

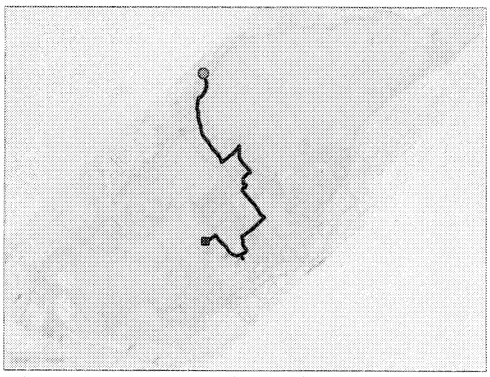


Figure 9-16 : Tracé du circuit 39

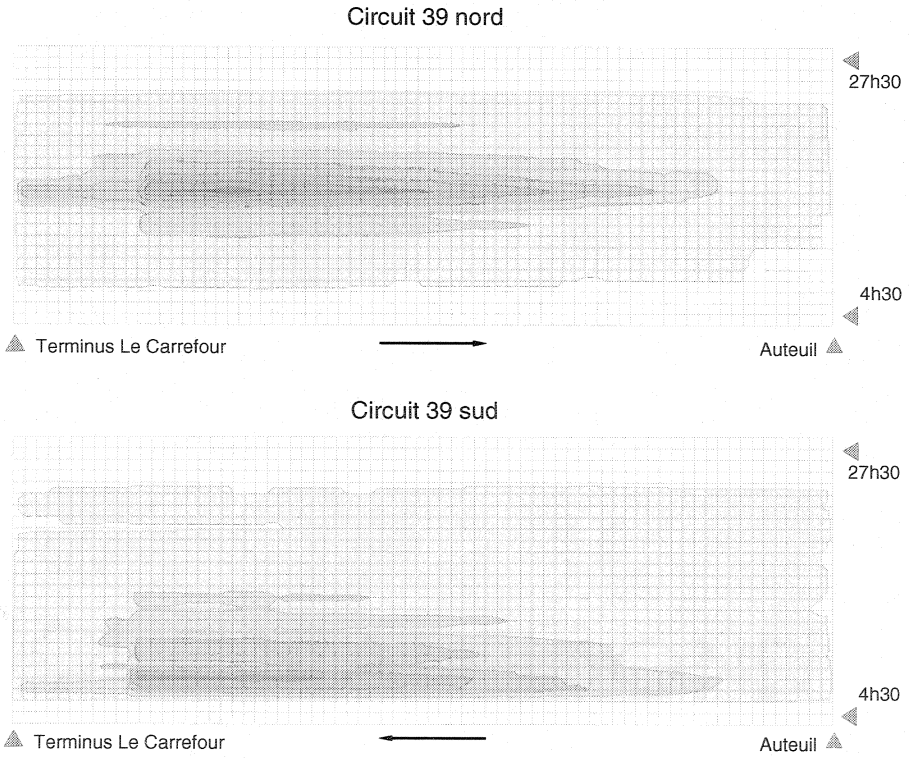


Figure 9-17 : Profils de charge 3D du circuit 39 aller-retour

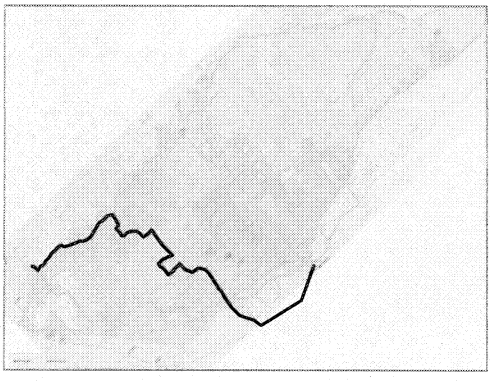


Figure 9-18 : Tracé du circuit 55

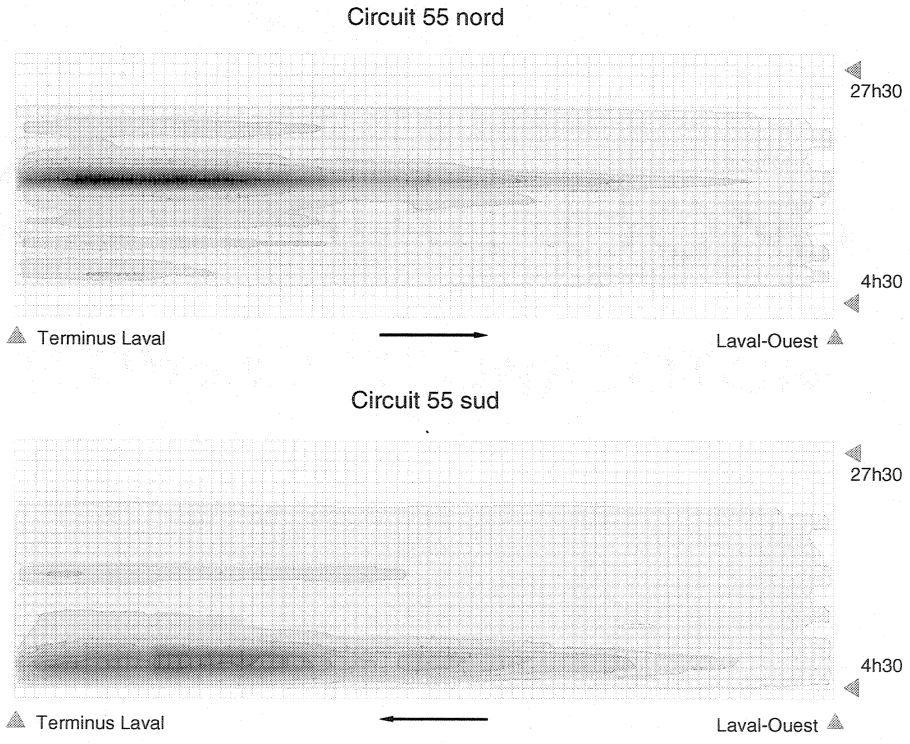


Figure 9-19 : Profils de charge 3D du circuit 55 aller-retour

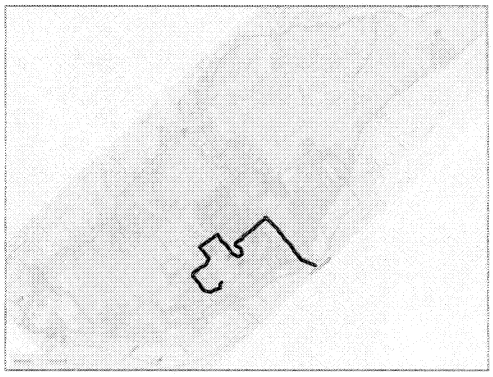


Figure 9-20 : Tracé du circuit 60

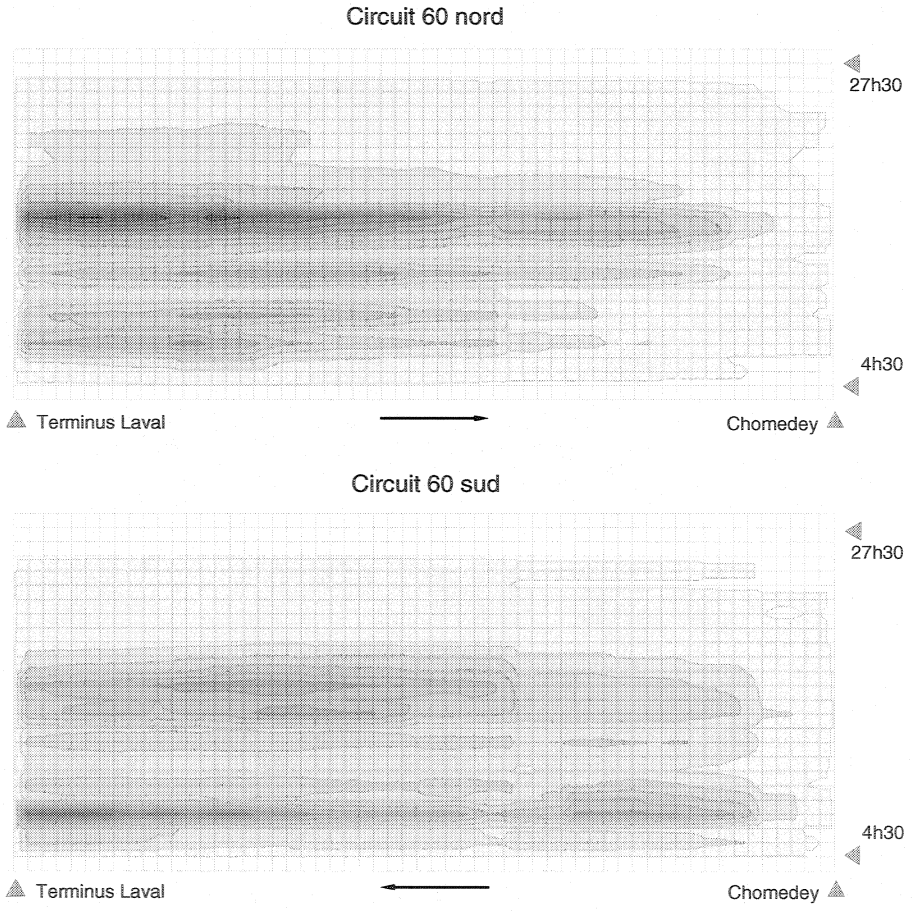


Figure 9-21 : Profils de charge 3D du circuit 60 aller-retour

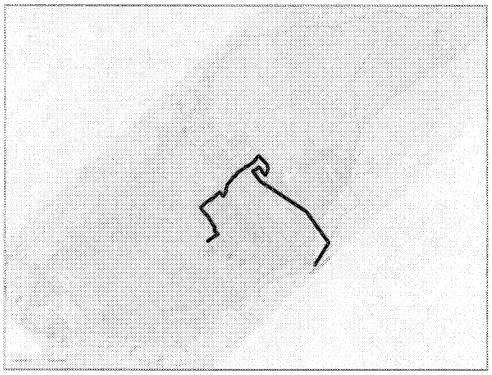


Figure 9-22 : Tracé du circuit 70

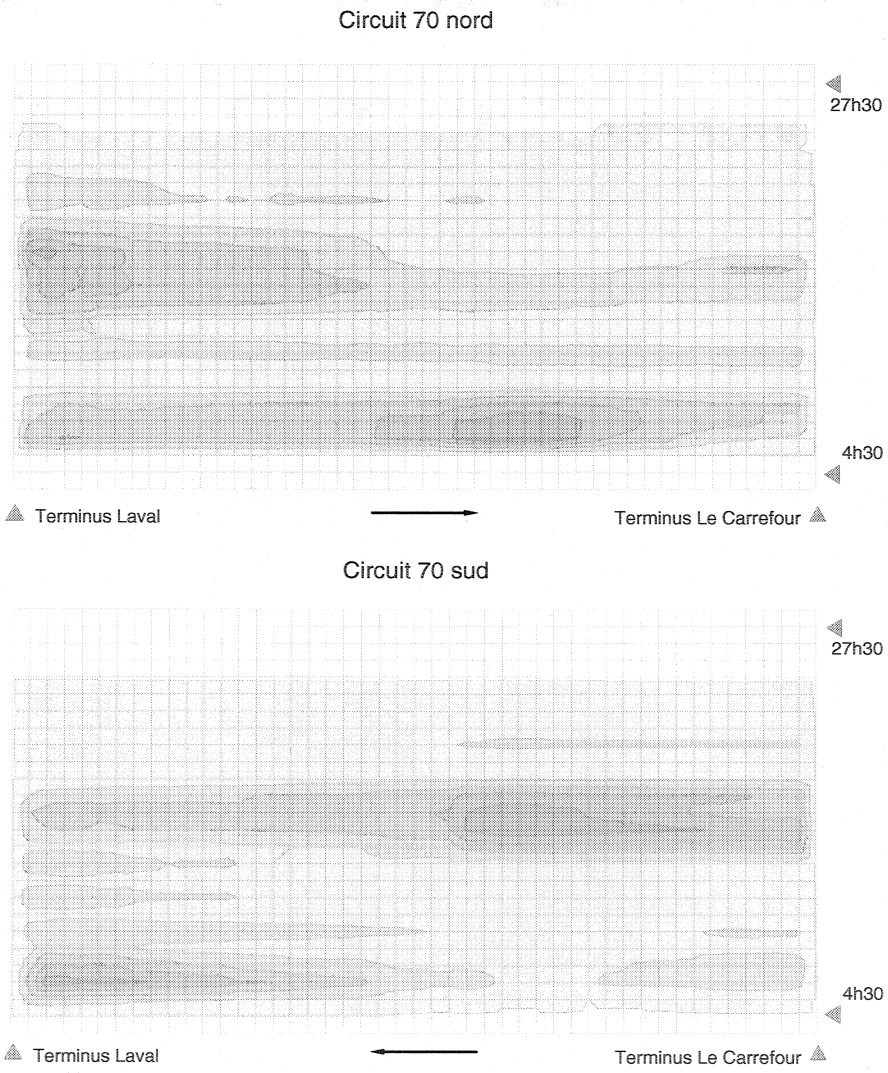


Figure 9-23 : Profils de charge 3D du circuit 70 aller-retour

CHAPITRE 10 : Conclusion

Ce document se veut une introduction à la collecte et au traitement des données GPS et de comptage en transport en commun; l'accent a toutefois été mis sur les méthodes utilisées et les obstacles rencontrés dans le traitement et l'analyse de ces données

Les données utilisées dans ce projet ont tous été fournies par la Société de transport de la Ville de Laval (STL). Les appareils qui ont servis à produire ces données ont été fabriqués par la compagnie Infodev. Le traitement des données à la STL se fait en deux étapes distincts : d'abord la reconnaissance des circuits parcourus avec la première version du logiciel BusStops d'Infodev, puis la vérification des horaires avec le Logiciel d'aide au développement du réseau d'autobus (LADRA), une application développée pour la STL.

Une première analyse des données collectées en avril 2002 et en avril 2003 a permis de révéler deux problématiques importantes qui ont par la suite été analysées en profondeur dans ce document : le comptage des montées et des descentes et la reconnaissance des tournées.

10.1 Comptage automatique des passagers

Le comptage des personnes est un défi complexe à relever et il n'existe pas encore aujourd'hui de système de comptage absolument parfait. Les systèmes Infodev utilisés par la STL commettent certaines erreurs et cela se voit principalement dans la différence entre le nombre de montées et celui des descentes. Certes, lorsque la charge résiduelle est très élevée, on considère que le comptage est peu crédible et on rejette les données. Cependant, et cela a été démontré au chapitre 5, un équilibre entre les montées et les descentes ne signifie pas que le comptage est exact.

Cela dit, il faut garder à l'esprit qu'un comptage parfait n'est pas suffisant pour évaluer l'achalandage sur un réseau. En fait, la composante stochastique (aléatoire) de la variation de l'achalandage peut être plus importante que l'écart de mesure des systèmes de comptage. Il est donc préférable de disposer d'un grand nombre de données imparfaites décrivant l'achalandage sur une tournée que de n'avoir que quelques relevés de systèmes « parfaits ». En général, on semble considérer une erreur équilibrée de 5 à 10 % comme étant raisonnable.

Il peut arriver que certains systèmes soient particulièrement enclins à commettre des erreurs, notamment à cause de l'usure, d'un choc ou d'un mauvais ajustement. Ces systèmes fautifs peuvent parfois être « démasqués » en utilisant les données de comptage automatique sur une période donnée (un jour, une semaine ou un mois). Le chapitre 8 propose quelques méthodes et met en évidence quelques-uns de ces problèmes. D'un autre côté, la STL procède régulièrement à des comptages manuels afin de déterminer l'efficacité de ses systèmes et, en cas de problème, Infodev effectue les réparations nécessaires. Ces données de comptage n'étaient malheureusement pas disponibles au moment où ce projet fut réalisé.

L'influence que peut avoir les erreurs de comptage sur l'analyse de l'achalandage est une question qui n'a pas été abordée dans ce projet. Il serait néanmoins intéressant d'avoir une réponse à ce sujet puisque cela permettrait de déterminer si les systèmes actuels sont suffisamment précis ou si des améliorations doivent être apportées aux systèmes de comptage.

10.2 Reconnaissance des tournées

Comme il a été souligné dans ce document, toutes les tournées effectuées par les autobus munis de systèmes de localisation ne sont pas identifiées comme

telles et la productivité des systèmes n'est donc pas optimale. Par contre, il n'est pas absolument nécessaire que *toutes* les tournées effectuées par les véhicules soient détectées. D'ailleurs, la recherche de la perfection dans ce cas-ci pourrait s'avérer inutilement compliquée, voire coûteuse.

Dans la mesure où ce qui nous intéresse est l'analyse de l'achalandage sur une période donnée, le plus important est d'obtenir la meilleure couverture possible du réseau. Cela implique que chaque tournée planifiée doit être détectée au moins une fois (préférablement plusieurs fois) pendant la période choisie. Il existe plusieurs moyens d'y parvenir :

- Choisir une période suffisamment longue, soit de un à quelques mois.
- S'assurer que le ou les logiciels utilisés soient capables de reconnaître tous les circuits (incluant les variantes) existants sur le réseau.
- Gérer l'utilisation des véhicules en tenant compte de l'état de la couverture du réseau pendant la période choisie – par exemple en assignant des véhicules munis de systèmes Infodev à des tournées qui ne sont pas (encore) représentées.
- Augmenter, au besoin, le nombre de systèmes embarqués.

