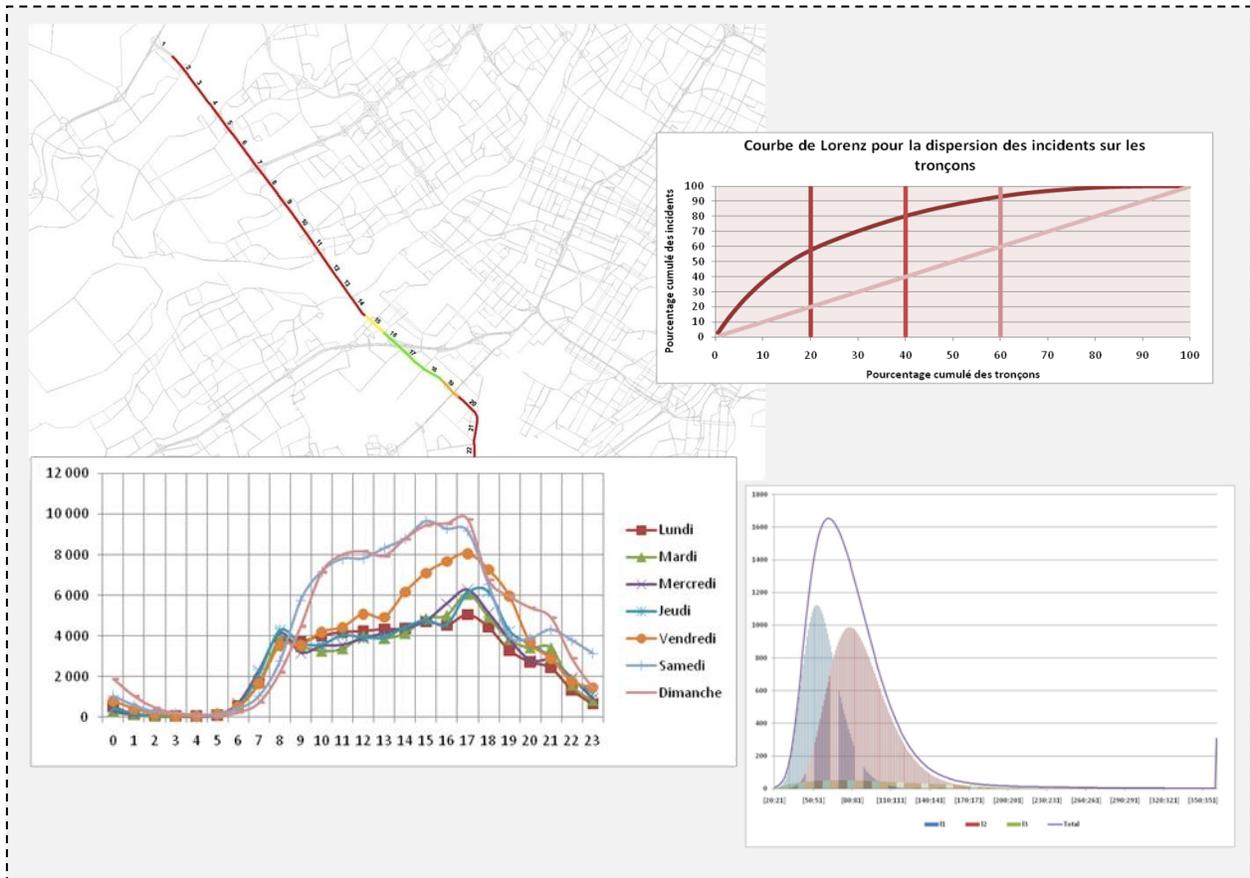


# Évaluation des enjeux analytiques et des potentialités des données de temps de parcours pour mesurer la fiabilité du réseau routier montréalais

## Rapport final (juin 2011)



# Évaluation des enjeux analytiques et des potentialités des données de temps de parcours pour mesurer la fiabilité du réseau routier montréalais

---

*Projet R.612, parrainé par*

*Service de la modélisation des systèmes de transport*

*Direction générale des politiques et de la sécurité en transport*

*et*

*Service de la planification et des stratégies d'intervention*

*Direction générale de Montréal et de l'Ouest*

*Transport Québec*

## Rapport final

---

*Pr. Catherine Morency*

*Pr. Martin Trépanier*

*Pr. Nicolas Saunier*

*École Polytechnique de Montréal*

### **Contributions de:**

*Hubert Verreault, associé de recherche*

*Pierre Loustau, gradué de Polytechnique (M.Sc.A)*

*Anouar Kalboussi, gradué de Polytechnique (M.Sc.A)*

*Yacine Lagha, stagiaire*

*Juin 2011*



# Table des matières

---

1.	Sommaire .....	1
1.1	Contenu du rapport .....	1
1.2	Contributions .....	1
2.	Modélisation des distributions fréquentielles des temps de parcours : synthèse .....	3
2.1	Analyse exploratoire d'un circuit .....	6
2.1.1	Présentation du circuit 1 .....	6
2.1.2	Localisation des points chauds du circuit 1 .....	6
2.1.3	Évolution temporelle des zones de congestion .....	8
2.1.4	Observation des temps de parcours selon le mois et l'année .....	10
2.1.5	Évolution d'indicateurs de fiabilité selon l'heure .....	11
2.1.6	Observations selon les conditions de la chaussée .....	12
2.1.7	Synthèse .....	14
2.2	Définition d'indicateurs et illustration .....	15
2.2.1	Indicateur de circulation anormale .....	15
2.2.2	Indicateur Moyenne-Variabilité (IMV) .....	15
2.2.3	Présentation de l'indicateur .....	16
2.2.4	Résultats : illustration .....	17
2.3	Analyse mathématique des circuits du réseau autoroutier montréalais. ....	21
2.3.1	Méthode de groupement par similarité .....	21
2.3.2	Résultats .....	22
2.3.3	Modélisation des distributions .....	25
3.	Étude des potentialités des différentes technologies de collecte de données .....	46
3.1	Sources de données et analyse descriptive .....	46
3.1.1	Données GPS des véhicules de Communauto inc. ....	46
3.1.2	Données de vitesses moyennes par segment RealTraffic .....	57
3.1.3	Données de temps de parcours INRIX .....	58
3.2	Mesure des conditions de circulation .....	61
3.2.1	Indicateurs .....	61
3.2.2	Données GPS .....	61
3.2.3	Données de vitesses moyennes par segment Real Traffic .....	65
3.2.4	Données de temps de parcours INRIX .....	67
3.3	Méthodologie d'analyse comparative .....	69
3.3.1	Définition des indicateurs .....	69

3.3.2	Comparaison des vitesses moyennes .....	70
3.3.3	Comparaison des distributions .....	72
3.4	Discussion.....	76
3.4.1	Synthèse.....	77
3.4.2	Limitations.....	77
3.4.3	Recommandations .....	77
4.	Analyse des segments routiers en fonction de leurs caractéristiques physiques .....	79
4.1.1	Montage de la base de données.....	80
4.1.2	Analyse descriptive du réseau routier .....	80
4.1.3	Analyse factorielle de correspondances multiples .....	82
4.1.4	Groupements FP .....	84
4.1.5	Confrontation des résultats (FP vs TP).....	84
4.1.6	Prédiction de l'appartenance.....	87
5.	Exploration des données d'incidents.....	88
5.1	Base de données .....	88
5.1.1	Concepts.....	88
5.1.2	Variables.....	88
5.2	Analyse descriptive .....	89
5.2.1	Distribution des incidents .....	89
5.2.2	Analyse de la durée d'incident.....	93
5.3	Analyse classifiée .....	94
5.4	Étude des incidents de type accidents.....	95
5.4.1	État des connaissances .....	95
5.5	Analyse des données d'accidents .....	96
5.5.1	Dispersion des accidents : courbe de Lorenz.....	97
5.6	Modélisation de survie d'un accident.....	97
5.6.1	Filtrage et transformation de la base de données.....	97
5.6.2	Analyse deltaT à travers l'effet d'une variable .....	99
5.6.3	Modèle de Kaplan et estimation des « probabilités de survie » d'incidents à travers le temps	100
5.6.4	Modèle de Cox : Rapports de risques entre les caractéristiques physiques des tronçons	103
5.6.5	Conclusion et perspectives .....	106
6.	Conclusion générale.....	107
6.1	Modélisation des temps de parcours .....	107

6.2	Comparaison des technologies de collecte de données.....	107
6.3	Facteurs physiques.....	108
6.4	Banque de données d'incidents.....	108
7.	Références complémentaires .....	109
8.	Annexes.....	110
8.1	Analyse statistique du circuit 1 au moyen de Statistica .....	110
8.1.1	Statistiques descriptives .....	110
8.1.2	Tables de fréquence.....	111
8.1.3	Moyennes par tronçon .....	112
8.1.4	Analyse de la variance selon l'année .....	115
8.1.5	Analyse de la variance selon le mois.....	120
8.1.6	Analyse de la variance selon les conditions météorologiques .....	125
8.1.7	Temps de parcours avec mois et année.....	128
8.2	Tableaux.....	135

## Liste des figures

---

Figure 1. Synthèse de la méthode de préparation et d'analyse des données de véhicules flottants .....	4
Figure 2. Démarche scientifique pour fins de modélisation des distributions de temps de parcours sur les tronçons d'un kilomètre .....	5
Figure 3. Circuit 1, direction et intersections majeures.....	6
Figure 4. Évolution des temps de parcours par tronçons du circuit 1 .....	7
Figure 5. Moyenne des temps de parcours du circuit 1 selon la période.....	7
Figure 6. Évolution temporelle du temps de parcours moyen par tronçon entre 6h00 et 8h20 AM (abscisse : numéro du tronçon, ordonnée : temps de parcours par kilomètre, en secondes).....	9
Figure 7. Moyenne des temps de parcours selon le mois .....	10
Figure 8. Évolution des temps de parcours pour les mois de février, juin et septembre 2004.....	11
Figure 9. Observations des TRI et TTI selon l'heure.....	12
Figure 10. Pourcentage de relevés rencontrés par condition de la chaussée, par année, pour le circuit 1 .....	13
Figure 11. Temps de parcours moyens par condition de la chaussée et par année pour le circuit 1 .....	13
Figure 12. Temps de parcours moyen selon la condition de la route pour le circuit 1 .....	14
Figure 13. Indicateurs de circulation anormale pour le circuit 1 .....	15
Figure 14. Représentation des 4 différentes zones de l'IMV .....	17
Figure 15. Valeurs de l'IMV par tronçon du circuit 1 AM.....	18
Figure 16. Valeurs de l'IMV par tronçon du circuit 1 PM.....	18
Figure 17. Indicateur du circuit 1 période AM, 8h00 .....	19
Figure 18. Indicateur du circuit 1 période AM, 8h10.....	20
Figure 19. Indicateur du circuit 1 période AM, 8h20 .....	20
Figure 20 Distribution fréquentielle de quelques groupes de tronçons pour la période AM .....	23
Figure 21. Répartition des tronçons selon les groupes avec moyenne, période AM .....	24
Figure 22. Répartition des tronçons selon les groupes avec moyenne, période PM .....	25
Figure 23. Distribution fréquentielle des temps de parcours pour le groupe AM-2 .....	26
Figure 24. Distribution fréquentielle des temps de parcours du groupe AM-6.....	28
Figure 25. Distribution fréquentielle des relevés de temps de parcours du groupe AM-8.....	29
Figure 26. Distribution du nombre d'observations disponibles par mois et par type de jour et comparaison des échantillons de semaine par rapport à l'échantillon de véhicules flottants de 2004 ....	48
Figure 27. Distribution temporelle du nombre d'observations disponibles par heure et par type de jour, tous mois confondus.....	49
Figure 28. Distribution fréquentielle du nombre d'observations de semaine par tronçon, pour les différents mois de 2010 .....	50
Figure 29. Distribution des tronçons (en %) en fonction du nombre d'observations disponibles par semaine (tous mois confondus).....	51
Figure 30. Comparaison des échantillons de données GPS et VF disponibles par mois, le lundi.....	52
Figure 31. Comparaison des échantillons de données GPS et VF disponibles par mois, le mardi .....	53
Figure 32. Comparaison des échantillons de données GPS et VF disponibles par mois, le mercredi .....	53
Figure 33. Comparaison des échantillons de données GPS et VF disponibles par mois, le jeudi.....	54
Figure 34. Comparaison des échantillons de données GPS et VF disponibles par mois, le vendredi.....	54
Figure 35. Comparaison des distributions temporelles des échantillons de données (jours de semaine) GPS et VF, mois de février.....	55
Figure 36. Comparaison des distributions temporelles des échantillons de données (jours de semaine) GPS et VF, mois de mai .....	55

Figure 37. Comparaison des distributions temporelles des échantillons de données (jours de semaine) GPS et VF, mois de septembre.....	56
Figure 38. Comparaison des tronçons de 1 km et des segments Real Traffic .....	58
Figure 39. Comparaison des tronçons de 1 km et des segments INRIX.....	60
Figure 40. Comparaison des segments routiers pour l'autoroute métropolitaine (A40) .....	60
Figure 41. Variation de la vitesse moyenne (tous jours de semaine confondus, excluant samedi et dimanche) sur l'autoroute 40 (direction EST) lors des différents mois. ....	62
Figure 42. Variation de la vitesse moyenne (tous jours de semaine confondus, excluant samedi et dimanche, tous mois confondus) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée.....	62
Figure 43. Variation de la vitesse moyenne (tous mois confondus) sur l'autoroute 40 (direction EST) lors des différents jours de la semaine.....	63
Figure 44. Variation de la vitesse moyenne (tous mois confondus, jours de la semaine excluant samedi et dimanche) sur les différents tronçons du réseau (sous-ensemble de directions, tel que représenté par les flèches), pour différentes périodes.....	64
Figure 45. Variation de la vitesse moyenne (tous jours de semaine confondus, août 2010) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée.....	65
Figure 46. Variation de la vitesse moyenne (tous jours de semaine confondus, tous mois confondus) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée .....	65
Figure 47. Variation de la vitesse moyenne (tous mois confondus, jours de la semaine excluant samedi et dimanche) sur les différents tronçons du réseau (sous-ensemble de directions, tel que représenté par les flèches), pour différentes périodes selon les données de Real Traffic.....	66
Figure 48. Variation de la vitesse moyenne (tous jours de semaine confondus, août 2010) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée.....	67
Figure 49. Variation de la vitesse moyenne (tous jours de semaine confondus, tous mois confondus) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée .....	67
Figure 50 : Distribution cumulée des tronçons selon l'amplitude des vitesses observées sur chacun d'eux pendant un jour de semaine.....	70
Figure 51 : Comparaison des vitesses observées entre les GPS de Communauto et INRIX par tronçon et par groupe d'heures (Jour de semaines, août 2010).....	71
Figure 52 : Comparaison des vitesses observées entre les GPS de Communauto et Real Trafic par tronçon et par groupe d'heures (Jour de semaines, août 2010).....	71
Figure 53 : Distribution des observations de vitesses GPS de Communauto (Tous les tronçons échantillonnés, jours de semaine seulement).....	72
Figure 54 : Distribution des observations de vitesses INRIX (Tous les tronçons échantillonnés, jours de semaine seulement).....	72
Figure 55 : Distribution des observations de vitesses Real Trafic (Tous les tronçons échantillonnés, jours de semaine seulement).....	73
Figure 56 : Position géographique du tronçon 34 de l'autoroute 40 en direction est. ....	73
Figure 57 : Distribution horaire des vitesses moyennes par heures sur le tronçon 34 en direction est (Jours de semaine, août 2010).....	74
Figure 58 : Distribution horaire des vitesses moyennes par heures sur le tronçon 34 en direction est (Jours de fin de semaine, août 2010).....	74
Figure 59 : Distribution de fréquence des vitesses observées par Communauto du tronçon 34 en direction est en pointe de l'après-midi (tous jours de semaine, août 2010) .....	75
Figure 60 : Distribution de fréquence des vitesses observées par Real Trafic du tronçon 34 en direction est en pointe de l'après-midi (tous jours de semaine, août 2010).....	75

Figure 61 : Distribution de fréquence des vitesses observées par INRIX du tronçon 34 en direction est en pointe de l'après-midi (tous jours de semaine, août 2010).....	76
Figure 82: Méthodologie de caractérisation des segments routiers selon leurs caractéristiques physiques .....	79
Figure 83: Vitesses autorisées par segment, selon leur groupe TP .....	81
Figure 84: Nombre d'intersections par segment, selon leur groupe TP .....	82
Figure 85: Nombre de voies par segment, selon leur groupe TP.....	82
Figure 86: Valeur propre, pourcentage d'inertie, pourcentage cumulé expliquée par les axes factoriels	83
Figure 87: Coordonnées des modalités des facteurs physiques.....	83
Figure 88: Dispersion intragroupe (w) et intergroupe (BSS) .....	84
Figure 89: Carte thématique représentant la contribution au khi-deux .....	86
Figure 90: Arbre de décision C4.5 complet .....	87
Figure 50. Distribution de l'ensemble des incidents entre 2002 et 2004, par mois de l'année .....	90
Figure 51. Répartition des incidents selon le jour de la semaine (données 2001-2004).....	90
Figure 52. Répartition des incidents selon l'heure de la journée .....	91
Figure 53. Répartition des incidents par axe routier (2001 à 2004) .....	92
Figure 54. Répartition des incidents : autoroute 40.....	92
Figure 55. Dispersion des incidents sur le réseau : courbe de Lorenz.....	93
Figure 56. Distribution fréquentielle des incidents selon leur durée .....	93
Figure 57. Occurrence des cinq types d'incidents les plus fréquents.....	94
Figure 58. Évolution du nombre d'incidents par année et de la durée d'incident moyenne.....	95
Figure 59. Cartographie des accidents sur le réseau autoroutier de Montréal entre janvier 2002 et décembre 2004 .....	96
Figure 60. Courbe de Lorenz pour les accidents.....	97
Figure 61. Courbe de survie pour la durée sans accident.....	101
Figure 62. Courbe de survie pour la durée sans accident, stratifiée par mois de l'année .....	102
Figure 63. Courbe de survie pour la durée sans accident, stratifiée par hiver/autres saisons .....	103
Figure 64. Histogramme des temps de parcours.....	111
Figure 65. Moyennes des temps de parcours selon le tronçon.....	113
Figure 66. Moyenne des temps de parcours selon la période.....	113
Figure 67. Moyennes des temps de parcours selon le tronçon et la période .....	114
Figure 68. Moyennes des temps de parcours selon l'année .....	116
Figure 69. Résidus sur échelle normale .....	117
Figure 70. Valeurs prédites vs résidus .....	118
Figure 71. Résidus vs Résidus exclus.....	118
Figure 72. Moyennes selon année et période .....	120
Figure 73. Moyennes des temps de parcours selon le mois.....	121
Figure 74. Résidus sur échelle normale .....	123
Figure 75. Résidus vs résidus exclus.....	123
Figure 76. Moyennes des temps de parcours selon le mois et la période .....	124
Figure 77. Moyennes des temps de parcours selon les conditions météo.....	125
Figure 78. Résidus sur échelle normale .....	127
Figure 79. Résidus vs résidus exclus.....	127
Figure 80. Moyennes selon mois par année .....	130
Figure 81. Moyennes et intervalles de confiance Mois(année*période) .....	133

## Liste des tableaux

---

Tableau 1. Valeurs caractéristiques des groupes AM.....	23
Tableau 2. Valeurs caractéristiques des groupes PM.....	24
Tableau 3. Paramètre du modèle de distribution fréquentielle pour les cas de « Pic étroit » .....	26
Tableau 4. Moyennes et intervalles de confiance des trois lois définissant le groupe AM-2 .....	27
Tableau 5. Définition des intervalles du groupe AM-2 .....	27
Tableau 6. Paramètres du modèle du groupe AM-6 .....	28
Tableau 7. Définition des intervalles du groupe AM-6 .....	28
Tableau 8. Définition des intervalles du groupe AM-8 .....	29
Tableau 9. Paramètres du modèle du groupe AM-8 .....	30
Tableau 10. Paramètres du modèle de distribution des temps de parcours AM-1 .....	30
Tableau 11. Paramètres du modèle de distribution des temps de parcours AM-2 .....	31
Tableau 12. Paramètres du modèle de distribution des temps de parcours AM-3 .....	32
Tableau 13. Paramètres du modèle de distribution des temps de parcours AM-4 .....	33
Tableau 14. Paramètres du modèle de distribution des temps de parcours AM-5 .....	34
Tableau 15. Paramètres du modèle de distribution des temps de parcours AM-6 .....	35
Tableau 16. Paramètres du modèle de distribution des temps de parcours AM-7 .....	36
Tableau 17. Paramètres du modèle de distribution des temps de parcours AM-8 .....	37
Tableau 18. Paramètres du modèle de distribution des temps de parcours PM-1.....	38
Tableau 19. Paramètres du modèle de distribution des temps de parcours PM-2.....	39
Tableau 20. Paramètres du modèle de distribution des temps de parcours PM-3.....	40
Tableau 21. Paramètres du modèle de distribution des temps de parcours PM-4.....	41
Tableau 22. Paramètres du modèle de distribution des temps de parcours PM-5.....	42
Tableau 23. Paramètres du modèle de distribution des temps de parcours PM-6.....	43
Tableau 24. Paramètres du modèle de distribution des temps de parcours PM-7.....	44
Tableau 25. Paramètres du modèle de distribution des temps de parcours PM-8.....	45
Tableau 26. Nombre d'observations par type de jour et période de la journée.....	47
Tableau 27. Nombre d'observations par type de jour et période de la journée.....	49
Tableau 28. Nombre de relevés effectués par véhicules flottants en 2004 .....	52
Tableau 29. Nombre de relevés effectués par véhicules flottants en 2004 .....	56
Tableau 30. Caractérisation des distributions pour le tronçon 34 en pointe de l'après-midi.....	76
Tableau 66: Contenu de la table de données sur les segments .....	80
Tableau 67. Valeurs caractéristiques des groupes AM.....	81
Tableau 68: Répartition des tronçons de chaque groupe FP selon les groupes TP .....	85
Tableau 70: Cartogramme des contributions au khi-deux .....	85
Tableau 32. Contenu de la base de données.....	88
Tableau 33. Durée moyenne des incidents par type .....	96
Tableau 34. Variables de caractérisation physiques des tronçons .....	98
Tableau 35. Variables décrivant les conditions météorologiques .....	99
Tableau 36. Tableau des estimations de survie de Kaplan-Meier .....	100
Tableau 37. Sortie SAS – test d'hypothèse BETA=0 .....	104
Tableau 38. Rapports de risques issus de la procédure.....	105
Tableau 39. Statistiques descriptives: temps de parcours circuit 1.....	110
Tableau 40. Fréquence des observations par année .....	111
Tableau 41. Fréquence des observations par mois .....	112
Tableau 42. ANOVA pour le tronçon et la période plus effet d'interaction .....	114

Tableau 43. ANOVA selon l'année .....	115
Tableau 44. Coefficients de la régression selon l'année .....	115
Tableau 45. Moyennes et intervalles de confiances selon l'année .....	116
Tableau 46. Calcul des coefficients de signification.....	117
Tableau 47. Test de LSD selon l'année.....	117
Tableau 48. Coefficients de corrélation .....	119
Tableau 49. ANOVA selon la période et l'année avec effet d'interaction .....	119
Tableau 50. ANOVA selon le mois.....	120
Tableau 51. Calcul du coefficient de signification.....	121
Tableau 52. Coefficients et intervalle de confiance par mois.....	122
Tableau 53. Test LSD sur les moyennes par mois .....	122
Tableau 54. Coefficients de corrélation .....	124
Tableau 55. ANOVA selon les conditions météo .....	125
Tableau 56. Coefficients de la régression et intervalles de confiance.....	126
Tableau 57. Calcul des coefficients de signification.....	126
Tableau 58. Test LSD .....	126
Tableau 59. Coefficients de corrélation .....	128
Tableau 60. Moyennes selon Mois par Année.....	128
Tableau 61. ANOVA Nested design .....	130
Tableau 62. Moyenne et intervalle de confiance mois(année*periode).....	131
Tableau 63. ANOVA Mois(Année*Période) .....	134
Tableau 64. Calcul du coefficient de signification.....	134
Tableau 65. ANOVA des temps de parcours selon Période et tronçons.....	135
Tableau 66. ANOVA Nested Design pour la variable Mois(Année).....	136
Tableau 67. ANOVA selon conditions de la chaussée.....	136

# 1. Sommaire

---

Ce rapport résume les différentes contributions réalisées dans le cadre du projet de recherche sur l'évaluation des enjeux analytiques et des potentialités des données de temps de parcours pour mesurer la fiabilité du réseau routier montréalais. Tel qu'il sera observé dans les pages qui suivent, le projet a pris des directions de recherche multiples, toutes liées à la mesure des temps de parcours sur les réseaux autoroutiers mais pas nécessairement en lien avec les données issues de véhicules flottants, base de données à l'origine du projet.