Dans le cas où certaines tournées ne sont pas représentées par des données pendant la période choisie, on doit alors compléter par des approximations. La solution qui est recommandée à la section 9.4 est l'utilisation de données historiques, par exemple des données obtenues l'année précédente et pondérées à partir de la variation annuelle de l'achalandage.

Un algorithme a été élaboré au chapitre 7 afin de démontrer que la productivité des systèmes pouvait être notablement améliorée. En outre, cet exercice fut un défi intéressant et stimulant pour l'auteur de ce mémoire et cela a permis de mettre à l'épreuve la méthode esquissée à la section 7.5. L'avantage particulier

de cet algorithme est qu'il utilise à la fois les données planifiées *spatiales* (circuits) et *temporelles* (horaires). Le nouveau logiciel BusStops développé par Infodev a lui aussi cet avantage et permet en plus l'utilisation de données *opérationnelles* (voir Figure 3-7).

10.3 Autres considérations

Le traitement des données GPS et de comptage permet, dans sa forme actuelle, l'établissement de l'offre de service sur un réseau de transport en commun par l'analyse de l'achalandage et de la qualité du service actuel. Cette information peut naturellement être comparée avec les résultats d'enquêtes origine-destination comme celles effectuées dans la grande région de Montréal à tous les cinq ans et qui couvre un territoire de plus en plus grand. Un tel couplage permettrait d'obtenir plus de précision sur les déplacements effectués ainsi que sur la clientèle concernée – dans ce cas-ci la population lavalloise.

Les données GPS et de comptage en transport en commun peuvent avantageusement être complété par d'autres technologies informationnelles et ainsi former des systèmes de transport intelligents (STI) plus élaborés. On pense par exemple aux systèmes électroniques de perception des tarifs qui effectuent du même coup le comptage des personnes montant à bord. On pense également à la communication en temps réel avec un éventuel centre de contrôle, comme cela se fait présentement avec les trains de banlieue de l'AMT. La préemption aux feux de circulation est aussi une avenue intéressante qui est déjà expérimentée par le RTL et qui fait ses preuves (RTL, 2004).

Une autre problématique qui a été abordée dans ce document, toutefois avec moins d'emphase et plus de réserves, est le caractère événementiel des données collectées par les systèmes embarqués. Actuellement, les données ne sont collectées qu'à des moments précis correspondant à des événements

particuliers comme les ouvertures de porte, les haltes et les balises virtuelles, et les informations sur l'achalandage se retrouvent discrétisées aux arrêts. Mais on peut affirmer qu'un jour, les informations seront collectées de façon continue partout sur le réseau, d'une façon purement désagrégée, et que la programmation d'événements tels que « ouverture de portes » perdront leur utilité puisque nous connaîtrons précisément le moment où chaque passager monte à bord du véhicule pour ensuite en redescendre…

RÉFÉRENCES

ANDERSSON David, FJELLSTRÖM Johan (2004), *Vehicle Positioning with Map Matching Using Integration of a Dead Reckoning System and GPS*, Examensarbete, Linköpings Universitet, 69 pages.

BLOUIN Richard (1984), *Analyse et modélisation de l'offre de service sur une ligne de transport en commun*, mémoire M.Sc.A, École Polytechnique de Montréal, 291 pages.

CHAPLEAU Robert (2002), *Cours CIV3601 – Systèmes de transport*, notes de cours et exercices, deuxième édition, École Polytechnique de Montréal.

CHAPLEAU Robert, PICHÉ Daniel, CHAPLEAU Marie-Andrée (2003), *Les stations de métro en réalité virtuelle : une autre technologie d'information pour l'usager du transport en commun*, 38e congrès annuel de l'Association québécoise du transport et des routes, Sherbrooke.

DRANE Chris, RIZOS Chris (1998), *Positioning Systems in Intelligent Transportation Systems*, Artech House, 369 pages.

DUPUIS Jérémie (2000), *L.A.D.R.A. (Logiciel d'Aide au Développement du Réseau d'Autobus)*, projet de fin d'études, École Polytechnique de Montréal, 32 pages.

GÉOMATIQUE CANADA (1998), *Guide pour le positionnement GPS*, Ministre des Approvisionnements et Services Canada, 120 pages.

INFODEV (2004), *Infodev, systèmes de comptage de personnes*, site web, www.infodev.ca.

INIT (2004), *Init, Innovation in Traffic Systems*, site web, www.init-ka.de.

IRIS (2004), *Iris Infrared & Intelligent Sensors*, site web, www.apc-irma.biz.

IRIS (2004), *Irma, Infrared Motion Analyzer, Automatic Passenger Counter (APC) System*, document technique, www.irisgmbh.de/dokumentation/IF0402E-htm/files/IF0402E_0000.htm.

JACQUOT Mathieu (2002), *Traitement de données GPS en transport collectif*, rapport de production, École Polytechnique de Montréal, 47 pages.

KWAN Mei-Po (2003), *Mei-Po Kwan's 3D GIS work – Time-Space Paths*, Ohio State University, geog-www.sbs.ohio-state.edu/faculty/mkwan/Gallery/STPaths.htm.

LAVIGUEUR Pierre (2001), *Instrumentation embarquée à la STL*, Routes et transports vol. 30 no 2, pages 21 à 25, www.aqtr.qc.ca/groupes/revue/instrumentationstl.pdf.

McDONOUGH Kim (2002), *La réalité virtuelle appliquée aux systèmes d'information de transport urbain collectif*, mémoire M.Sc.A, École Polytechnique de Montréal, 141 pages.

MINSKY Marvin (1986), *The Society of Mind*, Touchstone, New York, 339 pages.