Le projet de recherche s'est, grosso modo, articulé autour de cinq grandes phases :

- la recension des écrits et la compréhension des principaux concepts et indicateurs liés à la mesure des temps de parcours et de leur fiabilité;
- le traitement, la validation et l'analyse des bases de données de véhicules flottants, étape qui a exigé la formulation d'une nouvelle méthodologie d'analyse basée sur la segmentation du réseau routier en tronçons d'un kilomètre;
- l'accompagnement, lors de la préparation du plan d'échantillonnage, pour la réalisation d'une nouvelle opération de collecte de données avec des véhicules flottants;
- la classification des tronçons autoroutiers sur la base de la similarité des distributions fréquentielles des temps de parcours et modélisation des distributions pour différents groupes de tronçons (16 groupes différents);
- l'étude des potentialités de différentes technologies de collecte de données sur les temps de parcours et/ou vitesses, comparaison des mesures issues de ces différentes technologies et formulation de recommandations.

D'autres thématiques ont été abordées en parallèle.

## 1.1 [Contenu du rapport](#)

Le rapport insistera sur deux composantes : d'une part, la description du processus de modélisation développé pour reproduire les distributions fréquentielles de groupe de segments autoroutiers et d'autre part, la présentation des différentes analyses comparatives visant à évaluer les potentialités des différentes technologies de collecte de données. Ces analyses comparatives s'appuient sur les données GPS issues des véhicules de Commuauto, un sous-ensemble de données INRIX ainsi qu'un sous-ensemble de données RealTraffic. Aussi, le développement d'une base de données de facteurs physiques décrivant les tronçons ainsi que diverses analyses en découlant sont présentés. En outre, puisque des données d'incidents ont aussi été rendu disponibles, certaines perspectives de modélisation à l'aide de ces données sont aussi présentées. Le rapport se termine par une discussion ainsi que des recommandations.

## 1.2 [Contributions](#)

Ce projet de recherche a mené à la publication de différents travaux de recherche, impliquant différents étudiants. Ces différents documents seront fournis sous forme électronique au MTQ.

- Loustau, P. (2009). Modélisation des temps de parcours sur un réseau routier à l'aide de données de véhicules flottants, mémoire de maîtrise (M.Sc.A.), département de mathématiques et génie industriel, École Polytechnique de Montréal, 215 pages.

- Loustau, P., Morency, C., Trépanier, M. (2010). *Measuring, Describing and Modeling Travel Time Reliability*, présenté au 89th Annual Meeting of the Transportation Research Board, Washington D.C.
- Loustau, P., Morency, C., Trépanier, M., Gourvil, L. (2010). *Travel time reliability on a highway network: estimations using floating car data*, *Transportation Letters*, Vol. 2(1), pp. 27-37, 2010
- Loustau, P., Grasset, V., Morency, C., Trépanier, M. *Comparing floating car data and carsharing GPS data for travel time assessment*, présenté au 12th World Conference on Transport Research, Lisbonne, Portugal.
- Loustau, P., Morency, C., Trépanier, M., Gourvil, L. (2009). *Portrait de la fiabilité des temps de parcours sur le réseau autoroutier montréalais*, présenté au 44<sup>e</sup> Congrès annuel de l'Association québécoise du transport et des routes, Montréal.
- Loustau, P., Grasset, V., Morency, C., Trépanier, M. (2009). *Véhicules flottants vs GPS : analyse comparative des temps de parcours sur le réseau routier supérieur montréalais*, présenté à la CITE Conference, Montréal
- Saunier, N., Morency, C. (2011). *Comparing data from mobile and static traffic sensors for travel time assessment*, 1st Conference on Integrated Transportation & Development for a Better Tomorrow, Chicago
- Morency, C., Trépanier, M., Saunier, N., Akélaguélo, J.-P. (2011). *Enhancing the value of an incidents database with an interactive visualization tool*, 1st Conference on Integrated Transportation & Development for a Better Tomorrow, Chicago
- Verreault, H., Morency, C., Saunier, N. (2011). *Des espions GPS au service de la gestion des réseaux*, présenté aux Dixièmes rencontres francophones est-ouest en socio-économie des transports, École Polytechnique de Montréal, Montréal, Québec, Canada

## 2. Modélisation des distributions fréquentielles des temps de parcours : synthèse

---

Différents travaux d'analyse avaient déjà été tentés à l'aide des bases de données de temps de parcours recueillies par le MTQ via des véhicules flottants sans toutefois permettre de tirer des conclusions satisfaisantes. Dans ce contexte, nous avons choisi de développer une autre approche s'appuyant sur un découpage des circuits et relevés en tronçons unitaires d'un kilomètre. Cette approche d'uniformisation permet de comparer les temps de parcours sur les différents tronçons ainsi que de les classer pour fins d'analyse.

La méthodologie de préparation et d'analyse exploratoire des données déployée à l'amorce du projet est rappelée ci-dessous, sous forme schématique (Figure 1). Les principales étapes de cette méthodologie ont été décrites dans le premier rapport d'étape.

La démarche scientifique qui a permis de mener à la modélisation des distributions de temps de parcours sur les tronçons d'un kilomètre est aussi présentée sous forme schématique, à la Figure 2.

Les sections qui suivent détaillent la méthodologie de modélisation développée. Les étapes suivantes sont couvertes :

- Analyse exploratoire d'un circuit et évolution d'indicateurs de fiabilité;
- Définition d'indicateurs;
- Classification des tronçons par similarité des distributions fréquentielles;
- Synthèse des résultats.

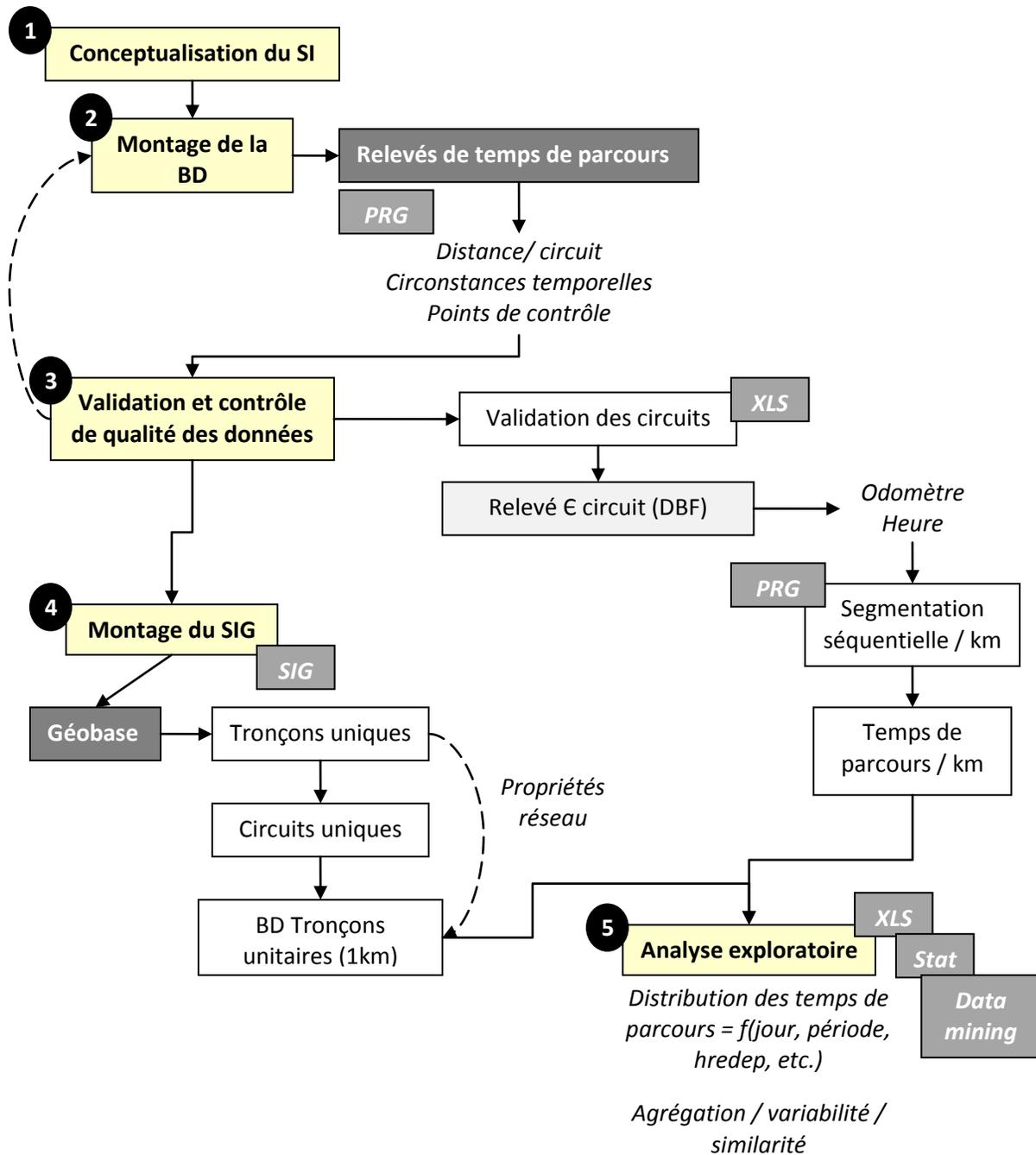


Figure 1. Synthèse de la méthode de préparation et d'analyse des données de véhicules flottants

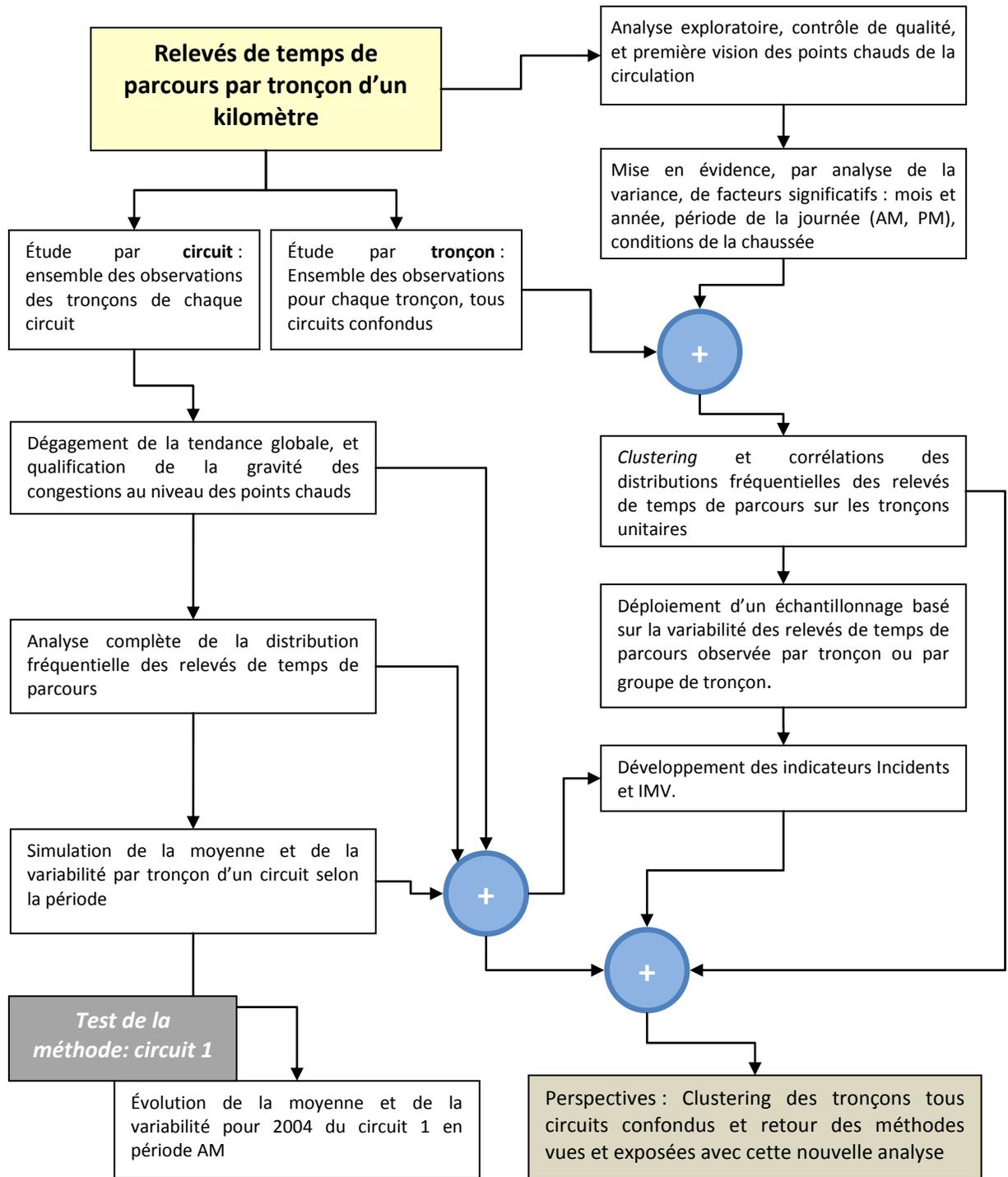


Figure 2. Démarche scientifique pour fins de modélisation des distributions de temps de parcours sur les tronçons d'un kilomètre

## 2.1 Analyse exploratoire d'un circuit

Les données de temps de parcours ont été examinées de différentes façons : globalement, par tronçon, par groupes de tronçon. Cette section vise à illustrer une démarche d'analyse exploratoire d'un circuit (des tronçons d'un circuit). Pour fins de démonstration le circuit 1 est utilisé comme base d'analyse : celui empruntant l'autoroute 13, direction sud. Différents résultats statistiques sont disponibles en Annexe A.

### 2.1.1 Présentation du circuit 1

Le circuit 1 emprunte l'autoroute 13, direction sud. Celui-ci commence donc au croisement de l'A640, coupe ensuite l'A440 au tronçon 7, l'A40 au 16, l'A520 au 19 et enfin l'A20 à la fin du circuit (voir Figure 3). Par conséquent, une voiture empruntant le circuit 1, se dirige vers Montréal. L'intérêt de ce circuit repose dans le fait que celui-ci dispose d'une propriété simple mais pratique pour fins d'analyse : la vitesse limite y est constante et égale à 100 km/h.

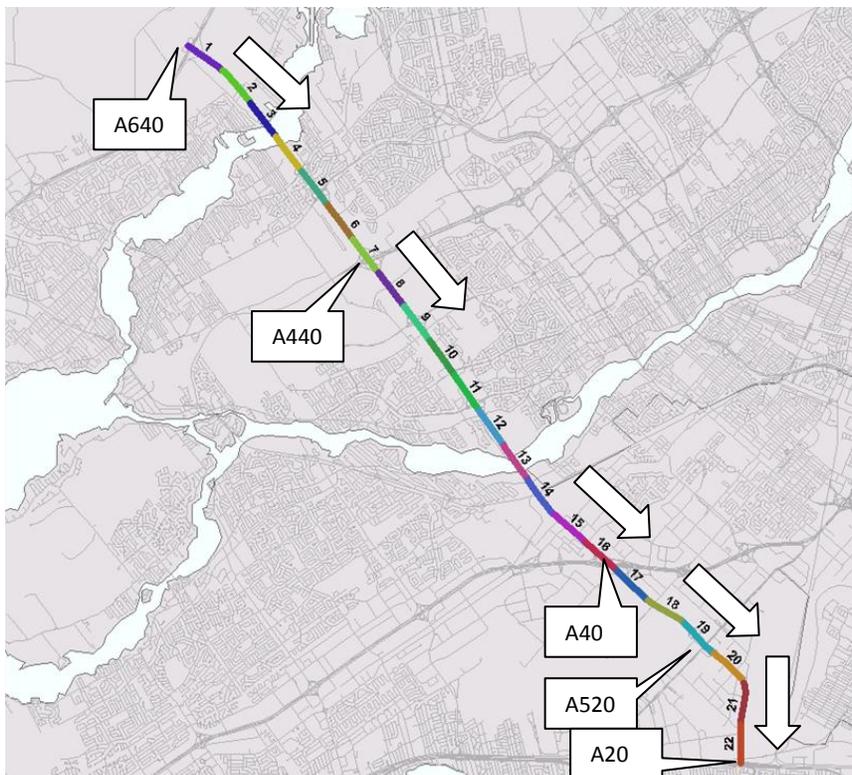


Figure 3. Circuit 1, direction et intersections majeures

### 2.1.2 Localisation des points chauds du circuit 1

Compte-tenu de la configuration spatiale et du sens de déplacement sur le circuit 1, il devient pertinent de considérer simplement la moyenne des temps de parcours tout au long de notre circuit, selon la période d'étude (AM ou PM)<sup>1</sup>. La Figure 4 présente le temps de parcours moyen par tronçon sur le circuit 1, en AM et PM. On observe une différence notable entre le comportement du temps de parcours moyen le matin (AM) et l'après-midi (PM). En période AM, la moyenne des temps de parcours tout au long du circuit 1 est plus importante que pour l'après-midi.

<sup>1</sup> Dans l'Annexe A, la signification d'un facteur à 2 niveaux (AM/PM) est démontrée par une analyse de la variance

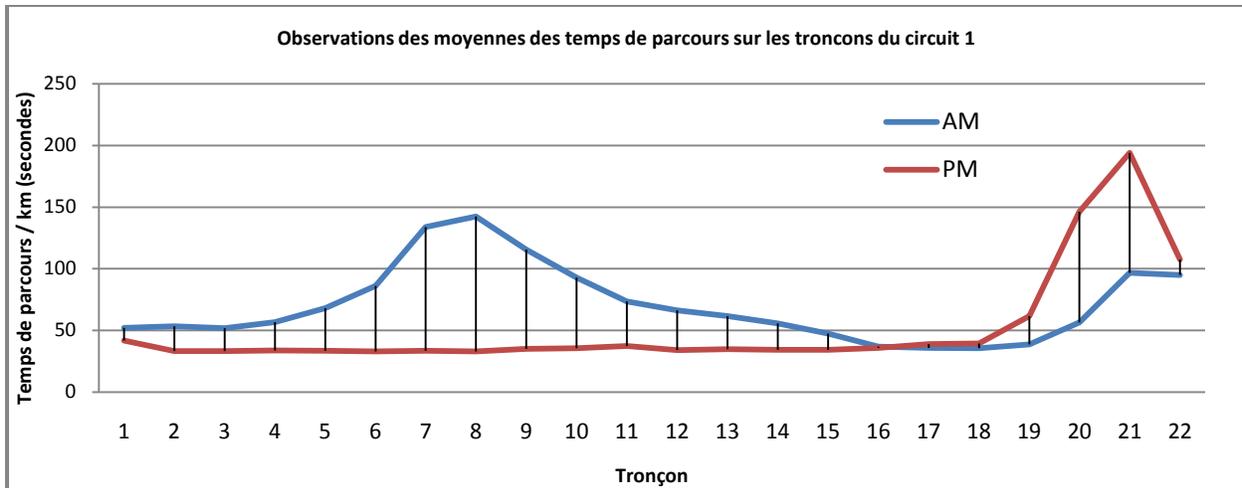


Figure 4. Évolution des temps de parcours par tronçons du circuit 1

Ce constat est conforme à notre connaissance des patrons de mobilité typiques des montréalais à savoir que ceux-ci convergent vers les parties centrales de Montréal en pointe du matin afin d'accéder aux pôles principaux d'emplois. Cette conclusion est statistiquement confirmée grâce à une analyse de variance (Figure 5).

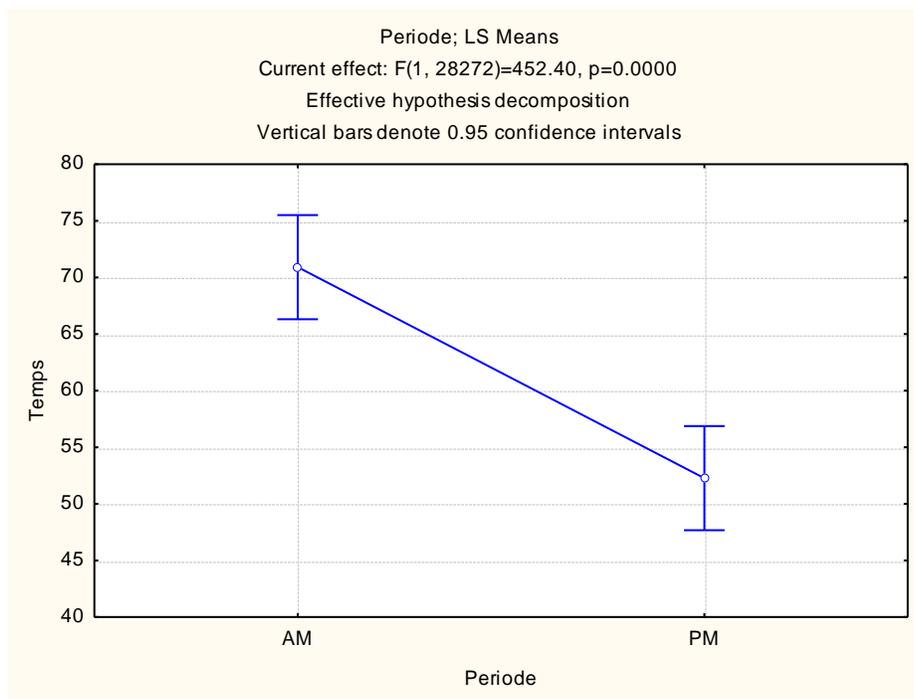


Figure 5. Moyenne des temps de parcours du circuit 1 selon la période

On observe que, durant la période AM, il y a deux zones de congestion. La première a lieu au tronçon 8 et s'étend entre le 7 et le 10 (la congestion s'étend grosso modo sur 4 km). La deuxième zone s'étend autour du tronçon 21 et est commune aux deux périodes. Rappelons que par rapport aux zones d'entrecroisement, nous pouvons émettre l'hypothèse que durant la matinée, de nombreuses voitures s'insèrent dans le circuit 1 au niveau de l'entrecroisement A13-A40 et semblent quitter ce circuit au

niveau de l'entrecroisement A13-A40. En effet, la moyenne des relevés de temps de parcours est décroissante du tronçon 8 au tronçon 16. Par ailleurs, la valeur de la moyenne des relevés des temps de parcours du tronçon 16 pour la période matinale, tous autres facteurs confondus, est de 37 secondes ce qui correspond à une vitesse 97 km/h. En clair, la vitesse limite étant constante et égale à 100km/h sur la totalité du circuit 1, on peut supposer que la circulation n'est plus congestionnée dès que l'échangeur de l'A40 est passé. Enfin, on retrouve un pic en fin de circuit qui correspond simplement à une fluctuation normale de sortie d'autoroute.

Nous venons de confirmer une différence significative entre le matin et l'après-midi avec une localisation spatiale pertinente des points chauds, puisque, effectivement, les pics de la courbe matinale des moyennes de temps de parcours correspondent très justement aux zones d'intersection avec d'autres axes autoroutiers.

### 2.1.3 Évolution temporelle des zones de congestion

---

Au moyen des relevés dont nous disposons, comment caractériser temporellement les zones de trafic congestionnées ? En effet, au cours de la matinée, un pic de congestion peut apparaître et par la suite disparaître, sa période de « vie » est forcément délimitée dans le temps mais nous pouvons retrouver ce pic plus tard, sur le même circuit.

Ainsi, pour observer l'évolution temporelle des moyennes des relevés de temps de parcours tout au long du circuit 1, nous choisissons de considérer des créneaux de 20 minutes soit 1200 secondes (vitesse de parcours de 5 km/h), ce qui permet d'éviter des chevauchements temporels (en effet très peu de relevés de temps de parcours ont des valeurs supérieurs à ce seuil). En clair, nous nous assurons que si une voiture est prise dans une congestion, vingt minutes plus tard, il est quasi certain que le véhicule a parcouru plus d'un kilomètre. Au niveau de la base de données, cela revient simplement à dire que si l'heure d'un relevé de temps de parcours est égale à 8h23, alors la valeur du créneau horaire auquel appartient ce relevé sera 8h20.

La Figure 6 illustre le phénomène d'évolution temporelle de la zone congestionnée en présentant le temps de parcours moyen par tronçon à différentes périodes de la journée. Pour chaque graphique, l'abscisse correspond au numéro de tronçon alors de l'ordonnée indique le temps de parcours par km (en secondes). Les courbes noires correspondent aux moyennes des relevés de temps de parcours par tronçon sur le circuit 1 durant le créneau horaire considéré, la zone grise correspond aux moyennes des relevés de temps de parcours sur la journée.

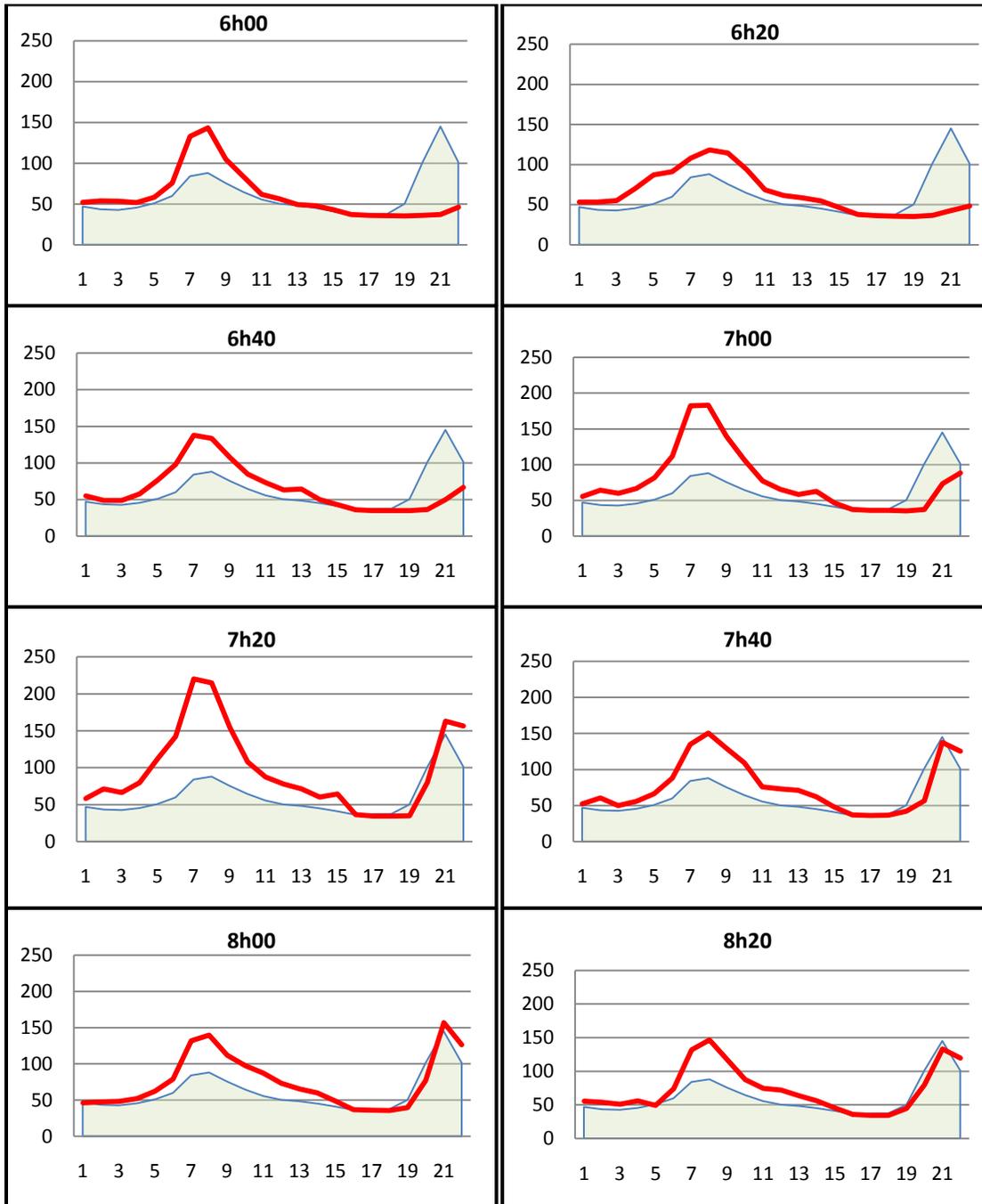


Figure 6. Évolution temporelle du temps de parcours moyen par tronçon entre 6h00 et 8h20 AM (abscisse : numéro du tronçon, ordonnée : temps de parcours par kilomètre, en secondes)

On constate qu'entre 6h40 et 8h20, la circulation est très congestionnée au niveau des tronçons 7 à 10 (intersection avec l'A440). Par conséquent, nous pouvons réitérer ce questionnement entre problème de congestion de la circulation sur les grands axes routiers et autoroutiers et le concept de déplacement OD (lignes de désir). De plus, le pic justement observé au niveau des tronçons, signifie que la situation est congestionnée de manière uniforme au cours du temps. Une des questions futures sera donc de qualifier la variabilité de cette congestion.

### 2.1.4 Observation des temps de parcours selon le mois et l'année

Tel que décrit dans le rapport précédent, tous les mois d'une année n'ont pas forcément été échantillonnés similairement. Néanmoins, des différences, intuitives, sont observées entre certains mois, septembre et les autres par exemple. Cette différence est confirmée par une analyse ANOVA (Figure 7). Ainsi, il n'est pas pertinent de comparer la moyenne des temps de parcours sur une base annuelle puisque l'échantillonnage pourrait être à la source des différences observées.

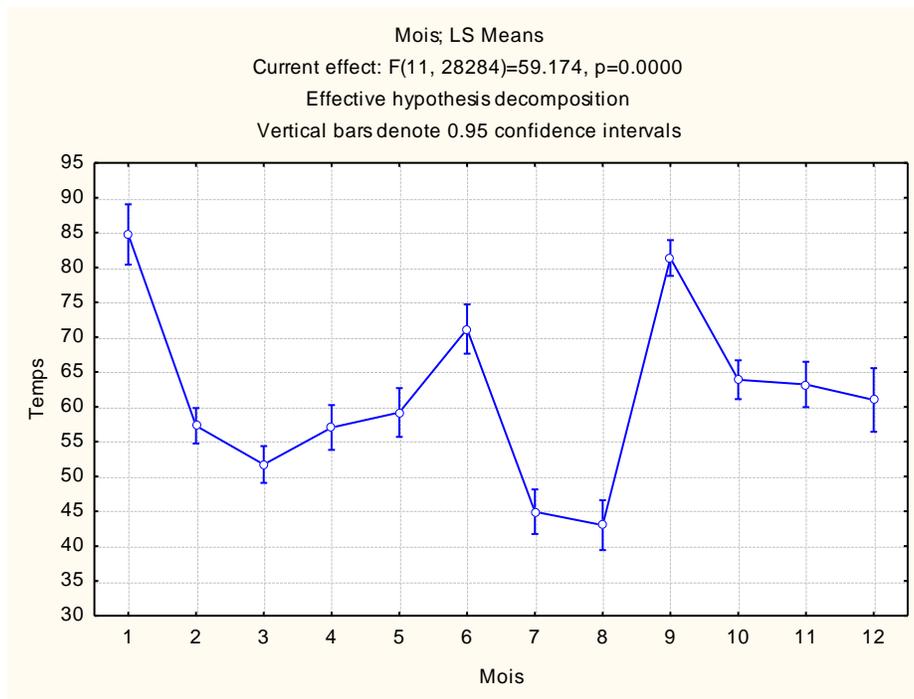


Figure 7. Moyenne des temps de parcours selon le mois

Il devient par contre très intéressant de concevoir un autre type d'analyse qui prend justement en compte le fait que tous les mois d'une même année ne sont pas échantillonnés de la même façon (ce type d'analyse statistique correspond à la régression ANOVA *Nested Design*). Le Tableau 69, disponible en Annexe, démontre que la fonction Mois(Année) est significative, les p-values étant très proches de 0. On peut alors observer le comportement moyen des temps de parcours, par exemple de l'année 2004, selon le mois de l'année. En clair, il existe une différence significative entre plusieurs mois d'une même année. Nous confirmons de plus que le mois de septembre est le mois le plus critique, et ce pour toutes les années. Les graphiques suivants (Figure 8), qui présentent l'évolution du temps de parcours, pour les deux périodes (AM et PM) et pour trois mois de l'année 2004, illustrent les différences observées entre mois. Pour chaque graphique, l'abscisse correspond au numéro de tronçon et l'ordonnée au temps de parcours par kilomètre, en secondes. Aussi, le trait correspond au temps de parcours moyen pour la

période et le mois indiqués alors que la surface indique le temps de parcours moyen annuel pour la période.

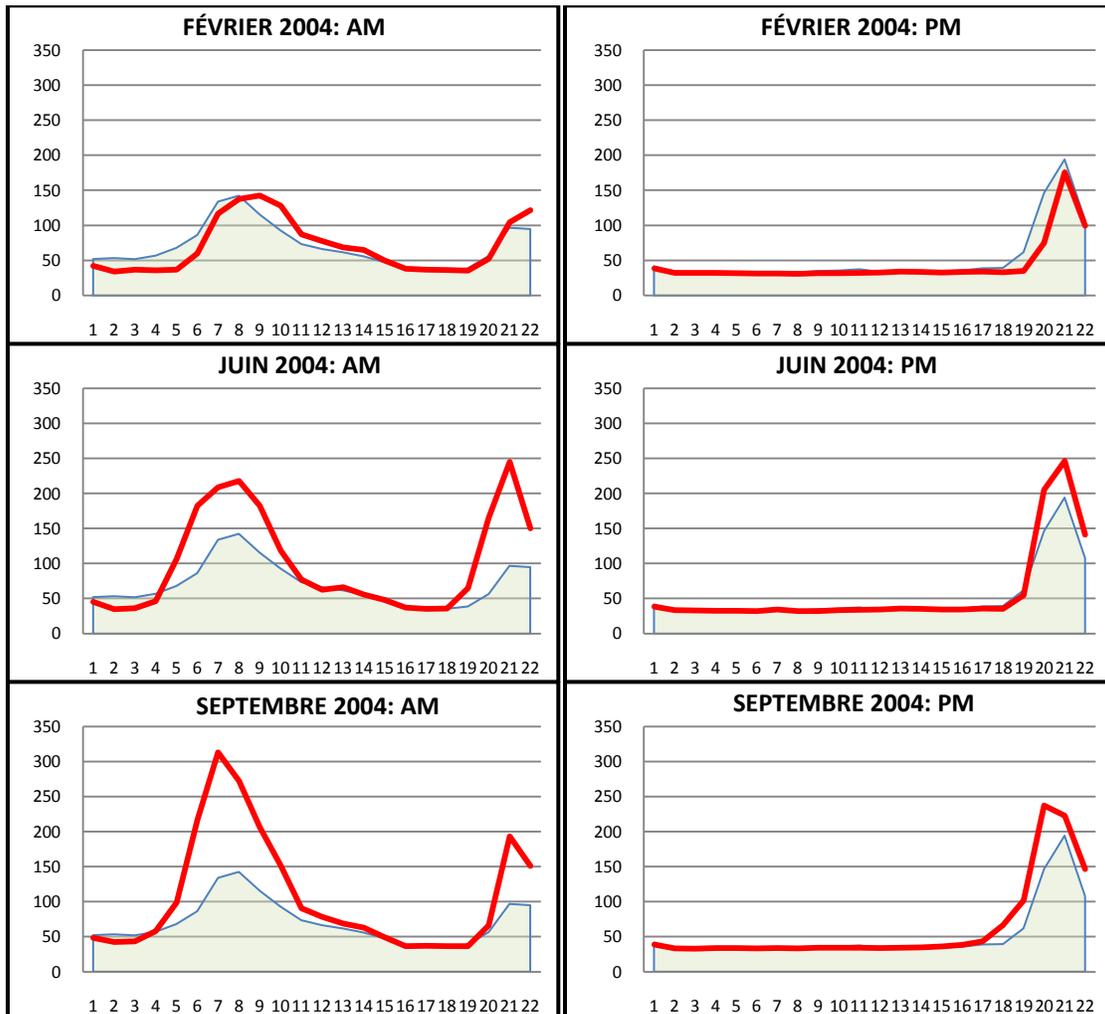


Figure 8. Évolution des temps de parcours pour les mois de février, juin et septembre 2004

### 2.1.5 Évolution d'indicateurs de fiabilité selon l'heure

Sachant qu'un des objectifs de cette étude est d'évaluer la fiabilité du réseau autoroutier montréalais, une application de quelques indicateurs de fiabilité, évoqués dans la revue de littérature, est proposée. Observons par exemple le *TTI* et le *TRI* pour les relevés réunis des tronçons 7 à 10 qui sont a priori de bons indicateurs de fiabilité.

Nous rappelons les formules suivantes,

$$TRI = \frac{T_{90} - T_{10}}{T_{50}} \text{ et } TTI = \frac{T_{90} - T_{50}}{T_{50}}, \text{ où } T(i) \text{ désigne le } i\text{ème centile.}$$

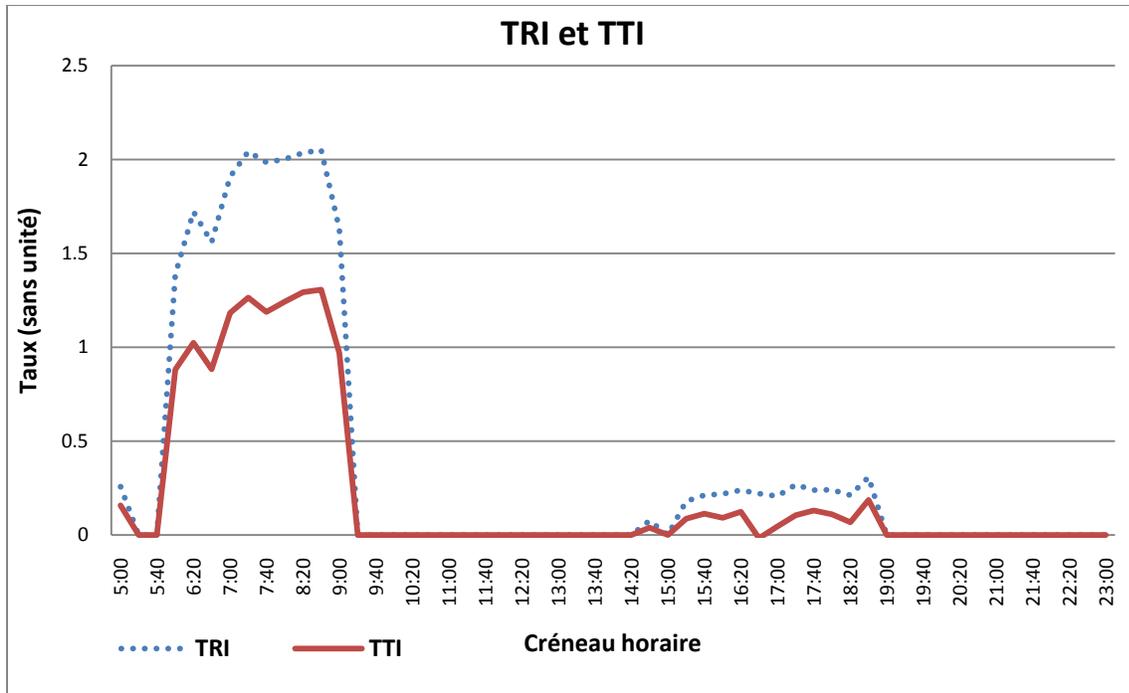


Figure 9. Observations des TRI et TTI selon l'heure

La Figure 9 expose de nouveau la fiabilité inférieure des tronçons 7 à 10. Il faut compter deux fois le temps de parcours moyen pour parcourir ne serait-ce, qu'un seul de ces tronçons durant la pointe matinale. Autrement dit, un conducteur empruntant ces tronçons doit multiplier son temps utile de parcours au niveau de ces tronçons entre 6h20 et 9h20 par 2. Signalons toutefois que nous ne connaissons pas l'état de ces indicateurs entre 9h40 et 15h00 d'où les valeurs nulles du graphique ci-dessus.

### 2.1.6 Observations selon les conditions de la chaussée

Une autre problématique importante de la circulation et de sa fiabilité est l'incidence des conditions de la chaussée sur le temps de parcours moyen sur un tronçon. Une analyse de variance a d'abord permis de valider la significativité des différences de temps de parcours selon différentes conditions de chaussée (voir Tableau 70 en Annexe).