MORENCY Catherine (2001), *Le système de positionnement global (GPS) et les transports : de multiples défis techniques en perspective*, examen général de synthèse, École Polytechnique de Montréal, 87 pages.
(Extrait : http://www.transport.polymtl.ca/person/EtudH02/cmorency.htm)

NCHRP (National Cooperative Highway Research Program) (1998), *Applications of GPS for surveying and other positioning needs in departments of transportation*, NCHRP Synthesis Reports 258, 46 pages.

PARKINSON Bradford W., SPILKER JR. James J. (1996), *Global Positioning System: Theory and Applications - volume 1*, American Institute of Aeronotics and Astronautics Inc., 793 pages.

RIENDEAU Daniel, CHAPLEAU Robert (2004), *Données GPS et transport en commun : vers un modèle spatio-temporel saisonnalisé de la demande en transport*, 39e Congrès annuel de l'Association québécoise du transport et des routes, Québec, 19 pages.

RTL (Réseau de transport de Longueuil) (2001), *Les technologies à bord des autobus de la STRSM*, Colloque de l'Association québécoise du transport et des routes sur les systèmes de transport intelligent, Montréal, 38 pages, www.aqtr.qc.ca/groupes/sti/B03_Cormier.pdf.

RTL (Réseau de transport de Longueuil) (2004), *Projet Préemption des feux de circulation, Axe du chemin de Chambly*, Congrès annuel de l'Association québécoise du transport et des routes, Québec, 19 pages.

STL (2004), *Société de transport de la Ville de Laval*, site web, www.stl.laval.qc.ca.

TRÉPANIER Martin (1999), *Modélisation totalement désagrégée et orientée-objet appliquée aux transports urbains*, thèse Ph.D. École Polytechnique de Montréal, 229 pages.

TRB (Transportation Research Board) (1998), *Passenger Counting Technologies and Procedures*, Transit Cooperative Research Program Synthesis 29, 50 pages.

UTRC2 (Region 2 University Transportation Research Center) (2003), *Impact of Congestion on Bus Operations and Costs*, Federal Highway Administration, New Jersey, 72 pages, www.utrc2.org/research/assets/67/buscongestion1.pdf.

WATZLAWICK Paul (1986), *Comment réussir à échouer*, Seuil, 118 pages.

ZHAO Yilin (1997), *Vehicle Location and Navigation Systems*, Artech House, 345 pages.

ANNEXE 1 : fichiers de données

Données brutes

Les données téléchargées à partir des ordinateurs embarqués sont originalement en format binaire. Un programme conçu par Infodev les convertit en format CSV (*comma-separated values*). Un exemple de fichier CSV ouvert dans Microsoft Excel est présenté à la Figure 3-3 de ce mémoire.

Les quatre premières données en entête d'un fichier brut correspondent au nom du fournisseur (« INFODEV »), au nom de l'acquisiteur de données (« DL10B »), à sa version (« 1.1 ») et au numéro du véhicule sur lequel est installé le système.

Le nom des fichiers bruts respecte la nomenclature suivante :

`[année] [mois] [jour] [unité] [suffixe].csv`

L'année est représentée par une lettre : A = 2000, B = 2001, C = 2002, etc. Le mois et le jour sont représentés par deux chiffres. L'unité est le numéro d'identification à deux chiffres de l'ordinateur embarqué. Le suffixe sert à distinguer les fichiers créés le même jour par le même véhicule : a = premier fichier, b = second fichier, c = troisième fichier, etc.

Le reste du fichier contient les informations suivantes :

Variables contenues dans les fichiers bruts

Variable	Signification	Événement déclencheur						
		B	E	H	P	v	V	X
Type	Type d'événement B = démarrage du moteur E = coupure du moteur H = halte de plus de 15 s sans ouverture de porte P = ouverture de portes v = pré-détection de balise (?) V = détection de balise X = transfert des données (fin de fichier)							
YY, MO, DD	Année, mois, jour	•	•					•
AHH, AMM, ASS	Heure, minute, seconde au début de l'événement	•	•	•	•	•	•	•
DHH, DMM, DSS	Heure, minute, seconde à la fin de l'événement			•	•		•	•
ODO	Relevé de l'odomètre au mètre près (débute à zéro au début du fichier)	•	•	•	•		•	
LAT, LON	Latitude, longitude en coordonnées MTM (NAD83)	•	•	•	•		•	
HEAD	Orientation du véhicule (0 à 359°)			•	•			
IN1R, OUT1R, IN2R, OUT2R, IN3F, OUT3F	Nombre de montées (in) et de descentes (out) aux portes arrières (r) et à la porte avant (f)				•			
GSTAT	Qualité de lecture du GPS (≈ nombre de satellites détectés)	•	•	•	•			
VZONE	Numéro de balise virtuelle					•	•	
GDIST	Distance parcourue (selon l'odomètre) lorsque aucune donnée GPS n'est disponible	•	•	•	•			

Données BusStops et LADRA

Chaque fichier brut est traité avec BusStops (couplage avec les circuits planifiés), lequel sauvegarde le résultat dans un autre fichier CSV. Ce nouveau fichier est ensuite traité avec LADRA (couplage avec les horaires planifiés), lequel sauvegarde le résultat dans un troisième fichier CSV. Les fichiers BusStops et LADRA possèdent le même nom que le fichier brut d'origine auquel est ajouté un préfixe. Les fichiers BusStops commencent par R tandis que les fichiers LADRA commencent par J.

Le contenu des fichiers BusStops et LADRA est semblable; LADRA ne fait qu'ajouter son résultat aux informations de BusStops.

Variables contenues dans les fichiers R et J

Variable	Signification	Fichier	
		R	J
Type	Type d'événement (tous les événements du fichier original sauf v et X)	●	●
Ligne	Numéro de ligne (clé primaire)	●	
Heure D	Moment de début d'événement (date et heure à la seconde près)	●	●
Heure F	Moment de fin d'événement (date et heure à la seconde près)	●	●
Trajet	Temps écoulé depuis la fin de l'événement précédent (arrondi à l'heure près dans le fichier J)	●	●
Arrêté	Temps écoulé entre le début et la fin de l'événement (arrondi à l'heure près dans le fichier J)	●	●
Distance	Distance parcourue depuis l'événement précédent (selon l'odomètre, au mètre près)	●	●
Parcours	Numéro de circuit, direction et variante (BusStops) (Direction : 0 = nord, 1 = sud, 2 = est, 3 = ouest)	●	●
Arrêt No	Numéro d'arrêt d'autobus (BusStops)	●	●
Arrêt	Nom de l'arrêt d'autobus (BusStops)	●	●
Montées	Nombre total de montées (en circuit seulement)	●	●
Descentes	Nombre total de descentes (en circuit seulement)	●	●
Charge	Nombre de personnes à bord (en circuit seulement) (remis à zéro à chaque début de circuit)	●	●
Gstat	Qualité de lecture du GPS (≈ nombre de satellites détectés)	●	●
LongDelta, LatDelta	Différence en mètre entre la position obtenue par GPS (*Long*, *Lat*) et la position théorique de l'arrêt d'autobus détecté	●	●
Long, Lat	Latitude, longitude en coordonnées MTM (NAD83)	●	●
Direction	Orientation du véhicule (0 à 359°)	●	●
In0, Out0, In1, Out1, In2, Out2	Nombre de montées (In) et de descentes (Out) aux portes arrière (0 et 1) et à la porte avant (2)	●	●
Départ	Heure de départ planifiée (Hastus)		●
Passage	Heure planifiée de passage à l'arrêt d'autobus (Hastus, interpolé)		●
Index	Numéro de séquence de l'arrêt dans le circuit (Hastus) (multiplié par 1024)		●
État	Code d'erreur (voir ci-dessous)		●
Matri	Matricule du chauffeur assigné (Hastus)		●
Assign	Numéro d'assignation (données opérationnelles)		●
Voiture	Numéro de voiture/véhicule (Hastus)		●
Chauffeur	Matricule du chauffeur en service (données opérationnelles)		●
PositionX, PositionY	Position théorique de l'arrêt d'autobus détecté en coordonnées MTM (NAD83)		●
GPSDist	Distance parcourue (selon l'odomètre) lorsque aucune donnée GPS n'est disponible	●	●
Contrôle	Numéro de balise virtuelle	●	●
Trouvé	Repère, découpage Infodev (?)	●	●

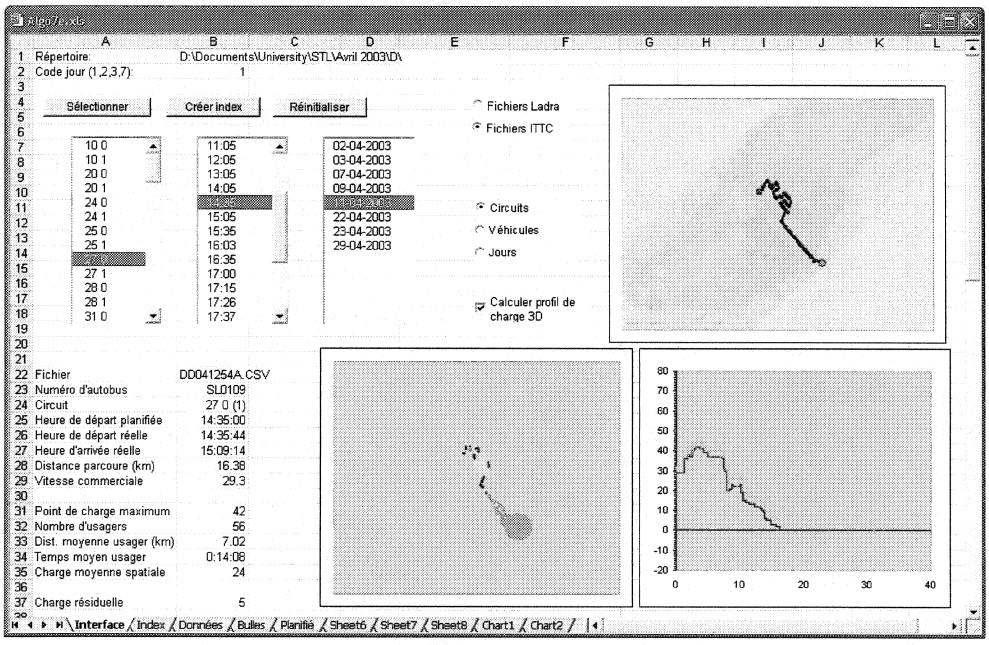
Liste des codes d'erreur de LADRA :

0 = correctement analysé, arrêt régulier
1 = correctement analysé, point de contrôle
2 = aucun circuit détecté
3 = booking non trouvé
4 = aucun chauffeur alloué pour l'autobus
5 = aucune assignation ou voiture trouvée
6 = mauvais circuit ou direction au départ
7 = aucun départ prévu à ce moment
8 = nom d'arrêt non présent dans le circuit
9 = événement H sans nom d'arrêt et sans montant ni descendant
10 = intervalle de passage à l'arrêt non respecté
11 = correctement analysé, point de contrôle de départ
12 = correctement analysé, point de contrôle d'arrivée
Vide = erreur, situation irrégulière

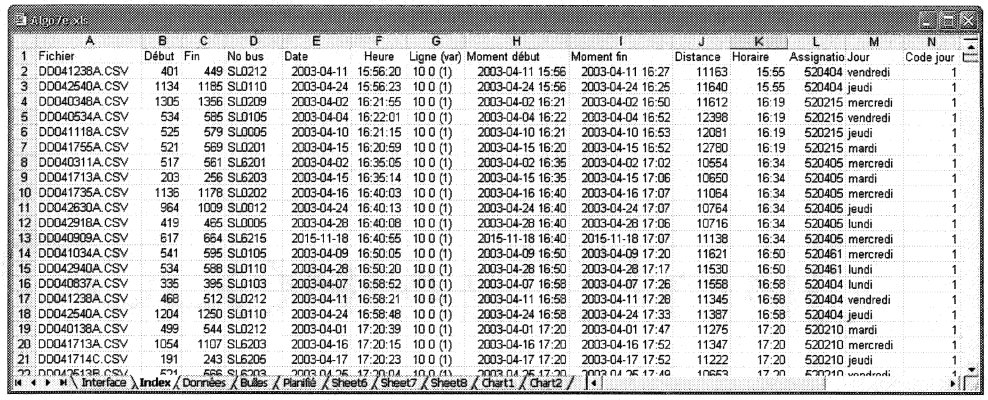
ANNEXE 2 : algorithme de visualisation de tournées

Afin de faciliter l'accès à la totalité des données contenues dans les fichiers CSV de LADRA, un algorithme fut conçu avec l'interface VBA de Microsoft Excel. Cet algorithme effectue les fonctions suivantes :

- Création d'un index des voyages effectués pendant une période donnée;
- Création de menus permettant la sélection d'un voyage à la fois;
- Pour chaque voyage sélectionné, présentation de résultats graphiques variés : tracé théorique et GPS, analyse des montants et des descendants (bulles), profils de charge mesuré et corrigé;
- Pour chaque circuit sélectionné, présentation de profils de charges 3D en fonction des tournées et du temps.