Différents graphiques témoignent de ces différences :

- La Figure 10 présente les proportions de relevés recueillis annuellement selon les différentes conditions de chaussée. On observe, pour toutes les années, une prédominance de relevés faits dans des conditions de chaussées sèches.
- La Figure 11 présente les temps de parcours moyens pour les différentes conditions de chaussée et année.
- La Figure 12 présente le temps de parcours moyen, toutes années confondues, selon les conditions de la chaussée.

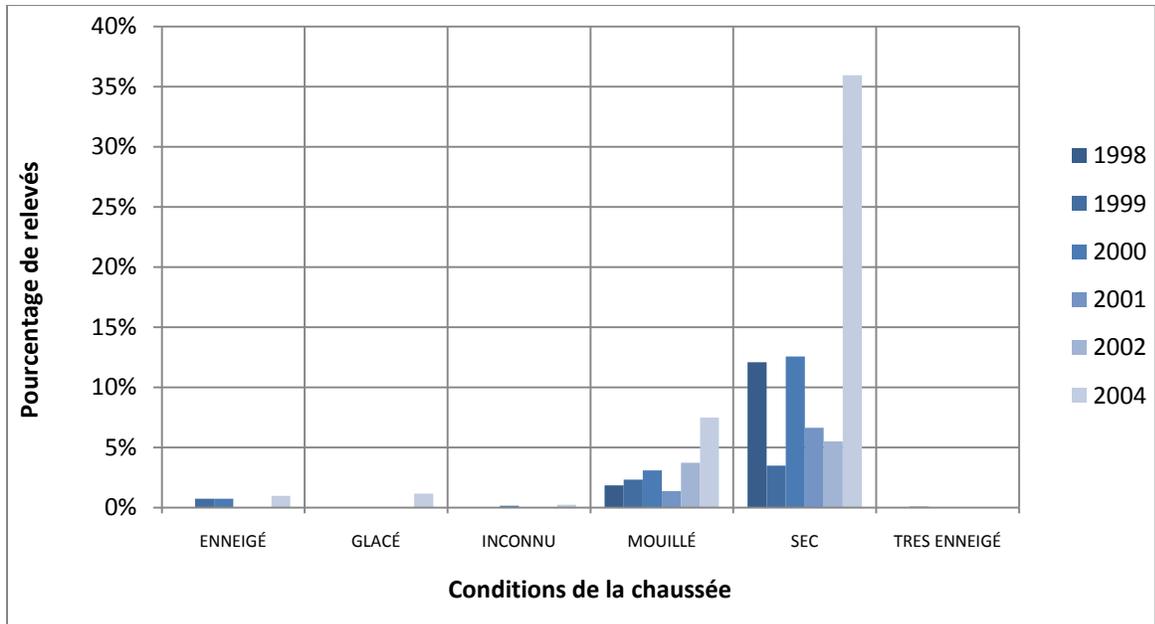


Figure 10. Pourcentage de relevés rencontrés par condition de la chaussée, par année, pour le circuit 1

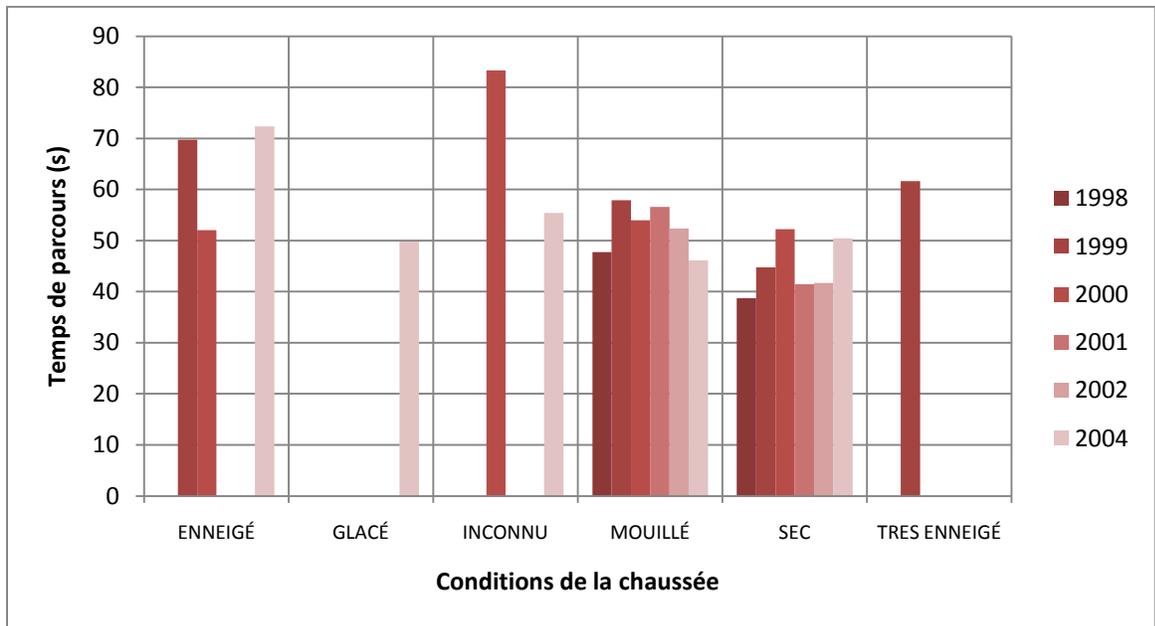


Figure 11. Temps de parcours moyens par condition de la chaussée et par année pour le circuit 1

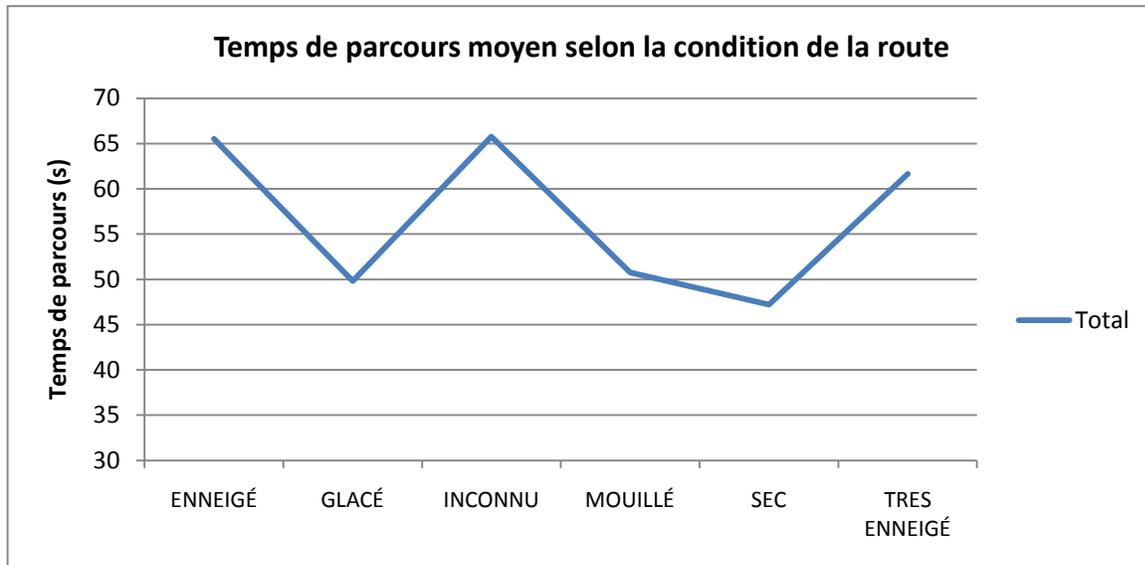


Figure 12. Temps de parcours moyen selon la condition de la route pour le circuit 1

Ces graphiques confirment que le temps de parcours moyen augmente lorsque la chaussée est enneigée ou très enneigée.

### 2.1.7 Synthèse

Ainsi, les éléments exposés ont permis de mettre en évidence la significativité statistique de nombreux facteurs. De plus l'analyse exploratoire a permis de dégager l'influence, sur les temps de parcours du circuit 1, de la période d'observation (AM vs PM), influence découlant notamment de la nature pendulaire des déplacements quotidiens.

Le cas du circuit 1 (Autoroute A13 Direction Sud) expose plusieurs conclusions propres à ce circuit : les différentes analyses ANOVA ont mis en exergue la signification de différents facteurs tels que le mois de l'année, la période d'observation, ou bien encore les conditions de la chaussée. Attention tout de même à ne pas en conclure que les facteurs cités précédemment sont significatifs pour tous les circuits.

## 2.2 Définition d'indicateurs et illustration

Un indicateur est une valeur permettant de donner une information pertinente pour son utilisateur. Celui-ci doit donc être objectif et fiable. Deux indicateurs sont proposés :

- le premier donnant une information sur l'occurrence de conditions de circulation anormales : indicateur de circulation anormale;
- le second unissant moyenne et variabilité simulées : IMV (indicateur moyenne-variabilité).

### 2.2.1 Indicateur de circulation anormale

L'indicateur de circulation anormale est la mesure du pourcentage de relevés reflétant des conditions de circulation très difficiles ou anormales. Il correspond aux relevés de temps de parcours dont la valeur est supérieure à 360 secondes, soit une vitesse inférieure à 10 km/h. Cet indicateur est directement issu des données observées.

La Figure 13 illustre cet indicateur pour le cas du circuit 1; on observe que les tronçons 2, 6, 7, 8, 19, 20, 21, ont de gros pourcentages d'observations reflétant des conditions anormales qui pourraient témoigner de la présence d'un incident. Cet indicateur permet d'identifier les tronçons les plus susceptibles d'être le lieu d'incidents ou autre phénomène similaire engendrant des conditions de circulation anormales.

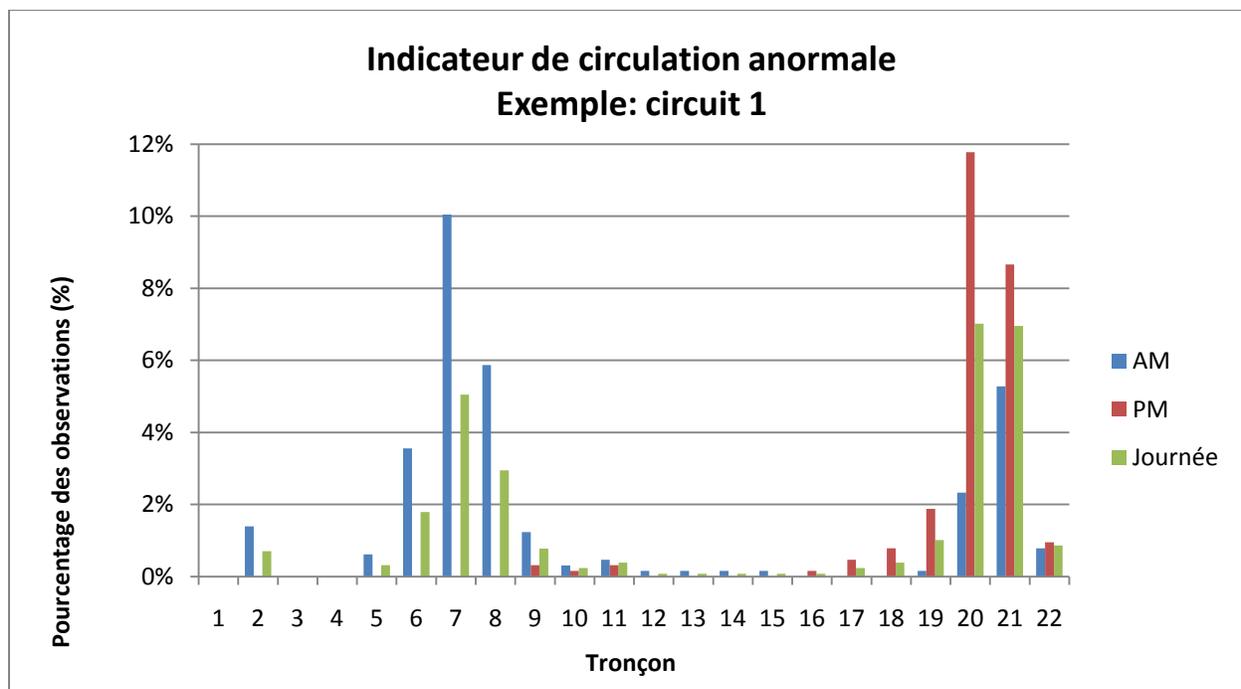


Figure 13. Indicateurs de circulation anormale pour le circuit 1

### 2.2.2 Indicateur Moyenne-Variabilité (IMV)

La moyenne et la variabilité des temps de parcours sont les principales mesures utilisées pour caractériser les conditions de circulation durant une période donnée. Il devient donc pertinent, dans un objectif de cerner la fiabilité des conditions routières, de créer un indicateur pouvant à la fois porter une information de performance (moyenne) et une information de fiabilité (variabilité).

L'objectif est donc de différencier quatre états de circulation sur un tronçon :

- Petite variabilité et petite moyenne : temps de parcours adéquats et assez constants;
- Petite variabilité et grande moyenne : temps de parcours élevés mais constants;
- Grande variabilité et petite moyenne : temps de parcours typiquement adéquats mais plus difficiles à prévoir;
- Forte variabilité et forte moyenne : temps de parcours typiquement élevés et difficiles à prévoir.

### 2.2.3 Présentation de l'indicateur

---

Pour différencier ces quatre états au moyen d'un simple graphe, nous choisissons de construire l'indicateur IMV, pour « indicateur de moyenne et de variabilité », comme un nombre complexe dont la formule suit :

$$IVM = (m - m_{seuil}) - i \cdot \frac{(v_{seuil} - v)}{v_{seuil}}$$
, où  $m$  désigne la moyenne,  $v$  la variabilité et  $i$  le nombre tel que  $i^2 = -1$ . Les valeurs seuils sont choisies de manière partielle ; on choisira pour la valeur seuil de la moyenne la borne supérieure de la circulation normale dans le cas de circuits du groupe 1 et 2, et la valeur seuil de la variabilité sera définie par l'utilisateur.

On pourrait proposer différentes formulations de cet indicateur (avec ou sans normalisation au dénominateur par exemple) mais fondamentalement, il s'agit de refléter les deux éléments clés de la circulation (moyenne et variabilité) et d'arriver à classer les principaux états (en quatre groupes par exemple) pour faciliter la comparaison et l'évolution des conditions.

Ainsi, dans le but de faciliter son utilisation, l'indicateur doit être catégorisé. Effectivement, la formule mathématique définissant l'indicateur IVM permet de caractériser un nouvel indicateur catégorique  $IVM_c$ , à 4 états selon les signes des parties réelle et imaginaire.

En voici la codification :

- Si  $\text{Re}(IVM) < 0$ , et  $\text{Im}(IVM) < 0$ , alors  $IVM_c = 0$  (petite moyenne, petite variabilité).
- Si  $\text{Re}(IVM) > 0$ , et  $\text{Im}(IVM) < 0$ , alors  $IVM_c = 1$  (grande moyenne, petite variabilité).
- Si  $\text{Re}(IVM) < 0$ , et  $\text{Im}(IVM) > 0$ , alors  $IVM_c = 2$  (petite moyenne, grande variabilité).
- Si  $\text{Re}(IVM) > 0$ , et  $\text{Im}(IVM) > 0$ , alors  $IVM_c = 3$  (grande moyenne, grande variabilité).

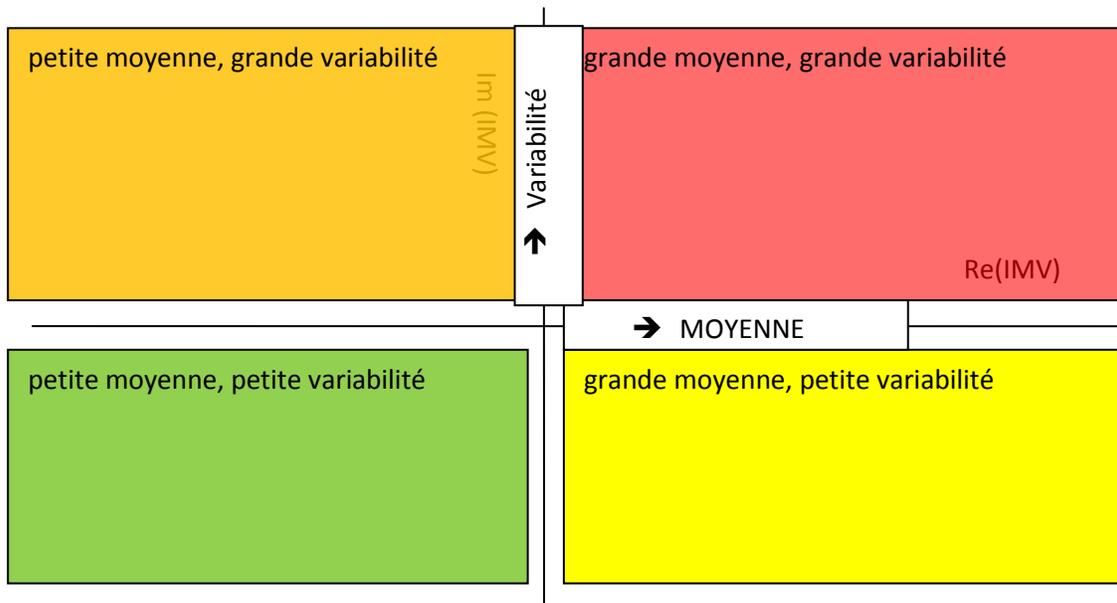


Figure 14. Représentation des 4 différentes zones de l'IMV

#### 2.2.4 Résultats : illustration

##### Valeurs de l'IMV selon la période

Nous présentons dans cette partie les valeurs de l'IMV pour le circuit 1 par tronçon. Les Figure 15 et Figure 16 présentent cet indicateur pour les périodes AM et PM respectivement. Rapidement, on observe les éléments suivants :

- AM : quatre tronçons ont des conditions idéales de petite moyenne et faible variabilité (16-17-18-19), quelque 6 tronçons semblent être régulièrement congestionnés (petite variabilité et grande moyenne : 1-11-12-13-14-15) alors que les autres tronçons présentent des conditions congestionnés et variables (plus difficiles à prévoir donc moins « fiables »).
- PM : la grande majorité des tronçons se retrouvent de justesse dans le cadran des conditions normales, avec des temps de parcours élevés et une faible variabilité (plausible en PM alors que le mouvement principal de circulation est sortant), les tronçons 19 et 20 sont « instables » avec des temps de parcours élevés et variables alors que les tronçons 22 et 21 sont typiquement congestionnés.

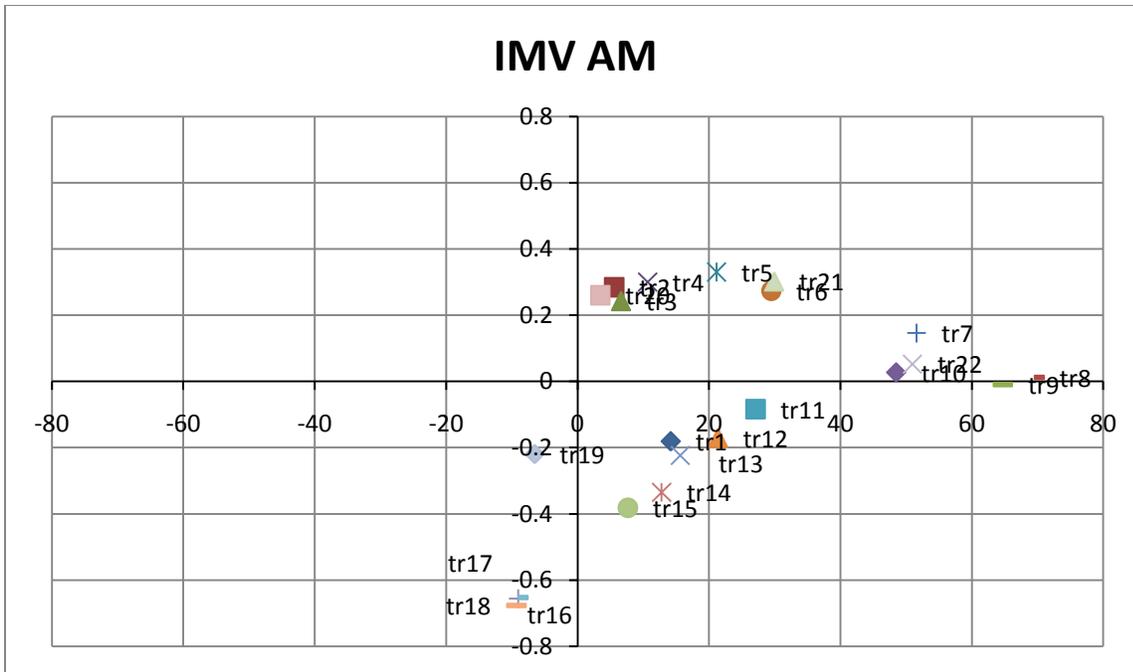


Figure 15. Valeurs de l'IMV par tronçon du circuit 1 AM

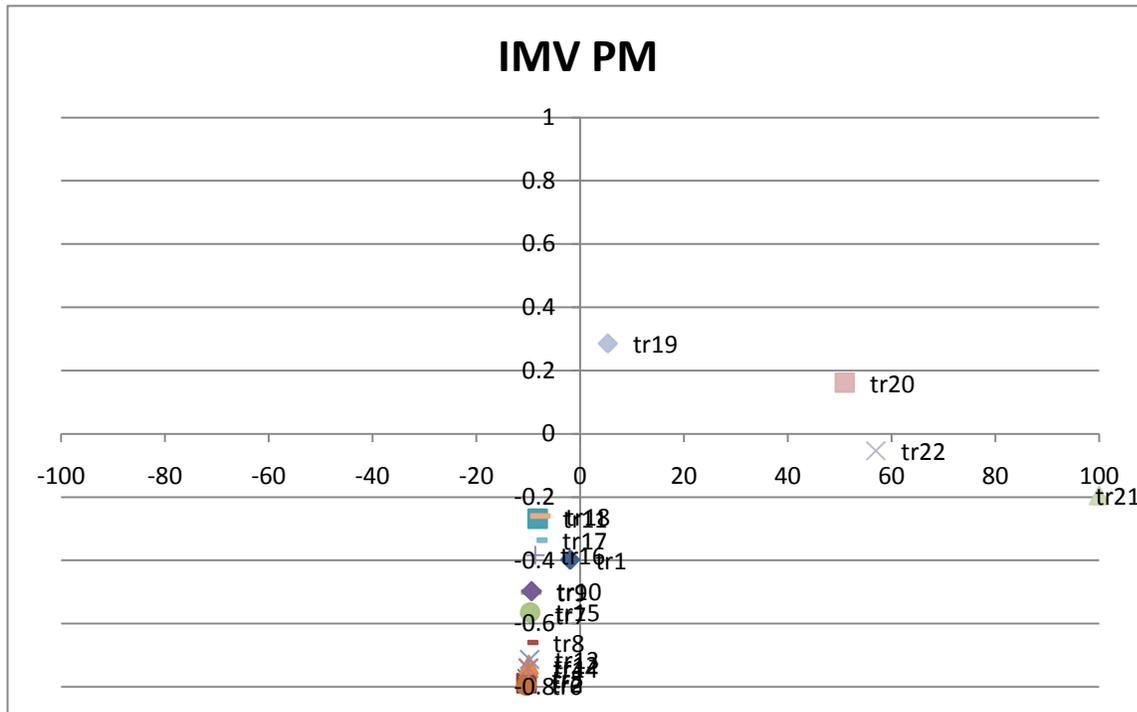


Figure 16. Valeurs de l'IMV par tronçon du circuit 1 PM

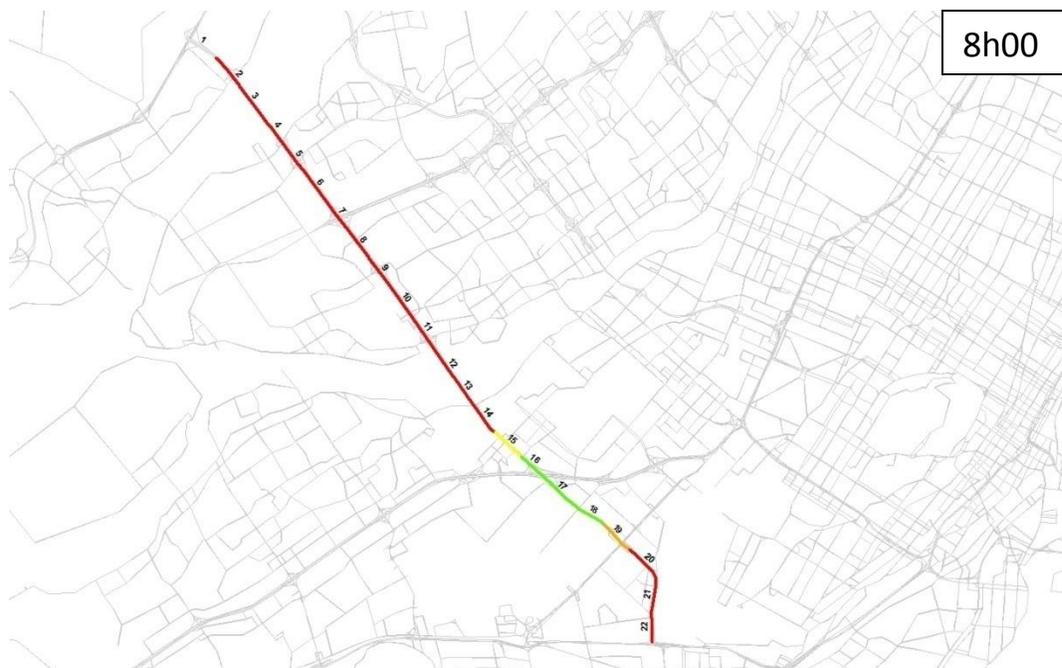
### IMV<sub>c</sub> par créneau de 10 minutes

Une observation de l'évolution de l'indicateur, par tranche de 10 minutes est maintenant réalisée pour le même circuit, en période AM. Dans la mesure où les tronçons ont été échantillonnés, l'indicateur est évalué sur tous les tronçons du circuit 1 en période matinale, toutes les dix minutes.

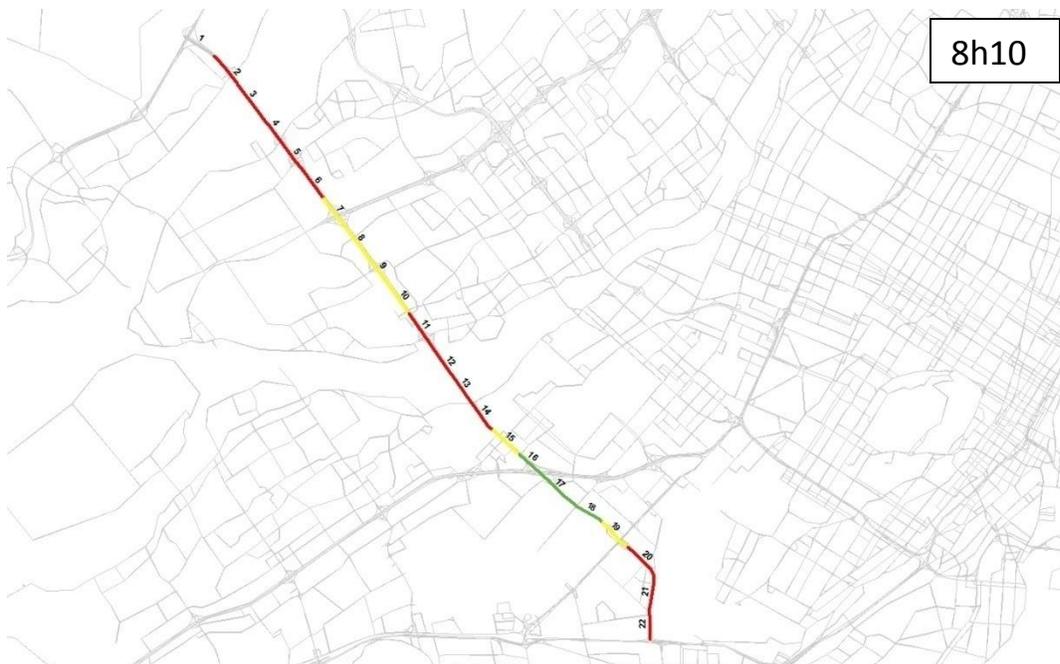
On rajoute de plus un code de couleur pour la représentation spatiale :

- $IVM_c = 0$ , couleur verte (petite moyenne, petite variabilité).
- $IVM_c = 1$ , couleur jaune (grande moyenne, petite variabilité).
- $IVM_c = 2$ , couleur orange (petite moyenne, grande variabilité).
- $IVM_c = 3$ , couleur rouge (grande moyenne, grande variabilité).

Dans le but ensuite de valider le modèle statistique ainsi que la pertinence de l'indicateur, on cherchera alors à comparer les valeurs de l'indicateur issu de la modélisation et les valeurs de l'indicateur selon les observations.



**Figure 17. Indicateur du circuit 1 période AM, 8h00**



**Figure 18. Indicateur du circuit 1 période AM, 8h10**



**Figure 19. Indicateur du circuit 1 période AM, 8h20**

Les valeurs de l'indicateur indiquent donc que les tronçons 2 à 15 sont sujets à de fortes variabilités et/ou sujets à de fortes congestions. Toutefois on constate que les tronçons 16 à 19 ne sont pas congestionnés. Effectivement une analyse complémentaire démontre que, au niveau de l'échangeur A13S – A40E, la situation est très congestionnée sur l'A40. De plus, la valeur de l'indicateur pour tous les créneaux de 10 minutes en matinée des 6 tronçons les plus proches de l'échangeur est 1, ce qui rappelons-le, équivaut à une faible variabilité et une forte moyenne des temps de parcours.

Ce type d'indicateur peut-être estimé directement à partir des données observées ou à partir de simulations (reconstruction de distributions fréquentielles de temps de parcours). En outre, différentes formulations peuvent-être testées afin que l'appartenance aux quadrants représente les seuils souhaités par le Ministère pour fins de dissémination ou analyse.

## 2.3 Analyse mathématique des circuits du réseau autoroutier montréalais.

En vue de valoriser correctement les données disponibles et ce, malgré l'échantillonnage variable en certains endroits, il devient souhaitable de créer un lot d'objets d'analyse. Les analyses précédentes ont permis d'observer que les distributions fréquentielles des temps de parcours de certains circuits sont très similaires ou différentes. Il apparaît donc opportun de grouper les tronçons présentant des distributions fréquentielles similaires et de procéder à la modélisation de ces distributions typiques par groupe (cluster), profitant ainsi d'un échantillon plus grand d'observations.

### 2.3.1 Méthode de groupement par similarité

L'étape de groupement des tronçons par similarité des distributions fréquentielles s'appuie sur une méthode issue du data mining dite clustering. Cette méthode permet de grouper des objets en fonction de la similarité de leurs propriétés. Dans le cas présent, les tronçons sont décrits par le biais de la distribution fréquentielle des temps de parcours observés, par période. Suite au groupement, un exercice de modélisation des distributions types par cluster est effectué.

#### Analyse fréquentielle

Choisissons  $tr_{j,c}^p$  le tronçon  $j$  du circuit  $c$ , pour la période d'observation  $p = AM$  ou  $PM$ , et soit une observation  $t$  sur ce tronçon.

On définit le vecteur  $T = [t_1, t_2, \dots, t_m] = [20, 25, \dots, 360]$ . Par conséquent, il existe  $t_i$  et  $t_{i+1}$  telle que  $t$  appartienne à  $[t_i ; t_{i+1})$ , où  $|t_i - t_{i+1}| = 5s$  si  $t < 360s$  (vitesse de 10km/h) ou  $[360 ; \infty[$  si  $t > 360s$ .

Soit  $n_{j,c}^p$  le nombre total d'observations pour le tronçon  $tr_{j,c}^p$ , on désignera alors par

$$f_{j,c,t_1,t_2}^p = \frac{n_{j,c,t_1,t_2}^p}{n_{j,c}^p},$$

où  $n_{j,c,t_1,t_2}^p$  désigne le nombre d'observations entre  $[t_1 ; t_2)$ , pour le tronçon  $j$  du circuit  $c$  durant la période  $p$ , la fréquence d'observations du tronçon  $tr_{j,c}^p$ , comprises entre  $[t_1 ; t_2]$ . Par conséquent, il vient que  $\sum_{k \in \llbracket 1; m \rrbracket} f_{j,c,t(k),t(k+1)}^p = 1$ .

#### Algorithme de Clustering

Les détails mathématiques de la méthode utilisée pour grouper les tronçons sont présentés ci-dessous.

Soit  $(x_1, x_2, \dots, x_n)$  un ensemble de vecteurs-observations. On désignera alors  $x_i^T = (x_i^1, x_i^2, \dots, x_i^m)$ , le vecteur de dimension  $m$  correspondant aux valeurs définissant le vecteur-observation  $x_i$ .

On rappelle que la distance euclidienne entre deux vecteurs  $x$  et  $y$  est égal à :

$$\|x - y\| = \sqrt{\sum_{k=1}^m (x^k - y^k)^2}$$

On définit alors une valeur réelle strictement positive  $\delta$ .

La construction des différents groupes s'effectue alors ainsi :

- Soit  $\Omega_l$ , le groupe singleton comprenant le vecteur  $x_l$  et de centre  $C_l = x_l$ .
- Comparons la distance euclidienne  $\|x_1 - x_2\|$  entre  $x_1$  et  $x_2$  :
  - Soit  $\|x_1 - x_2\| \leq \delta$ , dans ce cas alors on inclut dans le groupe  $\Omega_l$  le vecteur  $x_2$  et le centre  $C_l$  possède alors les nouvelles coordonnées  $\left[\frac{1}{2} (x_1^i + x_2^i)\right] = \frac{1}{2} (x_1 + x_2)$ .
  - Soit  $\|x_1 - x_2\| > \delta$ , alors on définit le nouveau groupe  $\Omega_2$  de centre  $C_2 = x_2$ ,
- Ainsi de suite, grosso modo, pour les  $n-2$  autres vecteurs.

Le résultat conséquent est la création d'un nombre de groupe distincts ( $\Omega_k$ ). Le groupe  $k$ -ième comprend alors  $l$  observations ( $x_{\Omega(1)}, x_{\Omega(2)}, \dots, x_{\Omega(n)}$ ) de centre  $C_k$  de coordonnées  $C_k^i = \frac{1}{l} \sum x_{\Omega(s)}^i$  et est tel que quelque soit le vecteur-observation  $x$  élément de  $\Omega_k$ , alors  $\|x - C_k\| = \sqrt{\sum (x^i - C_k^i)^2} \leq \delta$ .

### 2.3.2 Résultats

Différentes expériences de groupement ont été menées. D'abord, une classification en 10 clusters a été développée mais certaines distributions moyennes ont permis d'observer qu'il était nécessaire de modéliser séparément les conditions observées pendant les deux périodes (AM vs PM).

Pas conséquent, la modélisation retenue s'appuie sur deux ensembles de huit groupes, chaque ensemble permettant de résumer les distributions fréquentielles observées pendant les deux grandes périodes de la journée. La Figure 20 présente les distributions de quelques centres de groupes de la période AM. On observe facilement les différences de distributions :

- Le groupe AM-2 présente une distribution très concentrée (très faible variabilité) autour de petites valeurs de temps de parcours. Il s'agit donc d'un groupe constitué de tronçons sans congestion;
- Le groupe AM-1 a aussi une distribution concentrée (faible variabilité) mais autour de valeurs de temps plus élevées;
- Le groupe AM-5 a une distribution plus étalée et moins symétrique;
- Le groupe AM-7 a la distribution la plus étalée centrée autour de valeurs plus élevées de temps de parcours (tronçons plus congestionnés et de conditions plus variables)

Le Tableau 1 et la Figure 21 présentent les valeurs caractéristiques des groupes créés pour la période AM, caractéristiques qui permettent de nommer les groupes en fonction des conditions typiques de circulation.

- Près d'un tiers des tronçons présentent des conditions de circulation avec des vitesses élevées et une variabilité raisonnable (groupe AM-2);
- Les groupes 4 et 5 présentent les coefficients de variation les plus élevés et des vitesses très faibles;
- Le groupe 8 présente une vitesse faible et une variabilité acceptable : on peut penser que les tronçons liés à ce groupe ne sont pas des tronçons autoroutiers.

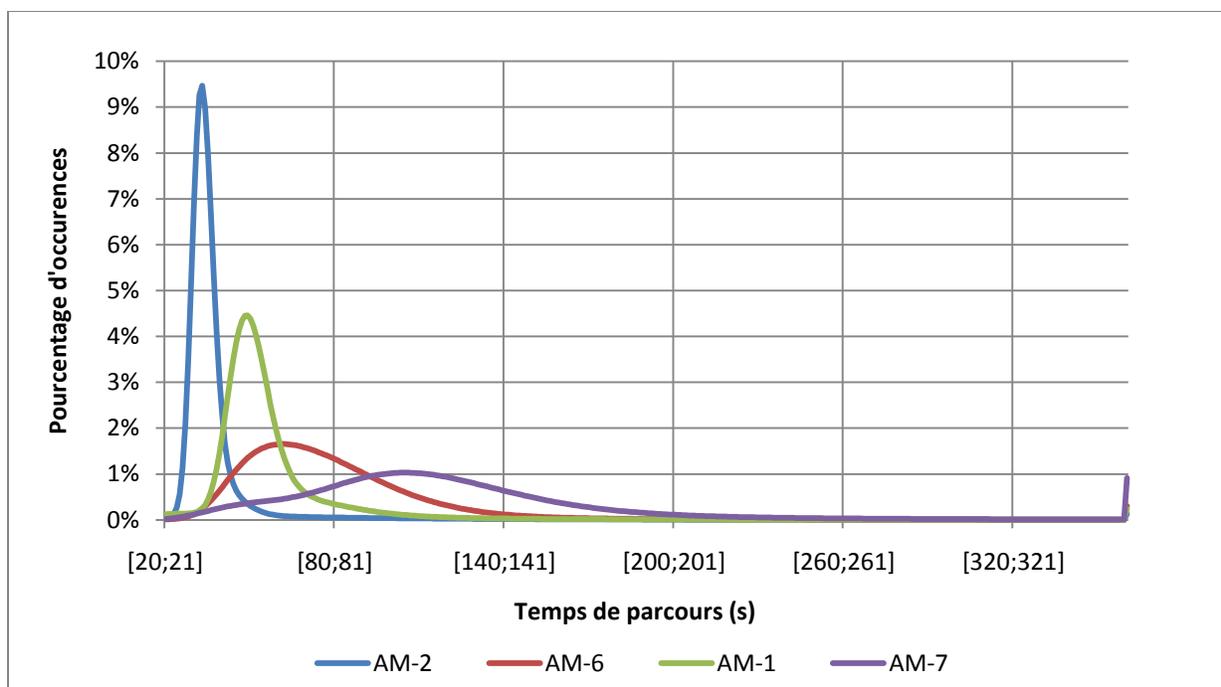


Figure 20. Distribution fréquentielle de quelques groupes de tronçons pour la période AM

Tableau 1. Valeurs caractéristiques des groupes AM

AM	Temps de parcours moyen par km (s)	Coefficient de Variation	Vitesse (km/h)	Pourcentage de tronçons	% observations > 360 sec
1	58.6	58.2%	61.4	10.9%	0.2%
2	39.2	63.8%	91.9	32.6%	0.2%
3	48.2	46.2%	74.7	21.3%	0.4%
4	156.2	114.8%	23.0	3.8%	11.0%
5	109.1	74.2%	33.0	6.5%	3.0%
6	79.8	44.2%	45.1	15.8%	0.3%
7	119.4	46.7%	30.1	6.9%	0.9%
8	189.7	45.7%	19.0	2.2%	9.0%

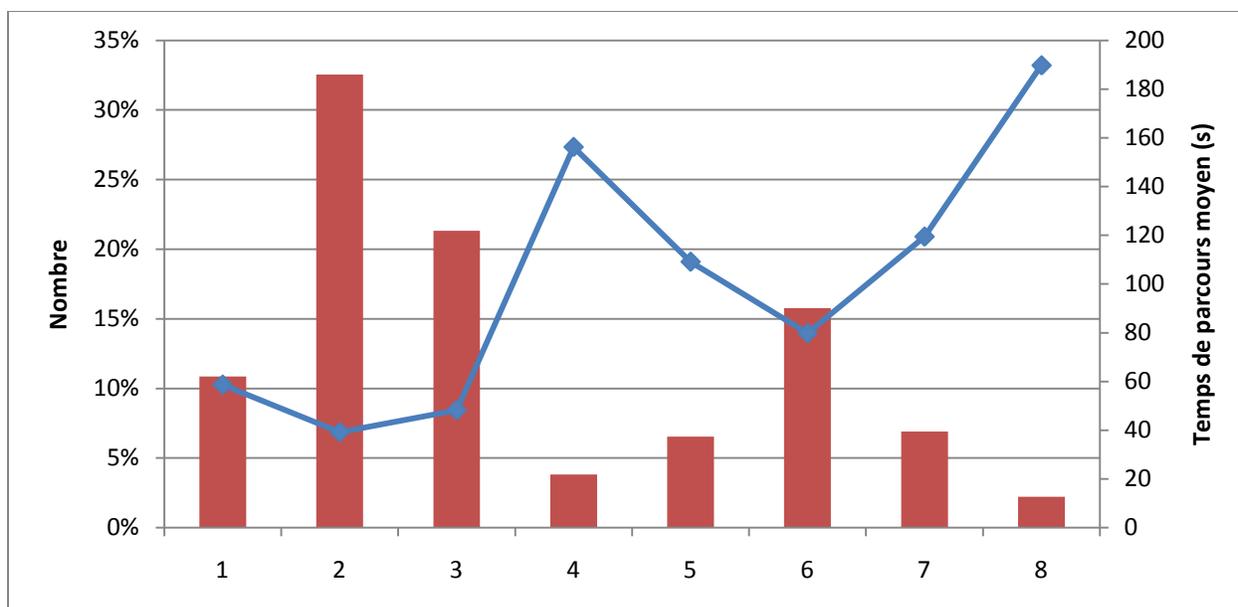


Figure 21. Répartition des tronçons selon les groupes avec moyenne, période AM (histogramme : nombre de tronçons, points : temps de parcours)

Le Tableau 2 et la Figure 22 présentent les valeurs caractéristiques des groupes créés pour la période PM : globalement, la répartition des tronçons dans les groupes est plus uniforme. Les groupes 6 et 8 ont les vitesses moyennes les plus élevées et se distinguent par leur niveau de variabilité.