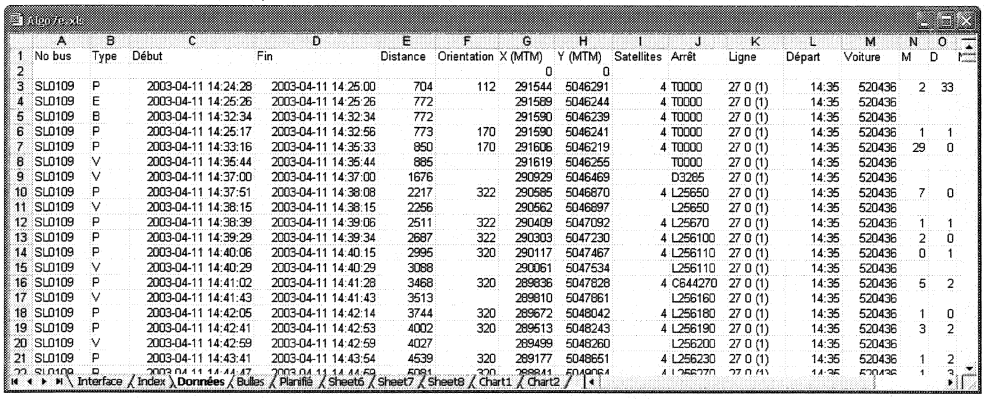
Feuille 1 : interface de visualisation



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Fichier	Début	Fin	No bus	Date	Heure	Ligne (var)	Moment début	Moment fin	Distance	Horaire	Assignatio	Jour	Code jour
2	DD041238A.CSV	401	449	SL0212	2003-04-11	15:56:20	10 0 (1)	2003-04-11 15:56	2003-04-11 16:27	11163	15:55	520404	vendredi	1
3	DD042540A.CSV	1134	1185	SL0110	2003-04-24	15:56:23	10 0 (1)	2003-04-24 15:56	2003-04-24 16:25	11640	15:55	520404	jeudi	1
4	DD040348A.CSV	1305	1356	SL0209	2003-04-02	16:21:55	10 0 (1)	2003-04-02 16:21	2003-04-02 16:50	11612	16:19	520215	mercredi	1
5	DD040534A.CSV	534	585	SL0105	2003-04-04	16:22:01	10 0 (1)	2003-04-04 16:22	2003-04-04 16:52	12398	16:19	520215	vendredi	1
6	DD041118A.CSV	525	579	SL0005	2003-04-10	16:21:15	10 0 (1)	2003-04-10 16:21	2003-04-10 16:53	12081	16:19	520215	jeudi	1
7	DD041755A.CSV	521	569	SL0201	2003-04-15	16:20:59	10 0 (1)	2003-04-15 16:20	2003-04-15 16:52	12780	16:19	520215	mardi	1
8	DD040311A.CSV	517	561	SL6201	2003-04-02	16:35:05	10 0 (1)	2003-04-02 16:35	2003-04-02 17:02	10554	16:34	520405	mercredi	1
9	DD041713A.CSV	203	256	SL6203	2003-04-15	16:35:14	10 0 (1)	2003-04-15 16:35	2003-04-15 17:06	10650	16:34	520405	mardi	1
10	DD041735A.CSV	1136	1178	SL0202	2003-04-16	16:40:03	10 0 (1)	2003-04-16 16:40	2003-04-16 17:07	11054	16:34	520405	mercredi	1
11	DD042630A.CSV	964	1009	SL0012	2003-04-24	16:40:13	10 0 (1)	2003-04-24 16:40	2003-04-24 17:07	10764	16:34	520405	jeudi	1
12	DD042918A.CSV	419	465	SL0005	2003-04-28	16:40:08	10 0 (1)	2003-04-28 16:40	2003-04-28 17:06	10716	16:34	520405	lundi	1
13	DD040909A.CSV	617	664	SL6215	2015-11-18	16:40:55	10 0 (1)	2015-11-18 16:40	2015-11-18 17:07	11138	16:34	520405	mercredi	1
14	DD041034A.CSV	541	595	SL0105	2003-04-09	16:50:05	10 0 (1)	2003-04-09 16:50	2003-04-09 17:20	11621	16:50	520461	mercredi	1
15	DD042940A.CSV	534	588	SL0110	2003-04-28	16:50:20	10 0 (1)	2003-04-28 16:50	2003-04-28 17:17	11530	16:50	520461	lundi	1
16	DD040837A.CSV	336	395	SL0103	2003-04-07	16:58:52	10 0 (1)	2003-04-07 16:58	2003-04-07 17:26	11558	16:58	520404	lundi	1
17	DD041238A.CSV	468	512	SL0212	2003-04-11	16:58:21	10 0 (1)	2003-04-11 16:58	2003-04-11 17:28	11345	16:58	520404	vendredi	1
18	DD042540A.CSV	1204	1250	SL0110	2003-04-24	16:58:48	10 0 (1)	2003-04-24 16:58	2003-04-24 17:33	11387	16:58	520404	jeudi	1
19	DD040138A.CSV	499	544	SL0212	2003-04-01	17:20:39	10 0 (1)	2003-04-01 17:20	2003-04-01 17:47	11275	17:20	520210	mardi	1
20	DD041713A.CSV	1054	1107	SL6203	2003-04-16	17:20:15	10 0 (1)	2003-04-16 17:20	2003-04-16 17:52	11347	17:20	520210	mercredi	1
21	DD041714C.CSV	191	243	SL6205	2003-04-17	17:20:23	10 0 (1)	2003-04-17 17:20	2003-04-17 17:52	11222	17:20	520210	jeudi	1
22	DD042513B.CSV	521	566	SL6203	2003-04-25	17:20:04	10 0 (1)	2003-04-25 17:20	2003-04-25 17:49	10653	17:20	520210	vendredi	1

Interface / Index / Données / Bulles / Planifié / Sheet6 / Sheet7 / Sheet8 / Chart1 / Chart2

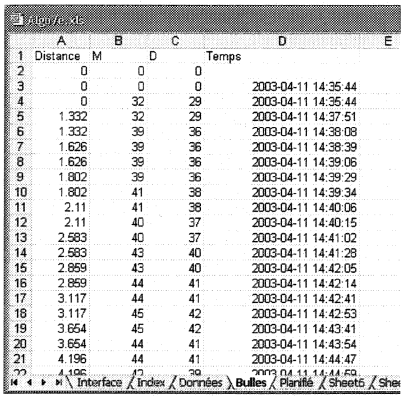
Feuille 2 : index des voyages



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	No bus	Type	Début	Fin	Distance	Orientation	X (MTM)	Y (MTM)	Satellites	Arrêt	Ligne	Départ	Voiture	M	D
2							0	0							
3	SL0109	P	2003-04-11 14:24:28	2003-04-11 14:25:00	704	112	291544	5046291	4	T0000	27 0 (1)	14:35	520436	2	33
4	SL0109	E	2003-04-11 14:25:26	2003-04-11 14:25:26	772		291589	5046244	4	T0000	27 0 (1)	14:35	520436		
5	SL0109	B	2003-04-11 14:32:34	2003-04-11 14:32:34	772		291590	5046239	4	T0000	27 0 (1)	14:35	520436		
6	SL0109	P	2003-04-11 14:25:17	2003-04-11 14:32:56	773	170	291590	5046241	4	T0000	27 0 (1)	14:35	520436	1	1
7	SL0109	P	2003-04-11 14:33:16	2003-04-11 14:35:33	850	170	291606	5046219	4	T0000	27 0 (1)	14:35	520436	29	0
8	SL0109	V	2003-04-11 14:35:44	2003-04-11 14:35:44	885		291619	5046255		T0000	27 0 (1)	14:35	520436		
9	SL0109	V	2003-04-11 14:37:00	2003-04-11 14:37:00	1676		290929	5046469		D3285	27 0 (1)	14:35	520436		
10	SL0109	P	2003-04-11 14:37:51	2003-04-11 14:38:08	2217	322	290585	5046870	4	L25650	27 0 (1)	14:35	520436	7	0
11	SL0109	V	2003-04-11 14:38:15	2003-04-11 14:38:15	2256		290562	5046897		L25650	27 0 (1)	14:35	520436		
12	SL0109	P	2003-04-11 14:38:39	2003-04-11 14:39:06	2511	322	290409	5047092	4	L25670	27 0 (1)	14:35	520436	1	1
13	SL0109	P	2003-04-11 14:39:29	2003-04-11 14:39:34	2687	322	290303	5047230	4	L256100	27 0 (1)	14:35	520436	2	0
14	SL0109	P	2003-04-11 14:40:06	2003-04-11 14:40:15	2995	320	290117	5047467	4	L256110	27 0 (1)	14:35	520436	0	1
15	SL0109	V	2003-04-11 14:40:29	2003-04-11 14:40:29	3088		290061	5047534		L256110	27 0 (1)	14:35	520436		
16	SL0109	P	2003-04-11 14:41:02	2003-04-11 14:41:28	3468	320	289836	5047828	4	C644270	27 0 (1)	14:35	520436	5	2
17	SL0109	V	2003-04-11 14:41:43	2003-04-11 14:41:43	3513		289810	5047861		L256160	27 0 (1)	14:35	520436		
18	SL0109	P	2003-04-11 14:42:05	2003-04-11 14:42:14	3744	320	289672	5048042	4	L256180	27 0 (1)	14:35	520436	1	0
19	SL0109	P	2003-04-11 14:42:41	2003-04-11 14:42:53	4002	320	289513	5048243	4	L256190	27 0 (1)	14:35	520436	3	2
20	SL0109	V	2003-04-11 14:42:59	2003-04-11 14:42:59	4027		289499	5048260		L256200	27 0 (1)	14:35	520436		
21	SL0109	P	2003-04-11 14:43:41	2003-04-11 14:43:54	4539	320	289177	5048651	4	L256230	27 0 (1)	14:35	520436	1	2
22	SL0109	P	2003-04-11 14:44:47	2003-04-11 14:44:59	5081	320	288841	5049064	4	L256270	27 0 (1)	14:35	520436	1	3

Interface / Index / Données / Bulles / Planifié / Sheet6 / Sheet7 / Sheet8 / Chart1 / Chart2