Tableau 2. Valeurs caractéristiques des groupes PM

PM	Moyenne (s)	Coefficient de Variation	Vitesse (km/h)	Pourcentage de tronçons	% observations > 360 sec
1	185.8	84.9%	19.4	5.7%	9.8%
2	63.1	52.8%	57.0	13.1%	0.3%
3	53.8	70.4%	66.9	15.6%	0.6%
4	123.5	46.9%	29.1	10.2%	0.9%
5	88.1	41.0%	40.9	13.2%	0.3%
6	37.1	52.8%	97.0	18.8%	0.1%
7	50.9	71.0%	70.8	13.2%	0.3%
8	36.7	67.1%	98.0	10.1%	0.0%

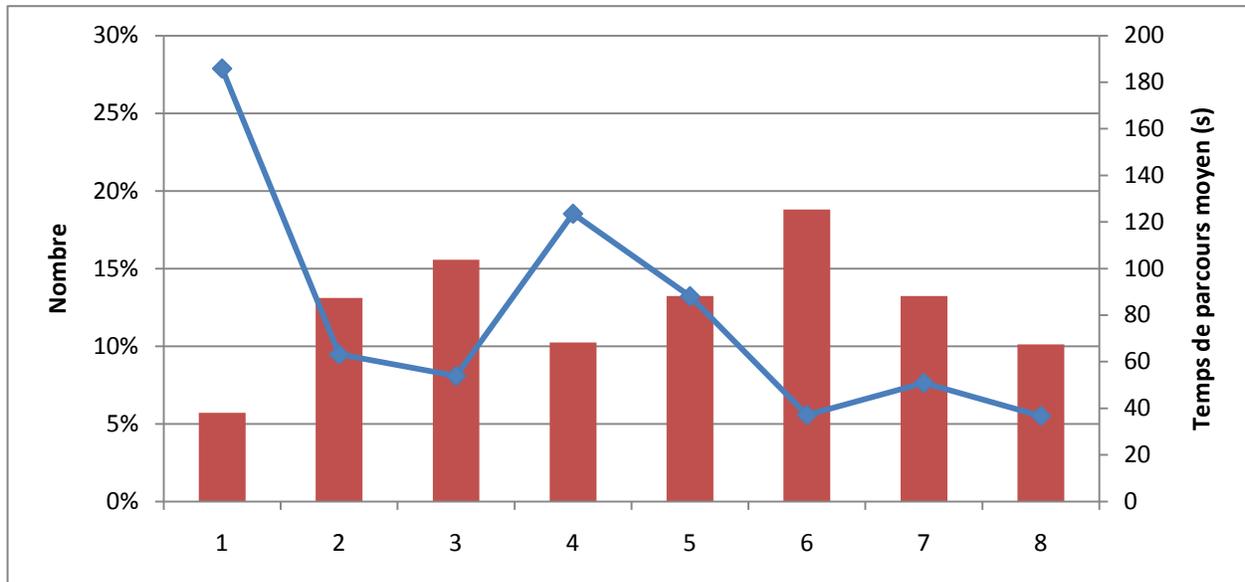


Figure 22. Répartition des tronçons selon les groupes avec moyenne, période PM (histogramme : nombre de tronçons, points : temps de parcours)

### 2.3.3 Modélisation des distributions

Sur la base des groupes précédemment créés, une modélisation de la distribution fréquentielle des relevés de temps de parcours est effectuée au moyen de trois lois lognormales. Le modèle à trois lois lognormales permet de modéliser pratiquement l'ensemble des observations, sans devoir traiter distinctement les observations témoignant de situations très congestionnées qui peuvent être liées à des incidents. En effet, pour la plupart des groupes, la modélisation au moyen de cette formule englobe plus de 95% des relevés.

Nous donnons la formule générique de la distribution fréquentielle théorique :

$$f(x) = \lambda_1 \cdot f_{\mu_1 \sigma_1}(x) + \lambda_2 \cdot f_{\mu_2 \sigma_2}(x) + \lambda_3 \cdot f_{\mu_3 \sigma_3}(x) \text{ où } f_{\mu_i \sigma_i} \text{ désigne une loi log} \\ \text{– normale de paramètres } \mu_i \sigma_i$$

Le développement d'un modèle statistique pour chaque groupe implique la détermination de valeurs seuils pour chaque loi lognormale. Dans le cas présent, trois types de distribution, appelant des mécanismes de calibration différents, ont été observés :

- les distributions avec une loi lognormale dominante qui seront appelées distributions avec pic étroit;
- les distributions avec deux lois lognormales importantes montrant une concentration importante de valeurs qui seront nommées distributions avec pic étalé;
- les distributions très dispersées avec une forte proportion de valeurs élevées de temps de parcours qui seront appelées distributions congestionnées.

Le processus de détermination des paramètres des modèles est illustré ci-dessous.

### Distributions avec "Pic étroit"

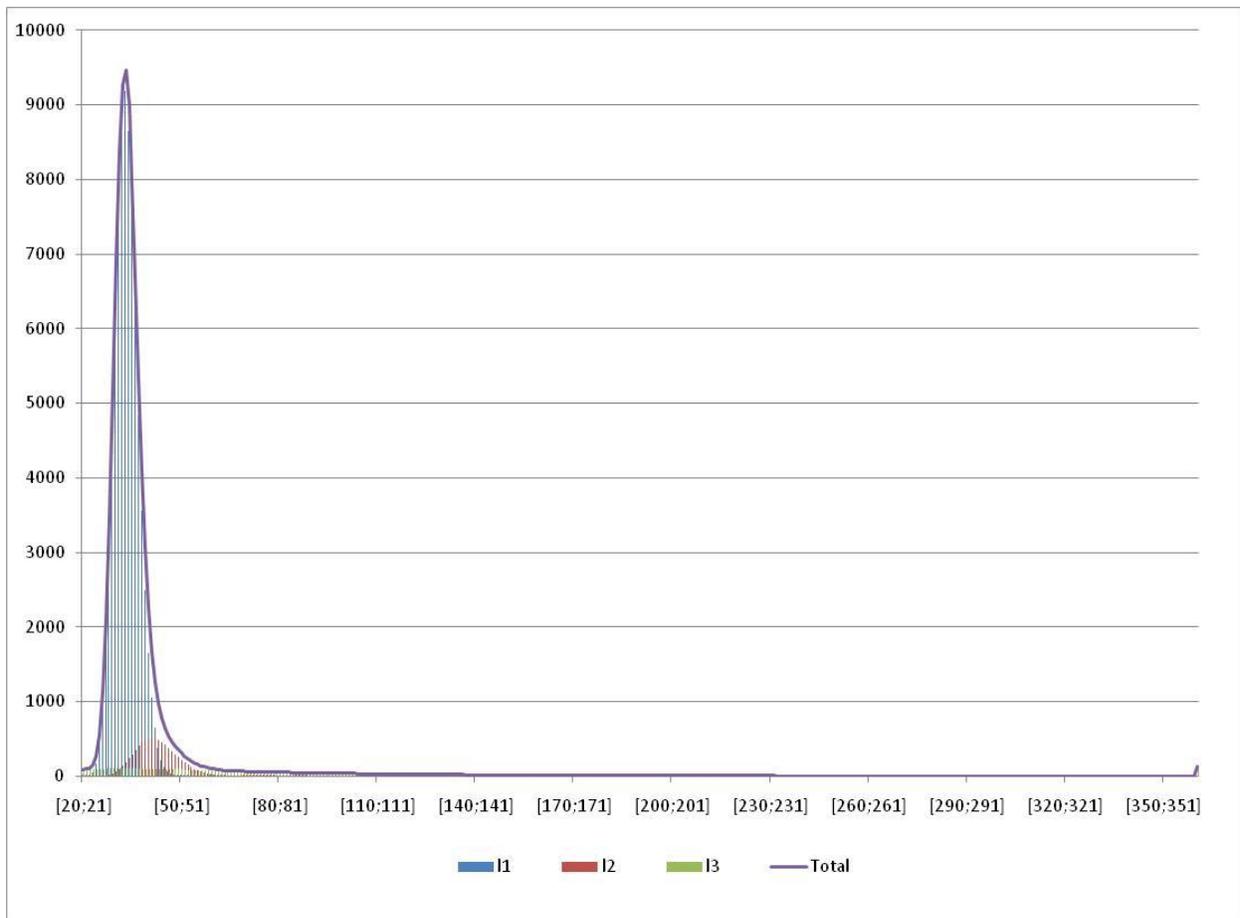
Dans le cas le plus simple, considérons l'exemple suivant, tel que le second groupe de tronçons durant la matinée. Les caractéristiques de sa distribution fréquentielle de temps de parcours sont :

**Tableau 3. Paramètres du modèle de distribution fréquentielle pour les cas de « Pic étroit »**

	Loi 1	Loi 2	Loi 3
Moyenne	3.514	3.748	4.108
Écart-type	0.107	0.149	0.811
Lambda	0.828	0.08	0.092

Temps de parcours moyen (s)	33.78	42.91	84.51
Vitesse moyenne (km/h)	106.59	83.90	42.60



**Figure 23. Distribution fréquentielle des temps de parcours pour le groupe AM-2**

On observe que la distribution présente un fort pic autour de la valeur 33s. L'intervalle de confiance à 95% autour de la moyenne de chaque loi lognormale est :

**Tableau 4. Moyennes et intervalles de confiance des trois lois définissant le groupe AM-2**

	I1	I2	I3
Moyenne	33.8	42.9	84.5
Borne inférieure	27.2	31.7	12.4
Borne supérieure	41.4	56.8	298.1

Les quatre intervalles définissent simplement les quatre catégories de circulation pour ce groupe :

- 0 si  $t$  est inclus dans [25;42], intervalle contenant les temps de parcours usuels
- 1 si  $t$  est inclus dans [42;57], intervalle contenant les temps de parcours tampons
- 2 si  $t$  est inclus dans [57;299], intervalle comprenant les grandes valeurs de temps de parcours
- 3 si  $t$  est inclus dans [299;360], intervalle contenant les cas extrêmes de temps de parcours

**Tableau 5. Définition des intervalles du groupe AM-2**

Définition des intervalles			
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4
[24;42]	[42;57]	[57;299]	[299;360]

Cette approche est viable pour les groupes suivants :

- groupes AM : 1, 2, 3, 4, 5
- groupes PM : 2, 3, 6, 7, 8

**Distributions avec "Pic étalé"**

En raison des distributions fréquentielles de temps de parcours observés, la méthode utilisée précédemment ne peut être appliquée à tous les groupes. Pour les groupes AM-6 et PM-5, la distribution fréquentielle présente un large pic. Dans ce cas, nous décidons de couper le pic en deux, chaque part correspondant à un intervalle où une loi lognormale domine les autres. Ensuite, la classification devient normale. En termes de coefficients  $\lambda$ , deux des trois coefficients sont proches.

Remarque : la valeur 25s (ce qui correspond à une vitesse de 145 km/h) est une convention ; il est considéré que toute valeur inférieure est le résultat d'une erreur expérimentale, ou de conducteurs atypiques.

La Figure 24 présente la distribution du groupe AM-6 ainsi que les contributions de chaque loi lognormale.

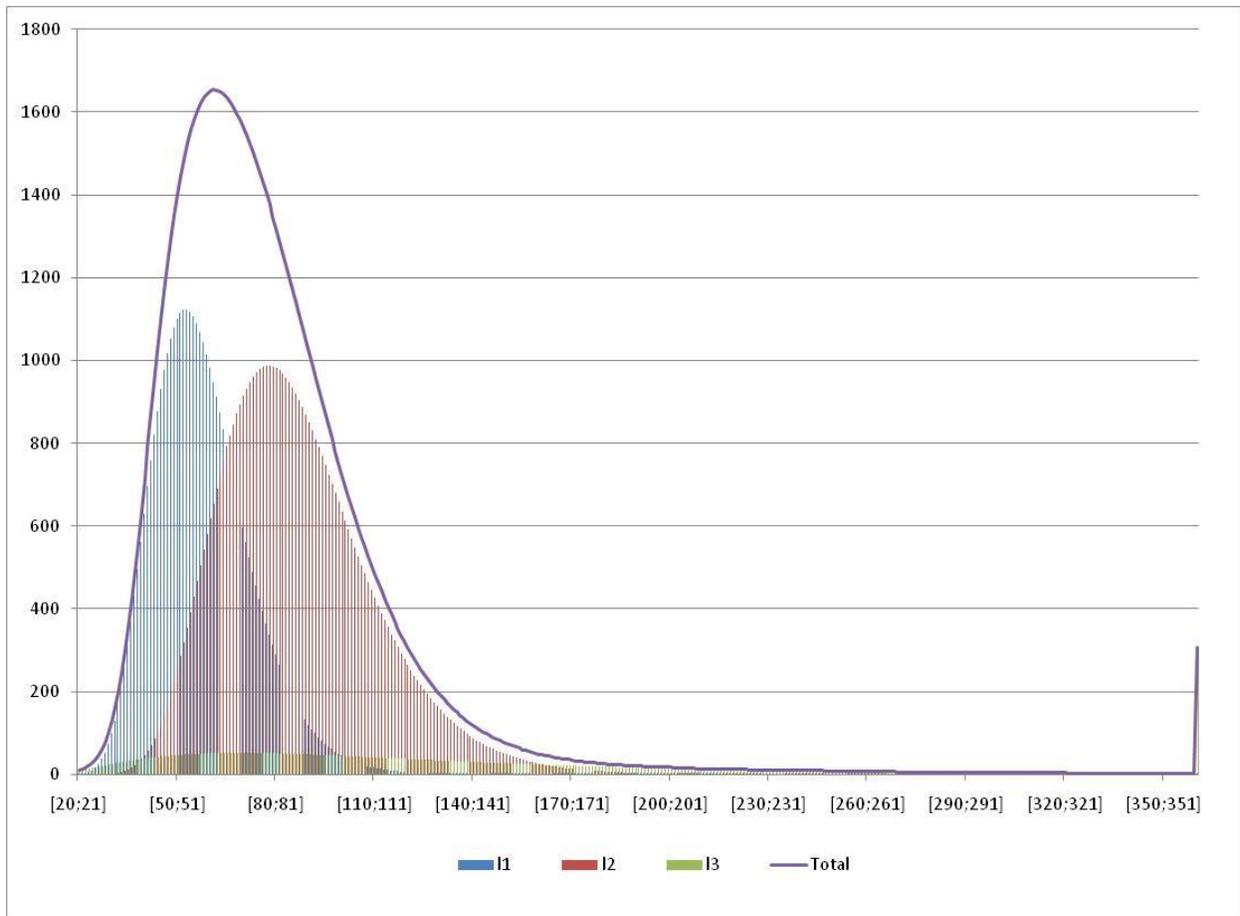


Figure 24. Distribution fréquentielle des temps de parcours du groupe AM-6

Tableau 6. Paramètres du modèle du groupe AM-6

	Loi 1	Loi 2	Loi 3
Moyenne	4.036	4.431	4.694
Écart-type	0.252	0.266	0.680
Lambda	0.389	0.534	0.077

Tableau 7. Définition des intervalles du groupe AM-6

Définition des intervalles			
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4
[24;65]	[65;93]	[93;142]	[142;360]

Dans ce cas, le premier intervalle correspondra à un temps de parcours usuel et optimiste, le second à un temps de parcours usuel mais pessimiste, ensuite, le troisième intervalle correspondra à un temps de parcours tampon, et le quatrième à un temps de parcours extrême.

Quatre groupes présentent ce type de distribution: AM-6, AM-7, PM-4 et PM-5.

### Distributions congestionnées

Les deux derniers groupes présentent une distribution fréquentielle des temps de parcours particulière : les valeurs des coefficients sont proches, et la distribution est très étendue. La classification est alors très simple : chaque intervalle sera défini comme l'intervalle de temps dans lequel une des trois lois dominant les deux autres. Nous rajouterons que ces groupes présentent de très nombreux relevés dont les valeurs excèdent la limite supérieure de 360s.

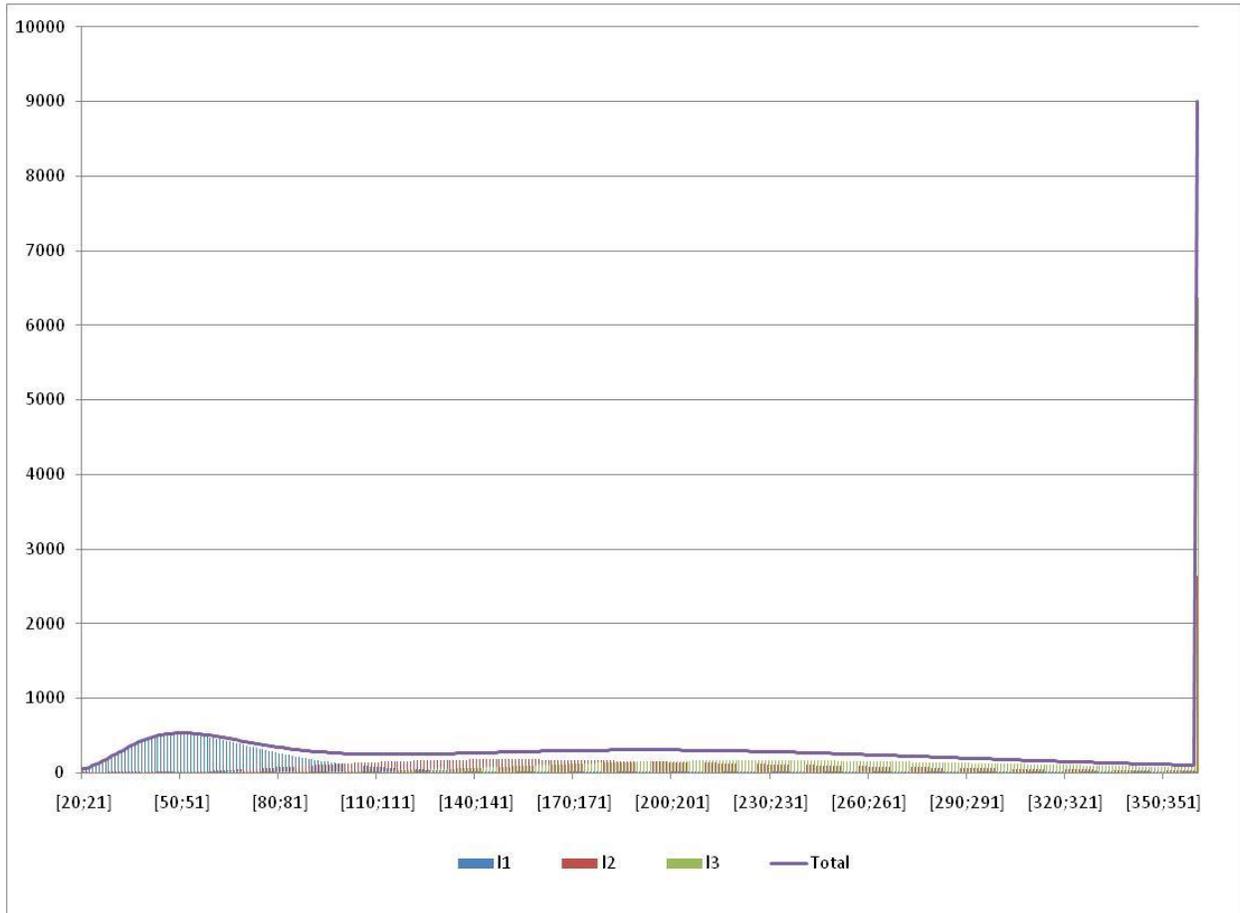


Figure 25. Distribution fréquentielle des relevés de temps de parcours du groupe AM-8

Tableau 8. Définition des intervalles du groupe AM-8

Définition des intervalles			
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4
[25;100]	[100;195]	[195;360]	No

**Tableau 9. Paramètres du modèle du groupe AM-8**

	Loi 1	Loi 2	Loi 3
Moyenne	4.044	5.219	5.553
Écart-type	0.410	0.466	0.357
Lambda	0.291	0.346	0.363

***Synthèse***

Les 16 groupes de tronçons (8 par période) ont été modélisés ainsi. Les paramètres de modèles développés sont synthétisés dans les tableaux qui suivent. Un dictionnaire permettant d'associer chaque tronçon d'un kilomètre aux groupes ainsi qu'aux numéros de relevés initiaux sera fourni selon les besoins.