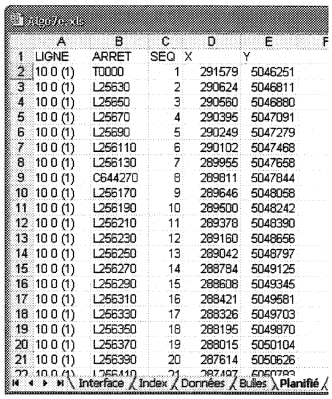
Feuille 3 : données correspondant au voyage sélectionné



	A	B	C	D	E
1	Distance	M	D	Temps	
2	0	0	0		
3	0	0	0	2003-04-11 14:35:44	
4	0	32	29	2003-04-11 14:35:44	
5	1.332	32	29	2003-04-11 14:37:51	
6	1.332	39	36	2003-04-11 14:38:08	
7	1.626	39	36	2003-04-11 14:38:39	
8	1.626	39	36	2003-04-11 14:39:06	
9	1.802	39	36	2003-04-11 14:39:29	
10	1.802	41	38	2003-04-11 14:39:34	
11	2.11	41	38	2003-04-11 14:40:06	
12	2.11	40	37	2003-04-11 14:40:15	
13	2.583	40	37	2003-04-11 14:41:02	
14	2.583	43	40	2003-04-11 14:41:28	
15	2.859	43	40	2003-04-11 14:42:05	
16	2.859	44	41	2003-04-11 14:42:14	
17	3.117	44	41	2003-04-11 14:42:41	
18	3.117	45	42	2003-04-11 14:42:53	
19	3.654	45	42	2003-04-11 14:43:41	
20	3.654	44	41	2003-04-11 14:43:54	
21	4.196	44	41	2003-04-11 14:44:47	
22	4.196	42	39	2003-04-11 14:44:59	

Interface / Index / Données / Bulles / Planifié / Sheet6 / She

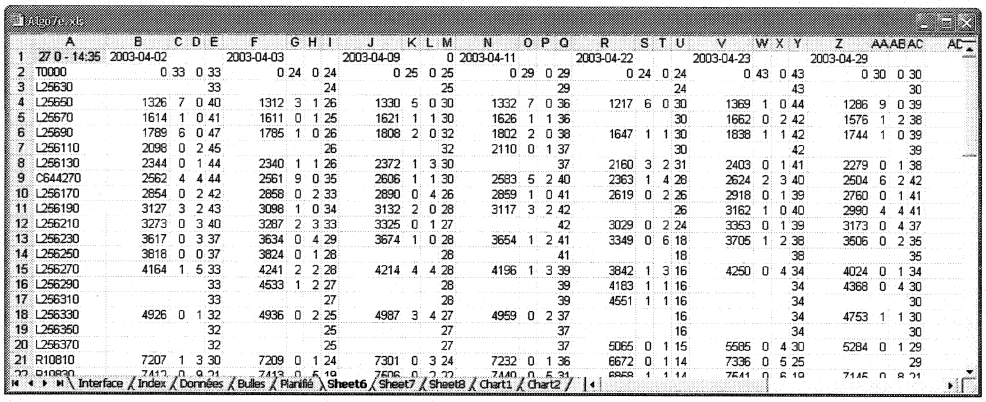
Feuille 4 : profil de charge du voyage sélectionné



	A	B	C	D	E	F
1	LIGNE	ARRET	SEQ	X	Y	
2	10 0 (1)	T0000	1	291579	5046251	
3	10 0 (1)	L25630	2	290624	5046811	
4	10 0 (1)	L25650	3	290580	5046880	
5	10 0 (1)	L25670	4	290395	5047091	
6	10 0 (1)	L25690	5	290249	5047279	
7	10 0 (1)	L256110	6	290102	5047468	
8	10 0 (1)	L256130	7	289955	5047658	
9	10 0 (1)	C644270	8	289811	5047844	
10	10 0 (1)	L256170	9	289646	5048058	
11	10 0 (1)	L256190	10	289500	5048242	
12	10 0 (1)	L256210	11	289378	5048390	
13	10 0 (1)	L256230	12	289160	5048656	
14	10 0 (1)	L256250	13	289042	5048797	
15	10 0 (1)	L256270	14	288784	5049125	
16	10 0 (1)	L256290	15	288608	5049345	
17	10 0 (1)	L256310	16	288421	5049581	
18	10 0 (1)	L256330	17	288326	5049703	
19	10 0 (1)	L256350	18	288195	5049870	
20	10 0 (1)	L256370	19	288015	5050104	
21	10 0 (1)	L256390	20	287614	5050626	
22	10 0 (1)	L256410	21	287497	5050783	

Interface / Index / Données / Bulles / Planifié /

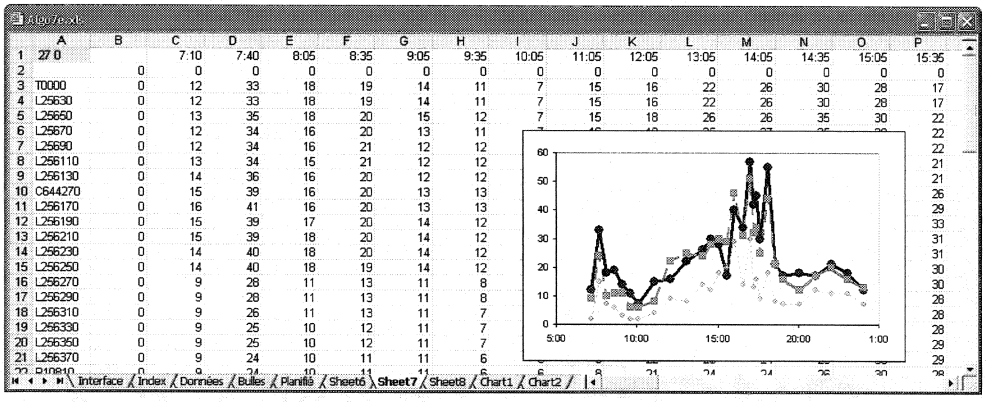
Feuille 5 : données planifiées



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AC
1	27 0 - 14:35	2003-04-02				2003-04-03				2003-04-09				0	2003-04-11			2003-04-22				2003-04-23				2003-04-29				
2	T0000		0	33	0	33		0	24	0	24		0	25	0	25		0	29	0	29		0	24	0	24		0	43	0	43	
3	L25630					33			24					25					29				24				43			30	
4	L25650	1326	7	0	40	1312	3	1	26	1330	5	0	30	1332	7	0	36	1217	6	0	30	1369	1	0	44	1286	9	0	39	
5	L25670	1614	1	0	41	1611	0	1	25	1621	1	1	30	1626	1	1	36				30	1662	0	2	42	1576	1	2	38	
6	L25690	1789	6	0	47	1785	1	0	26	1808	2	0	32	1802	2	0	38	1647	1	1	30	1838	1	1	42	1744	1	0	39	
7	L256110	2098	0	2	45				26				32	2110	0	1	37				30				42				39	
8	L256130	2344	0	1	44	2340	1	1	26	2372	1	3	30				37	2160	3	2	31	2403	0	1	41	2279	0	1	38	
9	C644270	2562	4	4	44	2561	9	0	35	2606	1	1	30	2583	5	2	40	2363	1	4	28	2624	2	3	40	2504	6	2	42	
10	L256170	2854	0	2	42	2858	0	2	33	2890	0	4	26	2859	1	0	41	2619	0	2	26	2918	0	1	39	2760	0	1	41	
11	L256190	3127	3	2	43	3098	1	0	34	3132	2	0	28	3117	3	2	42				26	3162	1	0	40	2990	4	4	41	
12	L256210	3273	0	3	40	3287	2	3	33	3325	0	1	27				42	3029	0	2	24	3353	0	1	39	3173	0	4	37	
13	L256230	3617	0	3	37	3634	0	4	29	3674	1	0	28	3654	1	2	41	3349	0	6	18	3705	1	2	38	3506	0	2	36	
14	L256250	3818	0	0	37	3824	0	1	28				28				41				18				38				36	
15	L256270	4164	1	5	33	4241	2	2	28	4214	4	4	28	4196	1	3	39	3842	1	3	16	4250	0	4	34	4024	0	1	34	
16	L256290				33	4533	1	2	27				28				39	4183	1	1	16				34	4368	0	4	30	
17	L256310				33				27				28				39	4551	1	1	16				34				30	
18	L256330	4926	0	1	32	4936	0	2	25	4987	3	4	27	4959	0	2	37				16				34	4753	1	1	30	
19	L256350				32				25				27				37				16				34				30	
20	L256370				32				25				27				37	5065	0	1	15	5585	0	4	30	5284	0	1	29	
21	R10810	7207	1	3	30	7209	0	1	24	7301	0	3	24	7232	0	1	36	6672	0	1	14	7336	0	5	25				29	
22	R10830	7412	0	9	21	7413	0	5	19	7506	0	2	22	7440	0	5	31	6868	1	1	14	7541	0	6	19	7145	0	8	21	

Interface / Index / Données / Bulles / Planifié / Sheet6 / Sheet7 / Sheet8 / Chart1 / Chart2 /

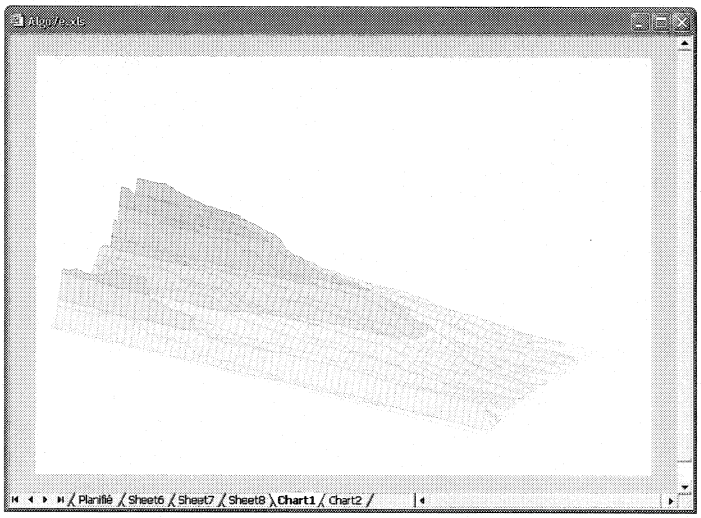
Feuille 6 : distance, montée, descente et charge à chaque arrêt pour toutes les occurrences d'une tournée sélectionnée



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	27 0		7:10	7:40	8:05	8:35	9:05	9:35	10:05	11:05	12:05	13:05	14:05	14:35	15:05	15:35
2		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	T0000	0	12	33	18	19	14	11	7	15	16	22	26	30	28	17
4	L25630	0	12	33	18	19	14	11	7	15	16	22	26	30	28	17
5	L25650	0	13	35	18	20	15	12	7	15	18	26	26	35	30	22
6	L25670	0	12	34	16	20	13	11								22
7	L25690	0	12	34	16	21	12	12								22
8	L256110	0	13	34	15	21	12	12								21
9	L256130	0	14	36	16	20	12	12								21
10	C644270	0	15	39	16	20	13	13								26
11	L256170	0	16	41	16	20	13	13								29
12	L256190	0	15	39	17	20	14	12								33
13	L256210	0	15	39	18	20	14	12								31
14	L256230	0	14	40	18	20	14	12								31
15	L256250	0	14	40	18	19	14	12								30
16	L256270	0	9	28	11	13	11	8								30
17	L256290	0	9	28	11	13	11	8								28
18	L256310	0	9	26	11	13	11	7								28
19	L256330	0	9	25	10	12	11	7								28
20	L256350	0	9	25	10	12	11	7								29
21	L256370	0	9	24	10	11	11	6								29
22	R10810	0	0	24	10	11	11	6								28