**Tableau 10. Paramètres du modèle de distribution des temps de parcours AM-1**

Caractéristiques du groupe	Groupe1 durant la période AM		
	Loi 1	Loi 2	Loi 3
Moyenne	3.91	4.253	4.121
Ecart-type	0.139	0.247	0.855
Lambda	0.72	0.153	0.127
Moyenne (en seconde)	50.3833367	72.4940508	88.81117618
Vitesse Moyenne (en km/h)	71.452195	49.6592474	40.53543883
Ecart-type (en seconde)	7.03724798	18.182641	92.17593231
Moyenne du Groupe (en seconde)	<b>58.6466116</b>		
Ecart-type du Groupe (en seconde)	34.1362368		
Coefficient de Variation	<b>0.58206665</b>		
Vitesse Moyenne (en km/h)	61.3846206		
Pourcentage d'incidents (relevés >360s)	0.2%		
Définition des intervalles			
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4
[30;65]	[65;120]	[120;360]	[;]

**Tableau 11. Paramètres du modèle de distribution des temps de parcours AM-2**

Caractéristiques du groupe		Groupe2			durant la période AM		
		Loi 1	Loi 2	Loi 3			
Moyenne		3.514	3.748	4.108			
Ecart-type		0.107	0.149	0.811			
Lambda		0.828	0.08	0.092			
Moyenne (en seconde)		33.7751221	42.9098112	84.5092081			
Vitesse Moyenne (en km/h)		106.587327	83.8968967	42.59890823			
Ecart-type (en seconde)		3.62430678	6.42921242	81.51472612			
Moyenne du Groupe (en seconde)		<b>39.1734332</b>					
Ecart-type du Groupe (en seconde)		25.0098163					
Coefficient de Variation		<b>0.6384382</b>					
Vitesse Moyenne (en km/h)		91.899017					
Pourcentage d'incidents (relevés >360s)		0.2%					
<b>Définition des intervalles</b>							
Intervalle 1		Intervalle 2	Intervalle 3	Intervalle 4			
[27;42]		[42;90]	[90;360]	[;]			

**Tableau 12. Paramètres du modèle de distribution des temps de parcours AM-3**

Caractéristiques du groupe		Groupe3			durant la période AM		
	Loi 1	Loi 2	Loi 3				
Moyenne	3.7	3.911	4.433				
Ecart-type	0.124	0.207	0.612				
Lambda	0.757	0.145	0.098				
Moyenne (en seconde)	40.7594616	51.0305514	101.5216423				
Vitesse Moyenne (en km/h)	88.3230508	70.545975	35.46041927				
Ecart-type (en seconde)	5.07366386	10.6774978	68.4295098				
Moyenne du Groupe (en seconde)	<b>48.2034633</b>						
Ecart-type du Groupe (en seconde)	22.2466318						
Coefficient de Variation	<b>0.46151522</b>						
Vitesse Moyenne (en km/h)	74.6834304						
Pourcentage d'incidents (relevés >360s)	0.4%						
Définition des intervalles							
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4				
[26;52]	[52;90]	[90;360]	[;]				

**Tableau 13. Paramètres du modèle de distribution des temps de parcours AM-4**

Caractéristiques du groupe		Groupe4 durant la période AM		
	Loi 1	Loi 2	Loi 3	
Moyenne	3.57	3.765	5.404	
Ecart-type	0.111	0.202	0.766	
Lambda	0.349	0.198	0.453	
Moyenne (en seconde)	35.7360685	44.0533765	298.0847971	
Vitesse Moyenne (en km/h)	100.738558	81.7190483	12.07710033	
Ecart-type (en seconde)	3.97895346	8.99033494	266.305988	
Moyenne du Groupe (en seconde)	<b>156.22687</b>			
Ecart-type du Groupe (en seconde)	179.298024			
Coefficient de Variation	<b>1.14767725</b>			
Vitesse Moyenne (en km/h)	23.0434112			
Pourcentage d'incidents (relevés >360s)	11.0%			
Définition des intervalles				
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4	
[25;62]	[62;150]	[150;360]	[;]	

**Tableau 14. Paramètres du modèle de distribution des temps de parcours AM-5**

Caractéristiques du groupe		Groupe5 durant la période AM		
	Loi 1	Loi 2	Loi 3	
Moyenne	3.69	4.095	4.874	
Ecart-type	0.146	0.267	0.621	
Lambda	0.274	0.178	0.548	
Moyenne (en seconde)	40.4739274	62.2180095	158.6690363	
Vitesse Moyenne (en km/h)	88.9461495	57.8610603	22.68873678	
Ecart-type (en seconde)	5.94082379	16.91272	108.8424076	
Moyenne du Groupe (en seconde)	<b>109.115294</b>			
Ecart-type du Groupe (en seconde)	80.9478877			
Coefficient de Variation	<b>0.74185648</b>			
Vitesse Moyenne (en km/h)	32.9926253			
Pourcentage d'incidents (relevés >360s)	3.0%			
Définition des intervalles				
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4	
[28;57]	[57;120]	[120;180]	[180;360]	

**Tableau 15. Paramètres du modèle de distribution des temps de parcours AM-6**

Caractéristiques du groupe		Groupe6 durant la période AM		
	Loi 1	Loi 2	Loi 3	
Moyenne	4.036	4.431	4.694	
Ecart-type	0.252	0.266	0.68	
Lambda	0.389	0.534	0.077	
Moyenne (en seconde)	58.4254744	87.0408893	137.7168823	
Vitesse Moyenne (en km/h)	61.616958	41.3598715	26.14058597	
Ecart-type (en seconde)	14.9600875	23.5685295	105.5921533	
Moyenne du Groupe (en seconde)	<b>79.8115443</b>			
Ecart-type du Groupe (en seconde)	35.2450127			
Coefficient de Variation	<b>0.44160294</b>			
Vitesse Moyenne (en km/h)	45.1062566			
Pourcentage d'incidents (relevés >360s)	0.3%			
Définition des intervalles				
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4	
[25;87]	[87;105]	[105;165]	[165;360]	

**Tableau 16. Paramètres du modèle de distribution des temps de parcours AM-7**

Caractéristiques du groupe		Groupe7			durant la période AM
	Loi 1	Loi 2	Loi 3		
Moyenne	4.088	4.75	4.792		
Ecart-type	0.36	0.265	0.608		
Lambda	0.12	0.625	0.255		
Moyenne (en seconde)	63.6118644	119.71483	145.014248		
Vitesse Moyenne (en km/h)	56.5932163	30.0714623	24.82514683		
Ecart-type (en seconde)	23.6626682	32.289627	96.98063057		
Moyenne du Groupe (en seconde)	<b>119.433826</b>				
Ecart-type du Groupe (en seconde)	55.8315773				
Coefficient de Variation	<b>0.46746872</b>				
Vitesse Moyenne (en km/h)	30.1422145				
Pourcentage d'incidents (relevés >360s)	0.9%				
Définition des intervalles					
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4		
[27;70]	[70;165]	[165;210]	[210;360]		

**Tableau 17. Paramètres du modèle de distribution des temps de parcours AM-8**

Caractéristiques du groupe		Groupe8 durant la période AM		
	Loi 1	Loi 2	Loi 3	
Moyenne	4.077	5.219	5.553	
Ecart-type	0.41	0.466	0.357	
Lambda	0.291	0.346	0.363	
Moyenne (en seconde)	64.138833	205.938586	274.9871859	
Vitesse Moyenne (en km/h)	56.1282429	17.4809396	13.09151911	
Ecart-type (en seconde)	27.4417463	101.420922	101.3830186	
Moyenne du Groupe (en seconde)	<b>189.7395</b>			
Ecart-type du Groupe (en seconde)	86.6560168			
Coefficient de Variation	<b>0.45671047</b>			
Vitesse Moyenne (en km/h)	18.9733819			
Pourcentage d'incidents (relevés >360s)	9.0%			
Définition des intervalles				
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4	
[25;105]	[105;210]	[210;300]	[300;360]	

**Tableau 18. Paramètres du modèle de distribution des temps de parcours PM-1**

Caractéristiques du groupe		Groupe1			durant la période PM
	Loi 1	Loi 2	Loi 3		
Moyenne	3.669	5.014	5.352		
Ecart-type	0.14	0.812	0.333		
Lambda	0.16	0.571	0.269		
Moyenne (en seconde)	39.5988469	209.279776	223.0608268		
Vitesse Moyenne (en km/h)	90.9117381	17.2018532	16.13909557		
Ecart-type (en seconde)	5.57111463	202.204072	76.38681549		
Moyenne du Groupe (en seconde)	<b>185.83793</b>				
Ecart-type du Groupe (en seconde)	157.863074				
Coefficient de Variation	<b>0.84946638</b>				
Vitesse Moyenne (en km/h)	19.3717181				
Pourcentage d'incidents (relevés >360s)	9.8%				
Définition des intervalles					
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4		
[26;50]	[50;150]	[150;240]	[240;360]		

**Tableau 19. Paramètres du modèle de distribution des temps de parcours PM-2**

Caractéristiques du groupe		Groupe2			durant la période PM
	Loi 1	Loi 2	Loi 3		
Moyenne	3.948	4.088	4.51		
Ecart-type	0.141	0.298	0.703		
Lambda	0.511	0.38	0.109		
Moyenne (en seconde)	52.3494013	62.3274538	116.408377		
Vitesse Moyenne (en km/h)	68.7686948	57.7594588	30.92560942		
Ecart-type (en seconde)	7.41810472	18.9936648	93.0685606		
Moyenne du Groupe (en seconde)	<b>63.1234896</b>				
Ecart-type du Groupe (en seconde)	33.3067485				
Coefficient de Variation	<b>0.52764428</b>				
Vitesse Moyenne (en km/h)	57.0310676				
Pourcentage d'incidents (relevés >360s)	0.3%				
<b>Définition des intervalles</b>					
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4		
[26;75]	[75;105]	[105;360]	[;]		

**Tableau 20. Paramètres du modèle de distribution des temps de parcours PM-3**

Caractéristiques du groupe		Groupe3			durant la période PM
	Loi 1	Loi 2	Loi 3		
Moyenne	3.731	4.041	4.715		
Ecart-type	0.121	0.295	0.736		
Lambda	0.765	0.146	0.089		
Moyenne (en seconde)	42.0273459	59.4129669	146.3276082		
Vitesse Moyenne (en km/h)	85.658514	60.5928333	24.60232928		
Ecart-type (en seconde)	5.10397926	17.9151477	124.0699079		
Moyenne du Groupe (en seconde)	<b>53.8483699</b>				
Ecart-type du Groupe (en seconde)	37.9050818				
Coefficient de Variation	<b>0.70392255</b>				
Vitesse Moyenne (en km/h)	66.8543914				
Pourcentage d'incidents (relevés >360s)	0.6%				
Définition des intervalles					
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4		
[25;55]	[55;90]	[90;360]	[;]		

**Tableau 21. Paramètres du modèle de distribution des temps de parcours PM-4**

Caractéristiques du groupe		Groupe4			durant la période PM		
	Loi 1	Loi 2	Loi 3				
Moyenne	4.328	4.858	4.88				
Ecart-type	0.501	0.262	0.634				
Lambda	0.299	0.542	0.159				
Moyenne (en seconde)	85.9272063	133.262651	160.9313699				
Vitesse Moyenne (en km/h)	41.8959274	27.0143208	22.36978411				
Ecart-type (en seconde)	45.8976269	35.5226453	113.1961515				
Moyenne du Groupe (en seconde)	<b>123.508679</b>						
Ecart-type du Groupe (en seconde)	57.8888927						
Coefficient de Variation	<b>0.46870303</b>						
Vitesse Moyenne (en km/h)	29.1477491						
Pourcentage d'incidents (relevés >360s)	0.9%						
<b>Définition des intervalles</b>							
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4				
[25;75]	[75;165]	[165;210]	[210;360]				

**Tableau 22. Paramètres du modèle de distribution des temps de parcours PM-5**

Caractéristiques du groupe		Groupe5			durant la période PM		
		Loi 1	Loi 2	Loi 3			
Moyenne		4.206	4.53			4.75	
Ecart-type		0.299	0.263			0.594	
Lambda		0.469	0.431			0.1	
Moyenne (en seconde)		70.1545384	96.0226883			137.8847236	
Vitesse Moyenne (en km/h)		51.3152831	37.4911395			26.10876612	
Ecart-type (en seconde)		21.4538806	25.6970236			89.68844695	
Moyenne du Groupe (en seconde)		<b>88.0767295</b>					
Ecart-type du Groupe (en seconde)		36.1230282					
Coefficient de Variation		<b>0.41013135</b>					
Vitesse Moyenne (en km/h)		40.8734523					
Pourcentage d'incidents (relevés >360s)		0.3%					
<b>Définition des intervalles</b>							
Intervalle 1		Intervalle 2	Intervalle 3	Intervalle 4			
	[25;105]		[105;150]		[150;360]		[;]

**Tableau 23. Paramètres du modèle de distribution des temps de parcours PM-6**

Caractéristiques du groupe		Groupe6			durant la période PM
	Loi 1	Loi 2	Loi 3		
Moyenne	3.507	3.592	4.291		
Ecart-type	0.092	0.138	0.804		
Lambda	0.661	0.299	0.04		
Moyenne (en seconde)	33.4895014	36.6539794	100.9078778		
Vitesse Moyenne (en km/h)	107.496375	98.2158025	35.67610455		
Ecart-type (en seconde)	3.08756511	5.0824273	96.19055396		
Moyenne du Groupe (en seconde)	<b>37.1324154</b>				
Ecart-type du Groupe (en seconde)	19.5992282				
Coefficient de Variation	<b>0.52781991</b>				
Vitesse Moyenne (en km/h)	96.950332				
Pourcentage d'incidents (relevés >360s)	0.1%				
<b>Définition des intervalles</b>					
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4		
[24;40]	[40;75]	[75;360]	[;]		

**Tableau 24. Paramètres du modèle de distribution des temps de parcours PM-7**

Caractéristiques du groupe		Groupe7			durant la période PM
	Loi 1	Loi 2	Loi 3		
Moyenne	3.578	3.723	4.364		
Ecart-type	0.102	0.142	0.686		
Lambda	0.592	0.191	0.217		
Moyenne (en seconde)	35.988592	41.8077613	99.41450216		
Vitesse Moyenne (en km/h)	100.031699	86.1084135	36.2120206		
Ecart-type (en seconde)	3.68040496	5.96675512	77.06695783		
Moyenne du Groupe (en seconde)	<b>50.8634759</b>				
Ecart-type du Groupe (en seconde)	36.1060999				
Coefficient de Variation	<b>0.70986301</b>				
Vitesse Moyenne (en km/h)	70.7777032				
Pourcentage d'incidents (relevés >360s)	0.3%				
Définition des intervalles					
Intervalle 1	Intervalle 2	Intervalle 3	Intervalle 4		
[24;50]	[50;90]	[90;360]	[;]		

**Tableau 25. Paramètres du modèle de distribution des temps de parcours PM-8**

Caractéristiques du groupe		Groupe8			durant la période PM		
		Loi 1	Loi 2	Loi 3			
Moyenne		3.477	3.619				3.895
Ecart-type		0.105	0.177				0.922
Lambda		0.763	0.159				0.078
Moyenne (en seconde)		32.5413797	37.889139				75.19178628
Vitesse Moyenne (en km/h)		110.628376	95.0140357				47.87756985
Ecart-type (en seconde)		3.42628421	6.75924807				87.03569894
Moyenne du Groupe (en seconde)		<b>36.7184051</b>					
Ecart-type du Groupe (en seconde)		24.6391576					
Coefficient de Variation		<b>0.67103017</b>					
Vitesse Moyenne (en km/h)		98.0434741					
Pourcentage d'incidents (relevés >360s)		0.0%					
<b>Définition des intervalles</b>							
Intervalle 1		Intervalle 2	Intervalle 3	Intervalle 4			
[26;40]		[40;75]	[75;360]	[;]			

### 3. Étude des potentialités des différentes technologies de collecte de données

---

Un des objectifs émergents du projet est celui d'évaluer les potentialités et l'intérêt de différentes technologies de collecte de données sur les temps de parcours et/ou les vitesses en vue d'outiller le MTQ en matière de monitoring différé et modélisation des conditions de circulation sur le réseau routier. Dans les récentes années, plusieurs technologies et produits permettant soit de recueillir des données, soit d'obtenir des indicateurs, sont apparus. Les sections qui suivent visent à comprendre la pertinence et la qualité de certaines données et produits issus de diverses nouvelles technologies.

#### 3.1 Sources de données et analyse descriptive

Afin d'évaluer différentes sources de données et produits disponibles, une analyse comparative de mesures (temps de parcours et/ou vitesse) est conduite. Les bases de données examinées sont :

- Les données GPS des véhicules de la compagnie Communauto inc., compagnie qui offre les services d'autopartage au Québec (la plus ancienne compagnie d'autopartage en Amérique du Nord et actuellement une des plus importantes sur ce territoire);
- Un extrait de données reçu de la compagnie INRIX qui couvre certaines grandes autoroutes de la région montréalaise;
- Des données de vitesses moyennes extraites de bandes vidéo par la compagnie RealTraffic.

En outre, des analyses comparatives ponctuelles de données ont aussi été conduites (GPS, bandes vidéo, Bluetooth); celles-ci sont résumées ainsi que disponibles dans Saunier et Morency (2011).

##### 3.1.1 Données GPS des véhicules de Communauto inc.

---

Depuis quelques années, la compagnie Communauto inc. équipe une partie de sa flotte de véhicules de systèmes GPS en vue d'assurer leur localisation lors de circonstances particulières et de faire la validation des distances parcourues par les usagers. Actuellement, c'est grosso modo 400 véhicules qui sont équipés de tels systèmes. Ces véhicules sont dispersés dans les différentes stations du réseau montréalais et sont utilisés par de multiples membres du service, au gré de leurs besoins.

Une expérience de mesure a été tentée afin d'évaluer si ces données permettent de fournir un portrait suffisant de l'état de la circulation sur les principaux liens autoroutiers de la région de Montréal. Les traces des véhicules équipés pour l'année 2010 ont été rendues disponibles par Communauto à cette fin. A priori, ces données n'ont pas comme objectif de mesurer l'évolution des conditions de circulation sur un réseau routier et ont plutôt une utilité opérationnelle pour la compagnie. Par conséquent, le niveau de résolution temporelle des données actuellement stockées est de l'ordre de 2-5 minutes (selon les mois). En fait, suite à une requête de notre part, le niveau de résolution des données a été amélioré et ce afin de nous permettre d'évaluer leur pertinence dans un contexte d'analyse stratégique régionale.

##### Base de données brute

Les traces GPS des véhicules de Communauto sont disponibles pour les mois de janvier à novembre de l'année 2010. Le fichier complet de points compte près de 20 millions d'enregistrements. Bien évidemment, les cheminements des déplacements des usagers de Communauto ne s'effectuent pas tous sur le réseau autoroutier montréalais ou même à Montréal. En outre, comme dans tout ensemble de données, certaines données sont erronées et inutilisables. Il est donc nécessaire de faire un certain

filtrage des données. De plus, le niveau de résolution des données n'est pas suffisant pour permettre une reconnaissance simple des itinéraires détaillés empruntés par les conducteurs. Ces points fournissent plutôt une information de vitesse instantanée sur le réseau. Par conséquent, une méthodologie de couplage a été développée afin d'associer les points de vitesse instantanée aux tronçons d'un kilomètre exploités à l'étape de modélisation.

Le tableau suivant (Tableau 26) résume le contenu de la base de données complète, épurée des enregistrements contenant une information de vitesse nulle. Bien qu'il soit possible qu'une vitesse nulle soit enregistrée pour un véhicule sur le réseau routier, cette possibilité est nettement moindre lorsqu'on considère le réseau supérieur (autoroutes). On y dénombre le nombre de points avec vitesse non-nulle par mois, pour tout l'univers spatial de points, pour ceux localisés à l'intérieur du territoire de l'enquête OD 2008 ainsi que ceux qui ont été couplés à un tronçon du réseau typiquement échantillonné à l'aide de véhicules flottants.

**Tableau 26. Nombre d'observations par type de jour et période de la journée**

Nombre d'enregistrements par mois (vitesse>0)				
	Univers		Réseau typique d'échantillonnage	
	Nb. Enr.	Nb.Veh.	Nb. Enr.	Nb.Veh.
Janvier	367 325	381	41 533	376
Février	353 745	391	43 766	381
Mars	387 719	385	47 232	370
Avril	378 366	387	47 236	374
Mai	501 069	386	70 483	371
Juin	397 550	392	52 319	374
Juillet	477 656	384	56 483	369
Août	448 521	367	55 344	347
Septembre	364 264	364	45 390	344
Octobre	373 668	391	45 138	380
Novembre	431 259	408	45 083	380
<b>Total</b>	<b>4 481 142</b>	<b>420</b>	<b>550 007</b>	<b>408</b>

### Méthodologie de couplage avec les tronçons d'un kilomètre

Afin d'évaluer l'ampleur de l'échantillon de points disponibles pour documenter les vitesses sur le réseau routier, une méthodologie de couplage des points et des tronçons d'un kilomètre a été développée et appliquée systématiquement à la base de données de Communauto. Cette méthodologie implique les étapes suivantes :

- 1) Conversion des données en MTM nad83 pour l'ensemble de l'univers de points. Une automatisation par programme a été effectuée afin de transformer l'ensemble des points en lot.
- 2) Segmentation de la couche de tronçons en 2 pour avoir unicité de possibilité par lien (pour tenir compte des deux directions et de l'imprécision spatiale des localisations spatiales).

- 3) Association par proximité spatiale entre les points (GPS) et les tronçons d'un km. Pour ce faire, la distance entre le point et le tronçon le plus proche, à angle droit (perpendiculaire), a été calculée ainsi que l'adéquation des directions.
- 4) Extraction des points respectant deux critères :
  - a. distance point-tronçon inférieure à 20 m
  - b. différence d'angle inférieure à 30 degrés
- 5) Report des vitesses GPS sur les tronçons

### Synthèse de l'échantillon de données disponible

La méthodologie de couplage de données permet d'obtenir l'échantillon de points utiles pour l'évaluation des conditions de circulation sur l'univers de tronçons de 1 km typiquement échantillonnés par le MTQ (généralement à l'aide de véhicules flottants).

### Nombre d'observations disponibles par mois

En moyenne, chaque tronçon est couplé à quelque 715 données de vitesse instantanée sur une période de 11 mois, les plus échantillonnés recevant jusqu'à 6863 points. Au niveau mensuel, le nombre moyen de points par tronçon varie entre 55 et 92. Évidemment, la variabilité d'échantillonnage des tronçons est très grande et le niveau d'échantillonnage varie aussi beaucoup par type de jour.

La Figure 26 présente le nombre d'observations (points de vitesse instantanée) disponibles par mois selon le type de jour. Avec le niveau actuel de collecte de données (4-5 minutes entre chaque point) des véhicules de Communauto, un mois de données correspond à quelque 40 000 - 55 000 points. En outre, le pic observé au mois de mai est dû à une réduction soudaine de l'intervalle de temps entre deux observations successives, ce qui permet déjà d'évaluer la richesse potentielle accrue de cet univers de données découlant uniquement d'une modification aux paramètres de stockage des données. Un potentiel à ne pas négliger lors de l'étude des échantillons actuellement disponibles.