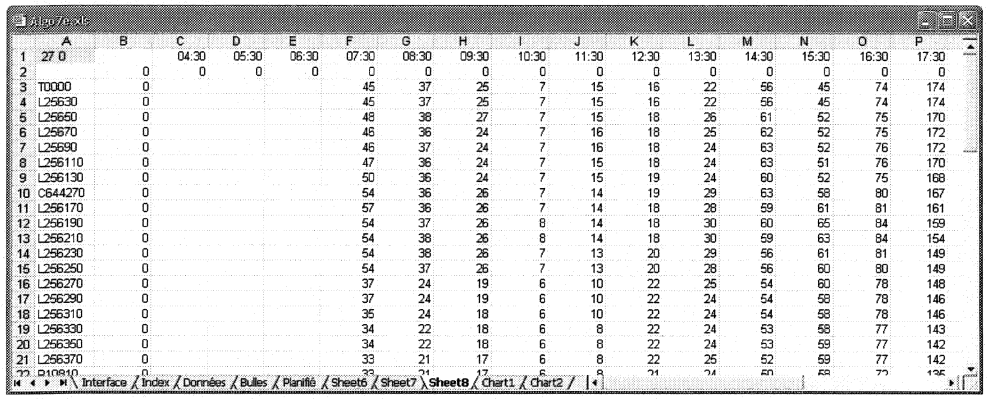
Interface / Index / Données / Bulles / Planifié / Sheet6 / Sheet7 / Sheet8 / Chart1 / Chart2 /

Feuille 7 : matrice de charge en fonction des arrêts et des tournées



Planifié / Sheet6 / Sheet7 / Sheet8 / Chart1 / Chart2 /

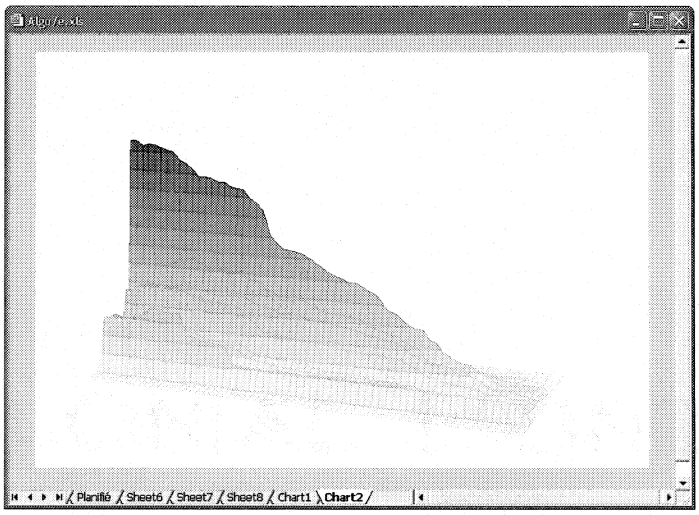
Graphique 1 : profil de charge 3D en fonction des tournées



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	27 0		04:30	05:30	06:30	07:30	08:30	09:30	10:30	11:30	12:30	13:30	14:30	15:30	16:30	17:30
2		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	T0000	0				45	37	25	7	15	16	22	56	45	74	174
4	L25630	0				45	37	25	7	15	16	22	56	45	74	174
5	L25650	0				48	38	27	7	15	18	26	61	52	75	170
6	L25670	0				46	36	24	7	16	18	25	62	52	75	172
7	L25690	0				46	37	24	7	16	18	24	63	52	76	172
8	L256110	0				47	36	24	7	15	18	24	63	51	76	170
9	L256130	0				50	36	24	7	15	19	24	60	52	75	168
10	C644270	0				54	36	26	7	14	19	29	63	58	80	167
11	L256170	0				57	36	26	7	14	18	28	59	61	81	161
12	L256190	0				54	37	26	8	14	18	30	60	65	84	159
13	L256210	0				54	38	26	8	14	18	30	59	63	84	154
14	L256230	0				54	38	26	7	13	20	29	56	61	81	149
15	L256250	0				54	37	26	7	13	20	28	56	60	80	149
16	L256270	0				37	24	19	6	10	22	25	54	60	78	148
17	L256290	0				37	24	19	6	10	22	24	54	58	78	146
18	L256310	0				35	24	18	6	10	22	24	54	58	78	146
19	L256330	0				34	22	18	6	8	22	24	53	58	77	143
20	L256350	0				34	22	18	6	8	22	24	53	59	77	142
21	L256370	0				33	21	17	6	8	22	25	52	59	77	142

Interface / Index / Données / Bulles / Planifié / Sheet6 / Sheet7 / **Sheet8** / Chart1 / Chart2 /

Feuille 8 : matrice de charge en fonction des arrêts et des intervalles d'une heure



Graphique 2 : profil de charge 3D en fonction du temps

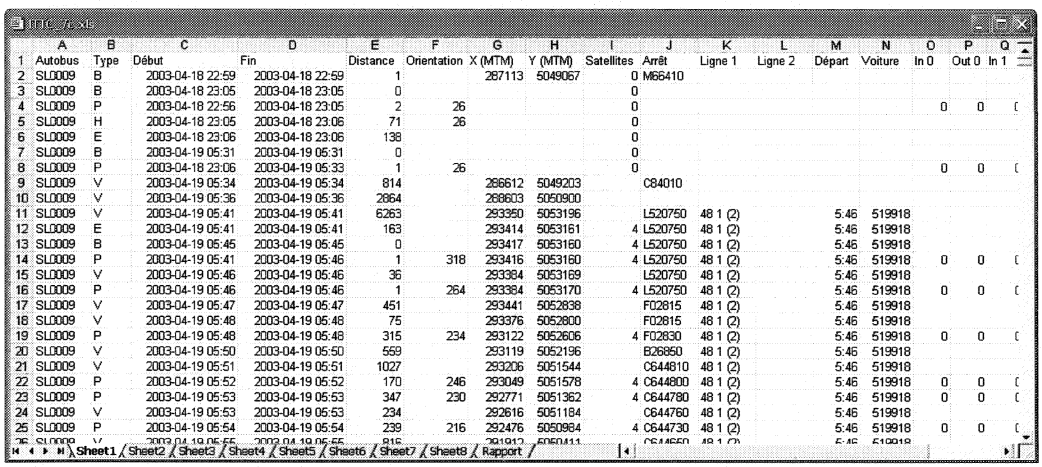
ANNEXE 3 : Identificateur de trajets en transport collectif

Logiciel

Le logiciel ITTC est une alternative aux logiciels BusStops et LADRA créée dans le cadre de ce projet de maîtrise afin de vérifier si le couplage avec les circuits et les horaires planifiés pouvait être amélioré.

ITTC a été conçu avec l'interface VBA de Microsoft Excel. Il fait la lecture des fichiers bruts un à un (il peut aussi lire les fichiers LADRA, mais ne tient pas compte des résultats déjà inclus dans ces fichiers) et effectue les opérations suivantes :

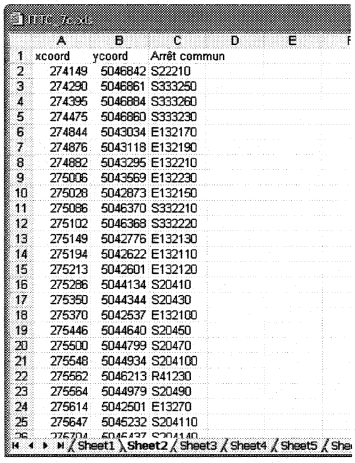
- Pour chaque événement enregistré, trouver l'arrêt le plus proche;
- Pour chaque événement, identifier les circuits passant par les arrêts trouvés;
- Vérifier chaque circuit un par un et éliminer les circuits non-conformes (section 7.5.3);
- En cas de circuits parallèles (plusieurs circuits possibles pour un même trajet), conserver le circuit le plus long et éliminer les autres circuits;
- Identifier les heures de départ et les numéros d'assignation (= numéros de *voiture*) de chaque circuit identifié;
- En cas de plusieurs tournées possibles pour un même circuit, comparer avec les assignations et conserver celui dont le numéro d'assignation possède le plus grand nombre d'occurrences.
- Compléter les données brutes avec les informations sur les tournées identifiées.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Autobus	Type	Début	Fin	Distance	Orientation	X (MTM)	Y (MTM)	Satellites	Arrêt	Ligne 1	Ligne 2	Départ	Voiture	In 0	Out 0	In 1
2	SL0009	B	2003-04-18 22:59	2003-04-18 22:59	1		267113	5049067	0	M66410							
3	SL0009	B	2003-04-18 23:05	2003-04-18 23:05	0				0								
4	SL0009	P	2003-04-18 22:56	2003-04-18 23:05	2	26			0						0	0	0
5	SL0009	H	2003-04-18 23:05	2003-04-18 23:06	71	26			0								
6	SL0009	E	2003-04-18 23:06	2003-04-18 23:06	138				0								
7	SL0009	B	2003-04-19 05:31	2003-04-19 05:31	0				0								
8	SL0009	P	2003-04-18 23:06	2003-04-19 05:33	1	26			0						0	0	0
9	SL0009	V	2003-04-19 05:34	2003-04-19 05:34	814		286612	5049203		C84010							
10	SL0009	V	2003-04-19 05:36	2003-04-19 05:36	2864		288603	5050900									
11	SL0009	V	2003-04-19 05:41	2003-04-19 05:41	6263		293350	5053196		L520750	48 1 (2)		5:46	519918			
12	SL0009	E	2003-04-19 05:41	2003-04-19 05:41	163		293414	5053161	4	L520750	48 1 (2)		5:46	519918			
13	SL0009	B	2003-04-19 05:45	2003-04-19 05:45	0		293417	5053160	4	L520750	48 1 (2)		5:46	519918			
14	SL0009	P	2003-04-19 05:41	2003-04-19 05:46	1	318	293416	5053160	4	L520750	48 1 (2)		5:46	519918	0	0	0
15	SL0009	V	2003-04-19 05:46	2003-04-19 05:46	36		293384	5053169		L520750	48 1 (2)		5:46	519918			
16	SL0009	P	2003-04-19 05:46	2003-04-19 05:46	1	264	293384	5053170	4	L520750	48 1 (2)		5:46	519918	0	0	0
17	SL0009	V	2003-04-19 05:47	2003-04-19 05:47	451		293441	5052838		F02815	48 1 (2)		5:46	519918			
18	SL0009	V	2003-04-19 05:48	2003-04-19 05:48	75		293376	5052800		F02815	48 1 (2)		5:46	519918			
19	SL0009	P	2003-04-19 05:48	2003-04-19 05:48	315	234	293122	5052606	4	F02830	48 1 (2)		5:46	519918	0	0	0
20	SL0009	V	2003-04-19 05:50	2003-04-19 05:50	559		293119	5052196		B26850	48 1 (2)		5:46	519918			
21	SL0009	V	2003-04-19 05:51	2003-04-19 05:51	1027		293206	5051544		C644810	48 1 (2)		5:46	519918			
22	SL0009	P	2003-04-19 05:52	2003-04-19 05:52	170	246	293049	5051578	4	C644800	48 1 (2)		5:46	519918	0	0	0
23	SL0009	P	2003-04-19 05:53	2003-04-19 05:53	347	230	292771	5051362	4	C644780	48 1 (2)		5:46	519918	0	0	0
24	SL0009	V	2003-04-19 05:53	2003-04-19 05:53	234		292616	5051184		C644760	48 1 (2)		5:46	519918			
25	SL0009	P	2003-04-19 05:54	2003-04-19 05:54	239	216	292476	5050984	4	C644730	48 1 (2)		5:46	519918	0	0	0

Sheet1 / Sheet2 / Sheet3 / Sheet4 / Sheet5 / Sheet6 / Sheet7 / Sheet8 / Rapport /

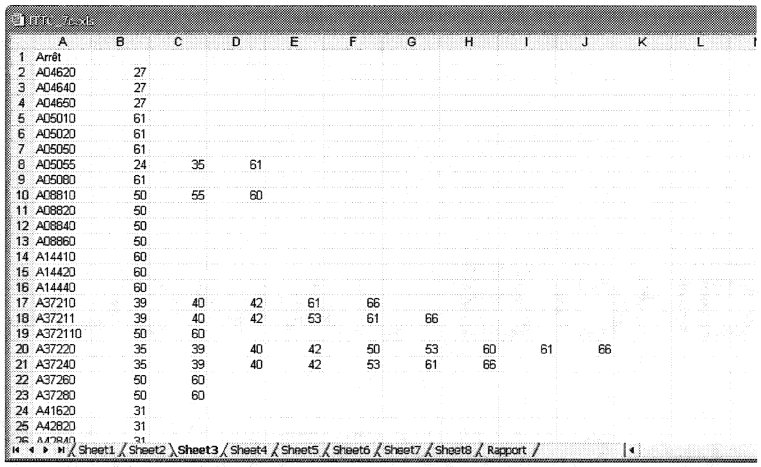
Feuille 1 : données Infodev complétées



	A	B	C	D	E
1	xcoord	ycoord	Arrêt commun		
2	274149	5046842	S22210		
3	274290	5046861	S333250		
4	274395	5046884	S333260		
5	274475	5046860	S333230		
6	274844	5043034	E132170		
7	274876	5043118	E132190		
8	274882	5043295	E132210		
9	275006	5043569	E132230		
10	275028	5042873	E132150		
11	275086	5046370	S332210		
12	275102	5046368	S332220		
13	275149	5042776	E132130		
14	275194	5042622	E132110		
15	275213	5042601	E132120		
16	275286	5044134	S20410		
17	275350	5044344	S20430		
18	275370	5042537	E132100		
19	275446	5044640	S20450		
20	275500	5044799	S20470		
21	275548	5044934	S204100		
22	275562	5046213	R41230		
23	275564	5044979	S20490		
24	275614	5042501	E13270		
25	275647	5045232	S204110		

Sheet1 / Sheet2 / Sheet3 / Sheet4 / Sheet5 / Sheet

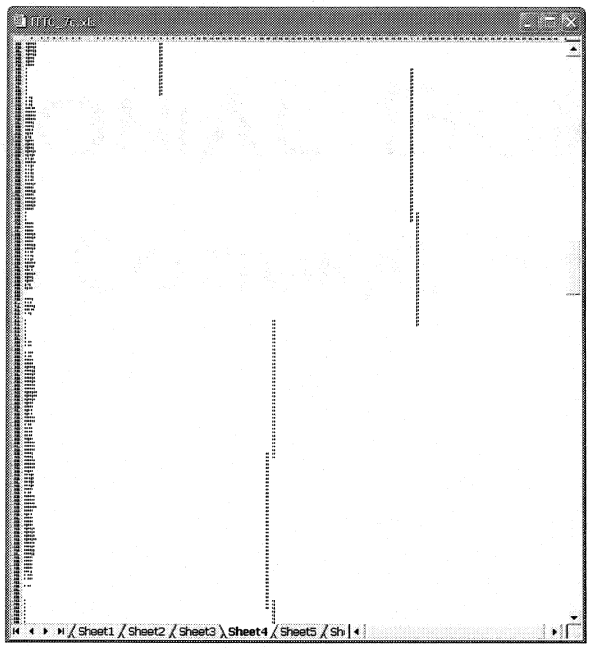
Feuille 2 : localisation théoriques des arrêts d'autobus



	A	B	C	D	E	F	G	H	I	J	K	L
1	Arrêt											
2	A04620	27										
3	A04640	27										
4	A04650	27										
5	A05010	61										
6	A05020	61										
7	A05050	61										
8	A05055	24	35	61								
9	A05080	61										
10	A08810	50	55	60								
11	A08820	50										
12	A08840	50										
13	A08860	50										
14	A14410	60										
15	A14420	60										
16	A14440	60										
17	A37210	39	40	42	61	66						
18	A37211	39	40	42	53	61	66					
19	A372110	50	60									
20	A37220	35	39	40	42	50	53	60	61	66		
21	A37240	35	39	40	42	53	61	66				
22	A37260	50	60									
23	A37280	50	60									
24	A41620	31										
25	A42820	31										

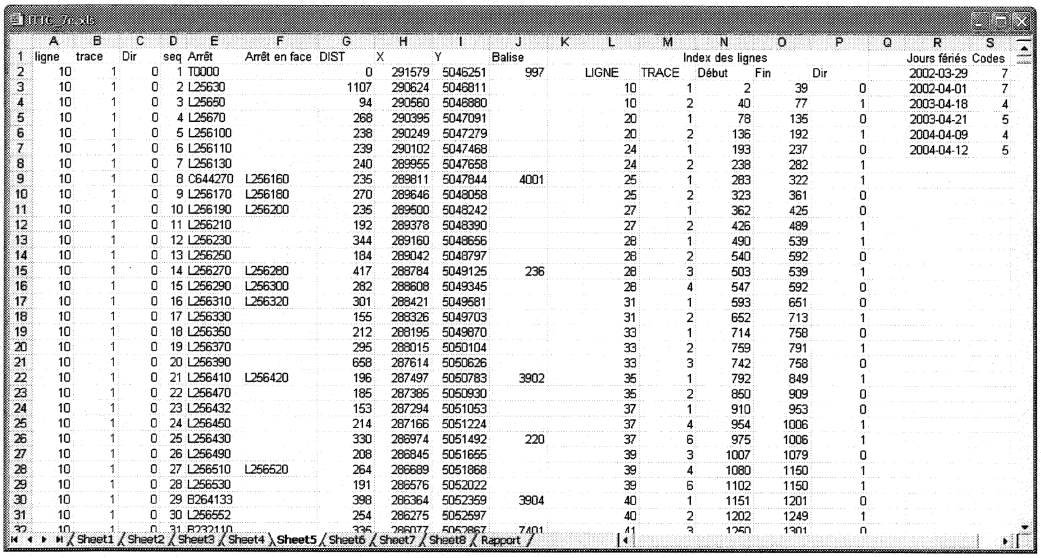
Sheet1 / Sheet2 / Sheet3 / Sheet4 / Sheet5 / Sheet6 / Sheet7 / Sheet8 / Rapport /

Feuille 3 : numéros de circuit pour chaque arrêt



Feuille 4 : matrice événement-circuit

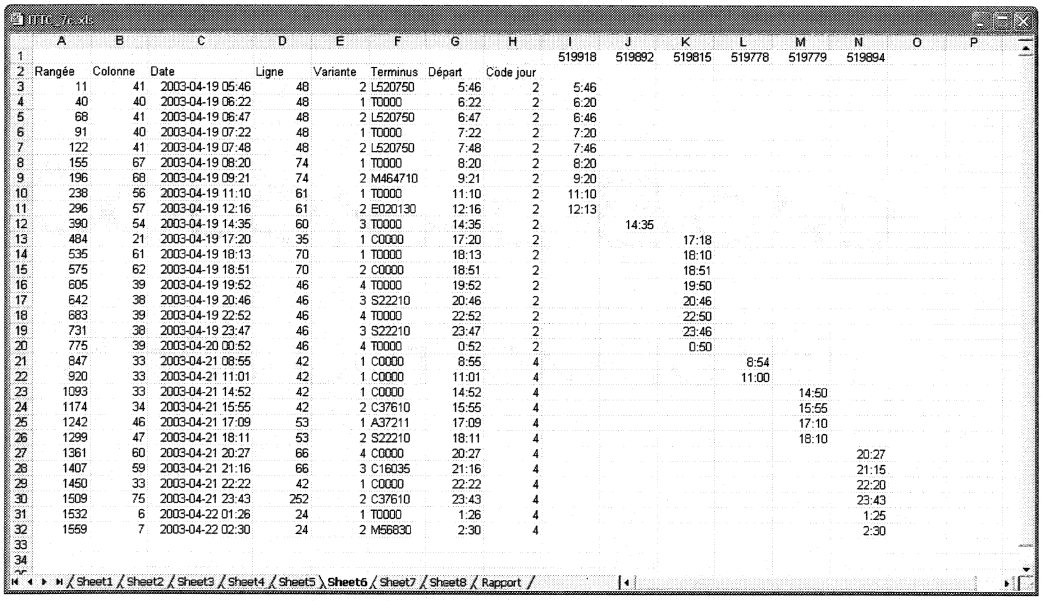
Pour chaque événement enregistré (rangée), le logiciel tente d'associer un circuit (colonne). Une rangée vide signifie qu'aucun circuit n'a été identifié. Plusieurs circuits se chevauchent aux terminus.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	ligne	trace	Dir	seq	Arrêt	Arrêt en face	DIST	X	Y	Balise			Index des lignes					Jours fériés	Codes
2	10	1	0	1	T0000		0	291579	5046251	997		LIGNE	TRACE	Début	Fin	Dir		2002-03-29	7
3	10	1	0	2	L25630		1107	290624	5046811			10	1	2	39	0		2002-04-01	7
4	10	1	0	3	L25650		94	290560	5046880			10	2	40	77	1		2003-04-18	4
5	10	1	0	4	L25670		268	290395	5047091			20	1	78	135	0		2003-04-21	5
6	10	1	0	5	L256100		238	290249	5047279			20	2	136	192	1		2004-04-09	4
7	10	1	0	6	L256110		239	290102	5047468			24	1	193	237	0		2004-04-12	5
8	10	1	0	7	L256130		240	289955	5047658			24	2	238	282	1			
9	10	1	0	8	C644270	L256160	235	289811	5047844	4001		25	1	283	322	1			
10	10	1	0	9	L256170	L256180	270	289646	5048058			25	2	323	361	0			
11	10	1	0	10	L256190	L256200	235	289500	5048242			27	1	362	425	0			
12	10	1	0	11	L256210		192	289378	5048390			27	2	426	489	1			
13	10	1	0	12	L256230		344	289160	5048656			28	1	490	539	1			
14	10	1	0	13	L256250		184	289042	5048797			28	2	540	592	0			
15	10	1	0	14	L256270	L256280	417	288784	5049125	236		28	3	503	539	1			
16	10	1	0	15	L256290	L256300	282	288608	5049345			28	4	547	592	0			
17	10	1	0	16	L256310	L256320	301	288421	5049581			31	1	593	651	0			
18	10	1	0	17	L256330		155	288326	5049703			31	2	652	713	1			
19	10	1	0	18	L256350		212	288195	5049870			33	1	714	758	0			
20	10	1	0	19	L256370		295	288015	5050104			33	2	759	791	1			
21	10	1	0	20	L256390		658	287614	5050626			33	3	742	758	0			
22	10	1	0	21	L256410	L256420	196	287497	5050783	3902		35	1	792	849	1			
23	10	1	0	22	L256470		185	287385	5050930			35	2	850	909	0			
24	10	1	0	23	L256432		153	287294	5051053			37	1	910	953	0			
25	10	1	0	24	L256450		214	287166	5051224			37	4	954	1006	1			
26	10	1	0	25	L256430		330	286974	5051492	220		37	6	975	1006	1			
27	10	1	0	26	L256490		208	288845	5051655			39	3	1007	1079	0			
28	10	1	0	27	L256510	L256520	264	286689	5051868			39	4	1080	1150	1			
29	10	1	0	28	L256530		191	286576	5052022			39	6	1102	1150	1			
30	10	1	0	29	B264133		398	286364	5052359	3904		40	1	1151	1201	0			
31	10	1	0	30	L256552		254	286275	5052597			40	2	1202	1249	1			
32	10	1	0	31	B232110		336	286077	5052867	7401		41	3	1250	1301	0			

Sheet1 / Sheet2 / Sheet3 / Sheet4 / **Sheet5** / Sheet6 / Sheet7 / Sheet8 / Rapport /

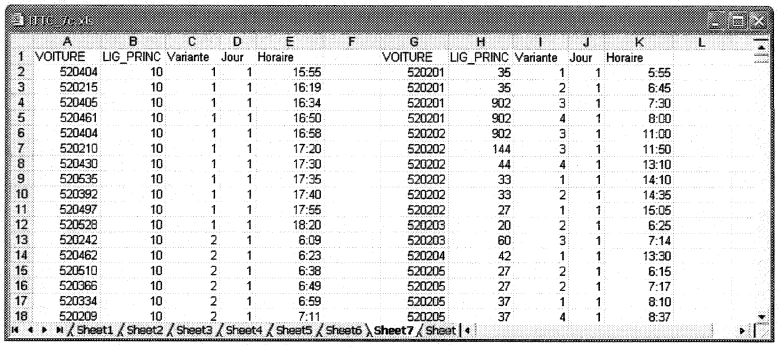
Feuille 5 : listes des circuits planifiés et des jours fériés



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1									519918	519892	519815	519778	519779	519894		
2	Rangée	Colonne	Date	Ligne	Variante	Terminus	Départ	Code jour								
3	11	41	2003-04-19 05:46	48	2	L520750	5:46	2	5:46							
4	40	40	2003-04-19 06:22	48	1	T0000	6:22	2	6:20							
5	68	41	2003-04-19 06:47	48	2	L520750	6:47	2	6:46							
6	91	40	2003-04-19 07:22	48	1	T0000	7:22	2	7:20							
7	122	41	2003-04-19 07:48	48	2	L520750	7:48	2	7:46							
8	155	67	2003-04-19 08:20	74	1	T0000	8:20	2	8:20							
9	196	68	2003-04-19 09:21	74	2	M464710	9:21	2	9:20							
10	238	56	2003-04-19 11:10	61	1	T0000	11:10	2	11:10							
11	296	57	2003-04-19 12:16	61	2	E020130	12:16	2	12:13							
12	390	54	2003-04-19 14:35	60	3	T0000	14:35	2		14:35						
13	484	21	2003-04-19 17:20	35	1	C0000	17:20	2			17:18					
14	535	61	2003-04-19 18:13	70	1	T0000	18:13	2			18:10					
15	575	62	2003-04-19 18:51	70	2	C0000	18:51	2			18:51					
16	605	39	2003-04-19 19:52	46	4	T0000	19:52	2			19:50					
17	642	38	2003-04-19 20:46	46	3	S22210	20:46	2			20:46					
18	683	39	2003-04-19 22:52	46	4	T0000	22:52	2			22:50					
19	731	38	2003-04-19 23:47	46	3	S22210	23:47	2			23:46					
20	775	39	2003-04-20 00:52	46	4	T0000	0:52	2			0:50					
21	847	33	2003-04-21 08:55	42	1	C0000	8:55	4				8:54				
22	920	33	2003-04-21 11:01	42	1	C0000	11:01	4				11:00				
23	1093	33	2003-04-21 14:52	42	1	C0000	14:52	4					14:50			
24	1174	34	2003-04-21 15:55	42	2	C37610	15:55	4					15:55			
25	1242	46	2003-04-21 17:09	53	1	A37211	17:09	4					17:10			
26	1299	47	2003-04-21 18:11	53	2	S22210	18:11	4					18:10			
27	1361	60	2003-04-21 20:27	66	4	C0000	20:27	4						20:27		
28	1407	59	2003-04-21 21:16	66	3	C16035	21:16	4						21:15		
29	1450	33	2003-04-21 22:22	42	1	C0000	22:22	4						22:20		
30	1509	75	2003-04-21 23:43	252	2	C37610	23:43	4						23:43		
31	1532	6	2003-04-22 01:26	24	1	T0000	1:26	4						1:25		
32	1559	7	2003-04-22 02:30	24	2	M56830	2:30	4						2:30		
33																
34																

Sheet1 / Sheet2 / Sheet3 / Sheet4 / Sheet5 / **Sheet6** / Sheet7 / Sheet8 / Rapport /

Feuille 6 : liste des circuits détectés avec assignations et heures de départ

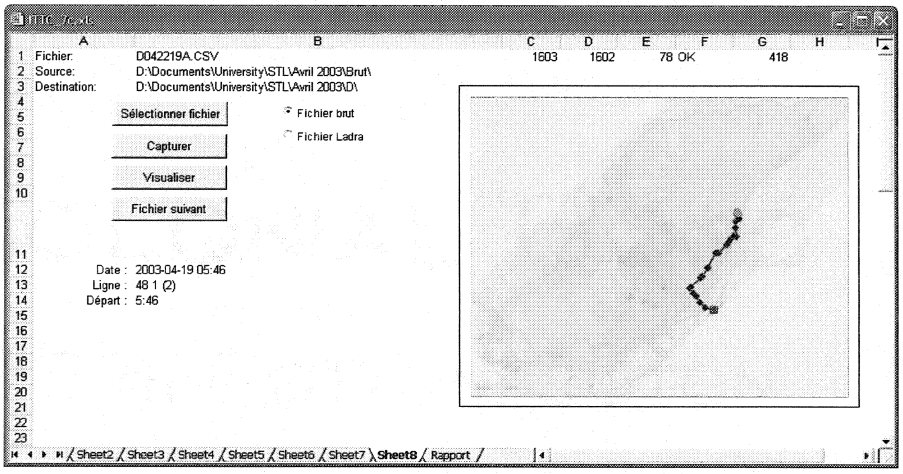


	A	B	C	D	E	F	G	H	I	J	K	L
1	VOITURE	LIG_PRINC	Variante	Jour	Horaire		VOITURE	LIG_PRINC	Variante	Jour	Horaire	
2	520404	10	1	1	15:55		520201	35	1	1	5:55	
3	520215	10	1	1	16:19		520201	35	2	1	6:45	
4	520405	10	1	1	16:34		520201	902	3	1	7:30	
5	520461	10	1	1	16:50		520201	902	4	1	8:00	
6	520404	10	1	1	16:58		520202	902	3	1	11:00	
7	520210	10	1	1	17:20		520202	144	3	1	11:50	
8	520430	10	1	1	17:30		520202	44	4	1	13:10	
9	520535	10	1	1	17:35		520202	33	1	1	14:10	
10	520392	10	1	1	17:40		520202	33	2	1	14:35	
11	520497	10	1	1	17:55		520202	27	1	1	15:05	
12	520528	10	1	1	18:20		520203	20	2	1	6:25	
13	520242	10	2	1	6:09		520203	60	3	1	7:14	
14	520462	10	2	1	6:23		520204	42	1	1	13:30	
15	520510	10	2	1	6:38		520205	27	2	1	6:15	
16	520366	10	2	1	6:49		520205	27	2	1	7:17	
17	520334	10	2	1	6:59		520205	37	1	1	8:10	
18	520209	10	2	1	7:11		520205	37	4	1	8:37	

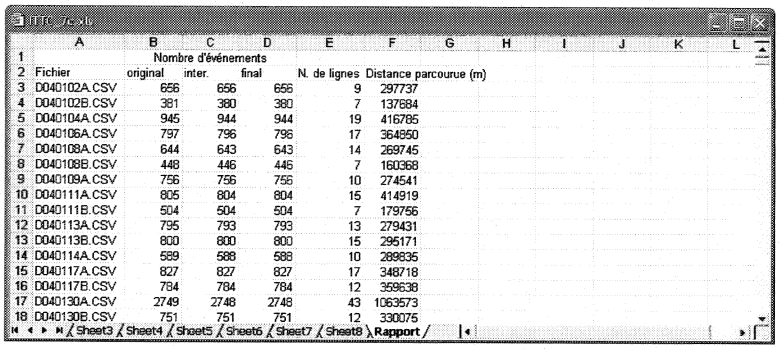
Sheet1 / Sheet2 / Sheet3 / Sheet4 / Sheet5 / Sheet6 \ **Sheet7** / Sheet

Feuille 7 : liste des horaires planifiés en ordre de circuit et en ordre d'assignation (*voiture*)

Les informations contenues dans les feuilles 2, 3 et 5 proviennent du fichier ligarr.dbf. Celles contenues dans la feuille 7 proviennent du fichier horaires.dbf. Ces deux fichiers en format Visual FoxPro ont été fournis gracieusement par la STL.



Feuille 8 : interface de visualisation



	A	B	C	D	E	F
1		Nombre d'événements				
2	Fichier	original	inter.	final	N. de lignes	Distance parcourue (m)
3	D040102A.CSV	656	656	656	9	297737
4	D040102B.CSV	381	380	380	7	137684
5	D040104A.CSV	945	944	944	19	416785
6	D040106A.CSV	797	796	796	17	364850
7	D040108A.CSV	644	643	643	14	269745
8	D040108B.CSV	448	446	446	7	160368
9	D040109A.CSV	756	756	756	10	274541
10	D040111A.CSV	805	804	804	15	414919
11	D040111B.CSV	504	504	504	7	179756
12	D040113A.CSV	795	793	793	13	279431
13	D040113B.CSV	800	800	800	15	295171
14	D040114A.CSV	589	588	588	10	289835
15	D040117A.CSV	827	827	827	17	348718
16	D040117B.CSV	784	784	784	12	359638
17	D040130A.CSV	2749	2748	2748	43	1063573
18	D040130B.CSV	751	751	751	12	330075

Sheet3 / Sheet4 / Sheet5 / Sheet6 / Sheet7 / Sheet8 \ Rapport /

Feuille 9 : rapport de production

Données

Chaque fichier brut est traité avec ITTC, lequel sauvegarde le résultat dans un autre fichier CSV. Ce nouveau fichier possède le même nom que le fichier brut d'origine auquel est ajouté le préfixe D.

Variables contenues dans les fichiers D

Variable	Signification
Autobus	Numéro de l'autobus
Type	Type d'événement (tous les événements du fichier original sauf v et X)
Début	Moment de début d'événement (date et heure à la seconde près)
Fin	Moment de fin d'événement (date et heure à la seconde près)
Distance	Distance parcourue depuis l'événement précédent (selon l'odomètre, au mètre près)
Orientation	Orientation du véhicule (0 à 359°)
X (MTM), Y (MTM)	Latitude, longitude en coordonnées MTM (NAD83)
Satellites	Qualité de lecture du GPS (≈ nombre de satellites détectés)
Arrêt	Numéro d'arrêt d'autobus
Ligne 1	Si l'arrêt est un terminus, alors circuit finissant. Sinon, circuit en cours (même format que dans BusStops et LADRA)
Ligne 2	Si l'arrêt est un terminus, alors circuit commençant. Sinon, rien. (À un terminus, les circuits finissant et commençant peuvent se chevaucher.)
Départ	Heure de départ planifiée
Voiture	Numéro de voiture
In 0, Out 0, In 1, Out 1, In 2, Out 2	Nombre de montées (In) et de descentes (Out) aux portes arrières (0 et 1) et à la porte avant (2)
Balise	Numéro de balise virtuelle