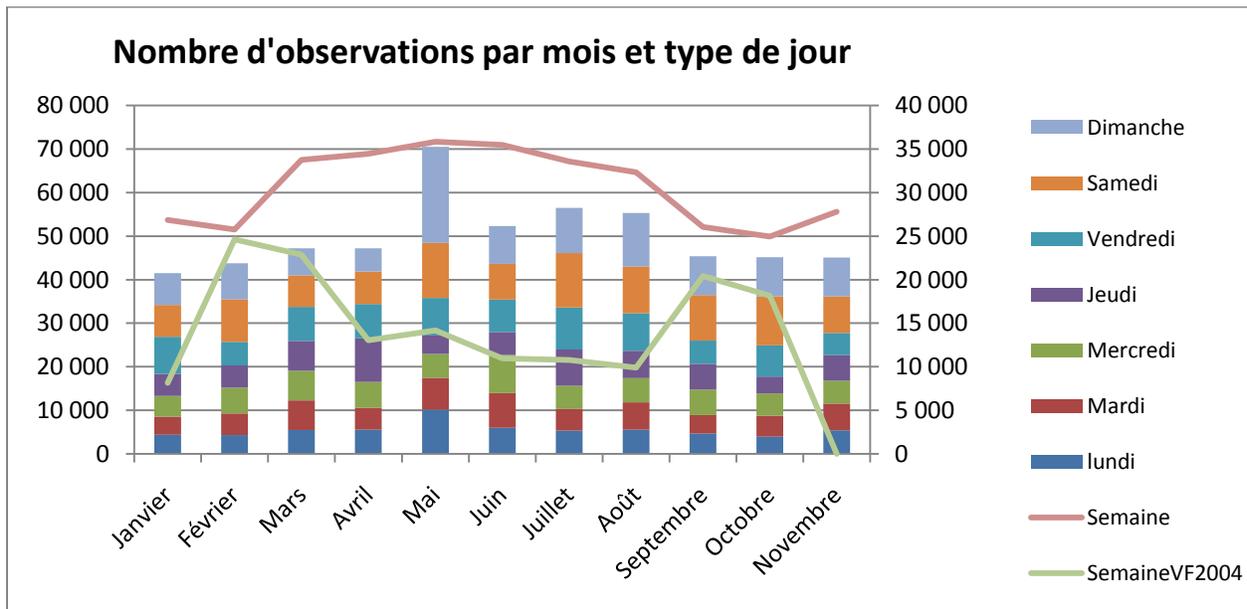


Figure 26. Distribution du nombre d'observations disponibles par mois et par type de jour et comparaison des échantillons de semaine par rapport à l'échantillon de véhicules flottants de 2004

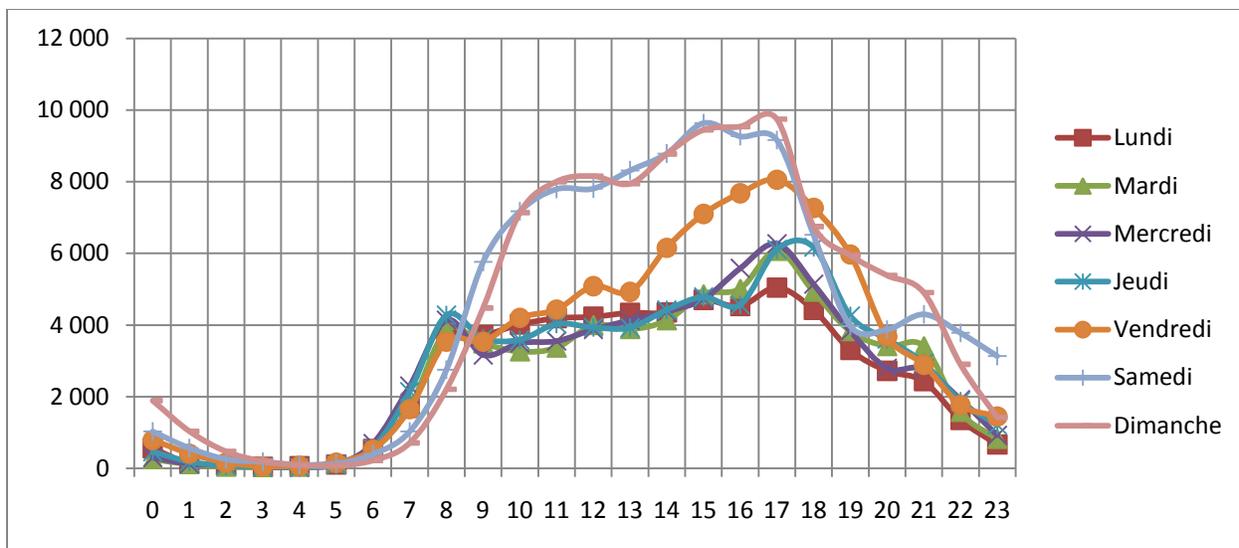
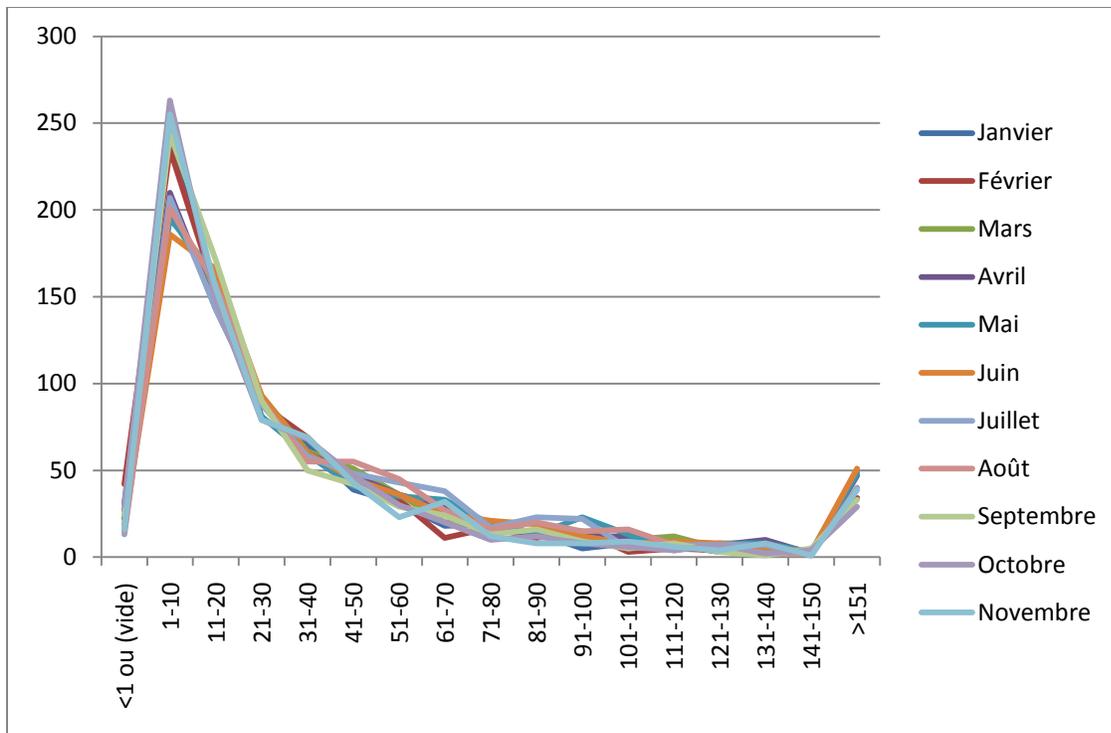


Figure 27. Distribution temporelle du nombre d'observations disponibles par heure et par type de jour, tous mois confondus

Tableau 27. Nombre d'observations par type de jour et période de la journée

Nombre d'observations par période et type de jour					
	<6h	6h-9h	9h-15h	15h-18h	>18h
Lundi	1000	5865	24 858	14 275	14 913
Mardi	683	6088	22 153	15 951	17 995
Mercredi	719	7132	22 612	16 585	17 407
Jeudi	922	6984	23 537	15 461	20 066
Vendredi	1663	5718	28 331	22 832	23 032
Samedi	2287	4190	45 644	28 070	25 589
Dimanche	3772	3147	44 490	28 727	27 309
Observations par heure					
Mardi	200	1955	4143	4758	2486
Mercredi	137	2029	3692	5317	2999
Jeudi	144	2377	3769	5528	2901
Vendredi	184	2328	3923	5154	3344
Samedi	333	1906	4722	7611	3839
Dimanche	457	1397	7607	9357	4265



**Figure 28. Distribution fréquentielle du nombre d'observations de semaine par tronçon, pour les différents mois de 2010**

Le graphique précédent confirme la stabilité des échantillonnages des différents tronçons mais ne permet pas de savoir si ce sont toujours les mêmes tronçons qui sont peu ou pas échantillonnés. Afin de documenter ceci, les tronçons sont maintenant caractérisés en termes de nombre de semaines avec jours (de semaine) observés. Le nombre maximum de semaines échantillonnées est de 49.

Globalement, c'est 39.5% des tronçons qui ont au moins une observation pour chacune des 49 semaines. La distribution complète est présentée à la Figure 29.

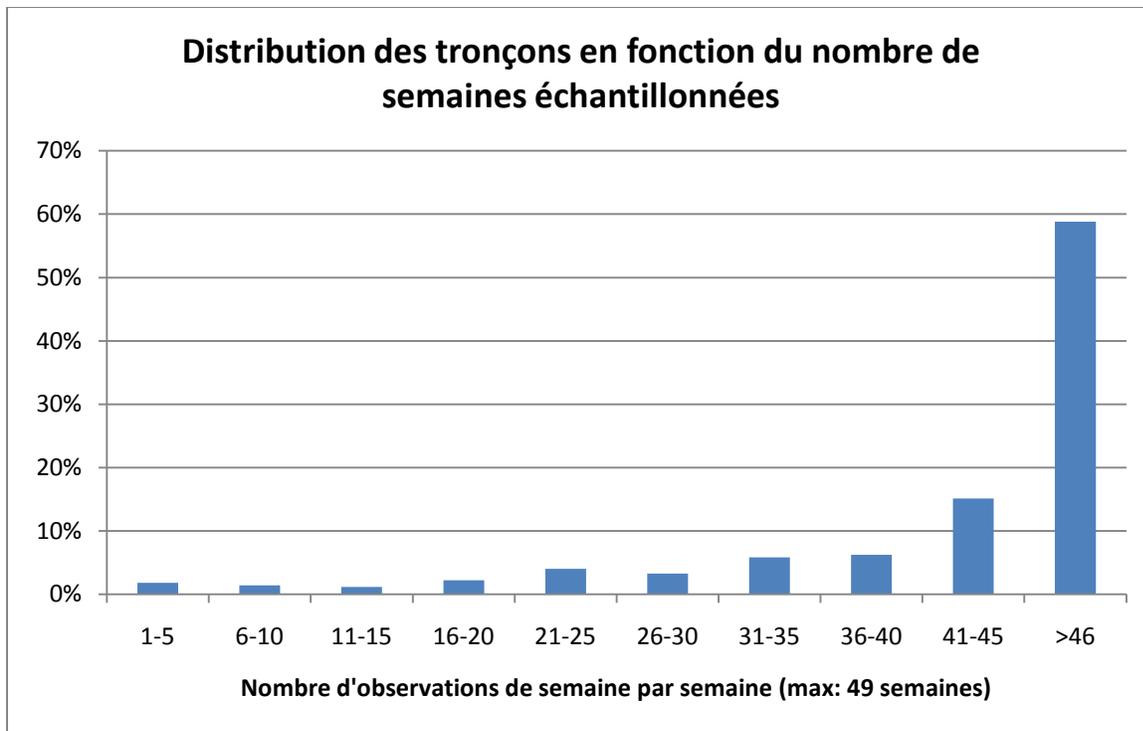


Figure 29. Distribution des tronçons (en %) en fonction du nombre de semaines échantillonnées (sur un maximum de 49).

Finalement, une cartographie des tronçons selon les niveaux d'échantillonnage de semaine est proposée.

#### Comparaison avec l'échantillon de données recueillies en 2004 avec les véhicules flottants

En vue de mieux évaluer les potentialités des données GPS de fournir une information suffisante pour mesurer les conditions de circulation et leur évolution dans le temps, nous proposons une comparaison avec l'échantillon de données recueillies par les véhicules flottants en 2004. Au début du projet, 10 mois de relevés de véhicules flottants ont été examinés pour l'année 2004. Typiquement, les véhicules suivent des routes prédéterminées afin de recueillir des données sur les temps de parcours sur ces itinéraires.

Avant d'amorcer l'analyse, il convient d'expliquer ce que nous entendons par « observation ». Dans le cas des données de GPS, la notion d'observation est simple et réfère à une donnée de vitesse instantanée. Dans le cas des véhicules flottants, une observation réfère à un temps de parcours par kilomètre (un passage de véhicule sur un tronçon d'un kilomètre). Dans ce dernier cas, les temps de parcours recueillis sur les différents tronçons, lors d'une même route ne sont pas nécessairement indépendants mais sont considérés comme tels pour les fins de la comparaison. Cela demeure toutefois un inconvénient.

Le Tableau 28 résume le nombre de relevés de véhicules flottants effectués de janvier à octobre en 2004.

Tableau 28. Nombre de relevés effectués par véhicules flottants en 2004

Nb. relevés	Janvier	Février	Mars	Avril	Mai	Juin	Juillet	Août	Septembre	Octobre	Novembre	Décembre	Total
2004	451	1065	868	577	568	524	467	436	799	712			6467

Les figures suivantes comparent les univers de données issues des données GPS et des véhicules flottants de 2004, pour chaque mois et pour chaque jour de la semaine (seuls ces jours sont échantillonnés avec les véhicules flottants). On observe que, de façon générale, les traces GPS fournissent au moins autant de points que les véhicules flottants de 2004, la différence étant très marquée pendant les mois d'avril à août. Une courbe du ratio entre le nombre d'observations GPS et le nombre d'observations VF (véhicules flottants) facilite la comparaison. Tel que mentionné précédemment, il ne faut pas négliger le fait que le niveau d'échantillonnage offert par les données GPS de Communauto pourrait facilement être augmenté uniquement en modifiant les paramètres de stockage des données (1 point par minute ou 30 secondes par exemple).

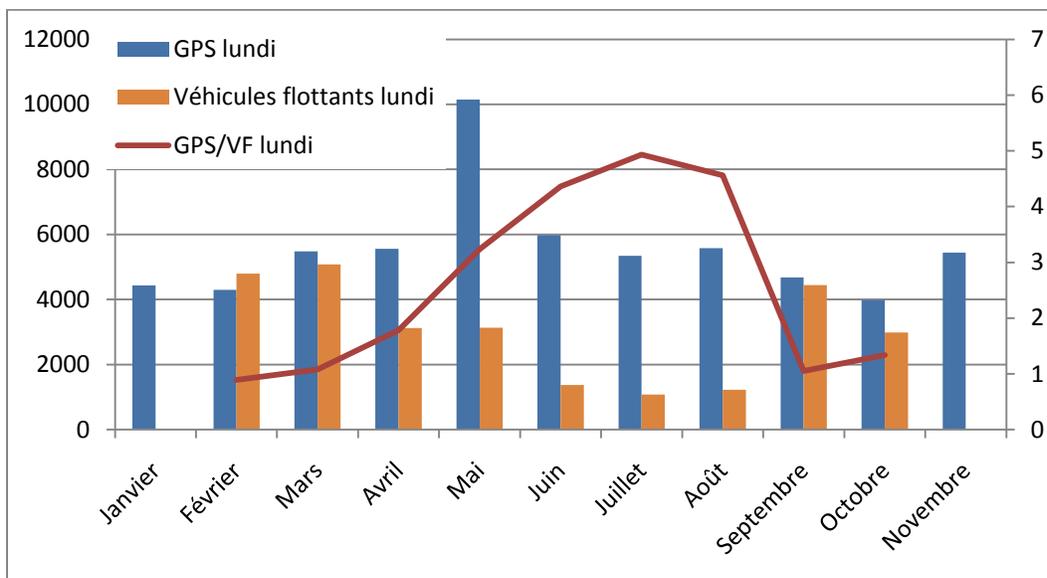


Figure 30. Comparaison des échantillons de données GPS et VF disponibles par mois, le lundi

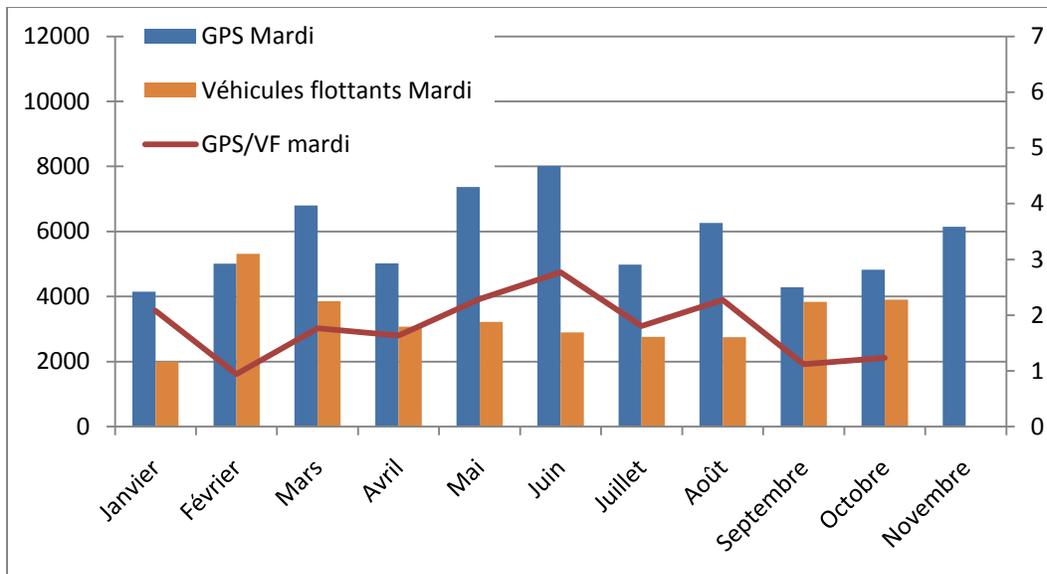


Figure 31. Comparaison des échantillons de données GPS et VF disponibles par mois, le mardi

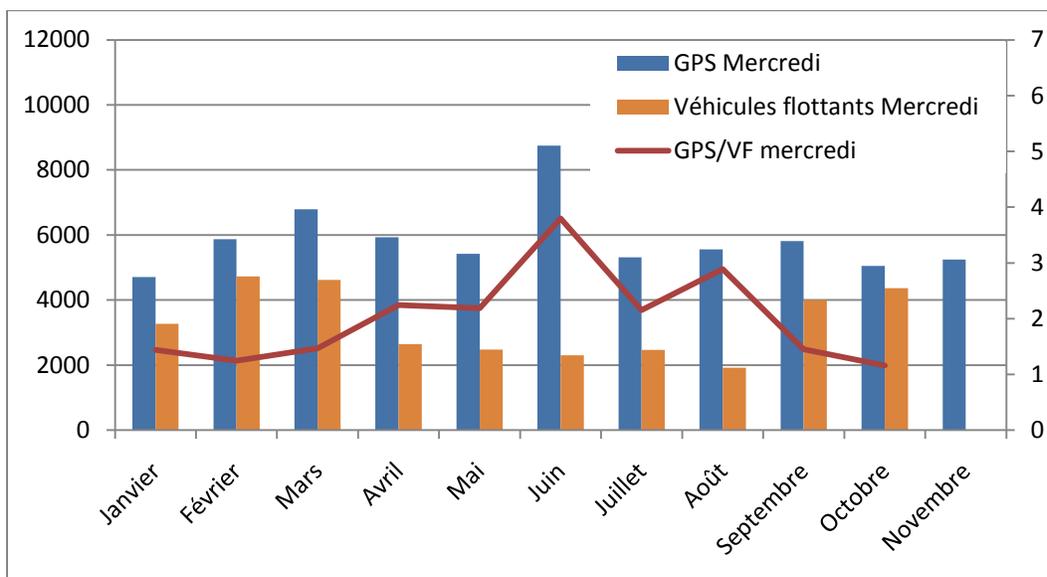


Figure 32. Comparaison des échantillons de données GPS et VF disponibles par mois, le mercredi

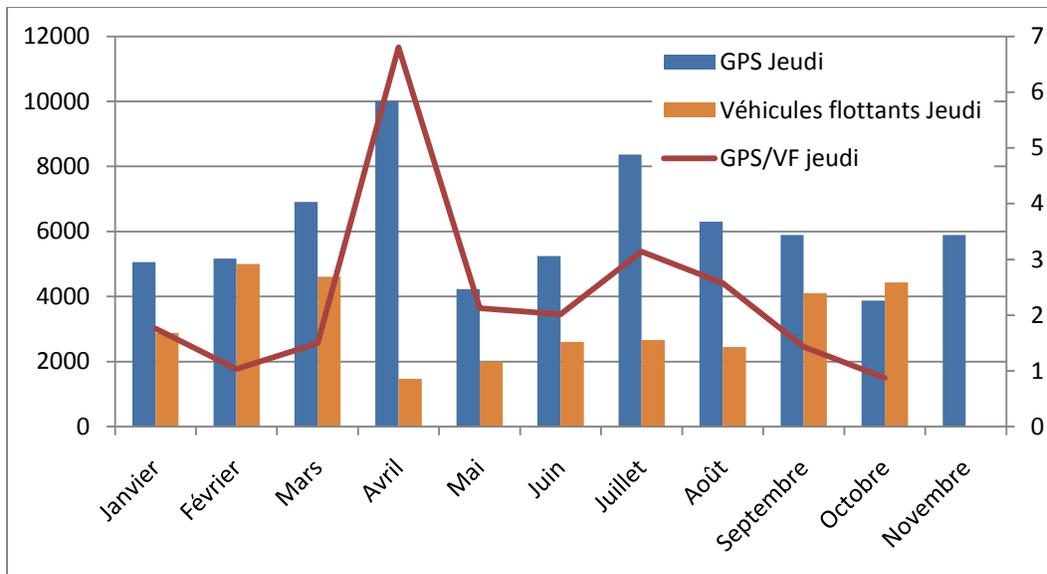


Figure 33. Comparaison des échantillons de données GPS et VF disponibles par mois, le jeudi

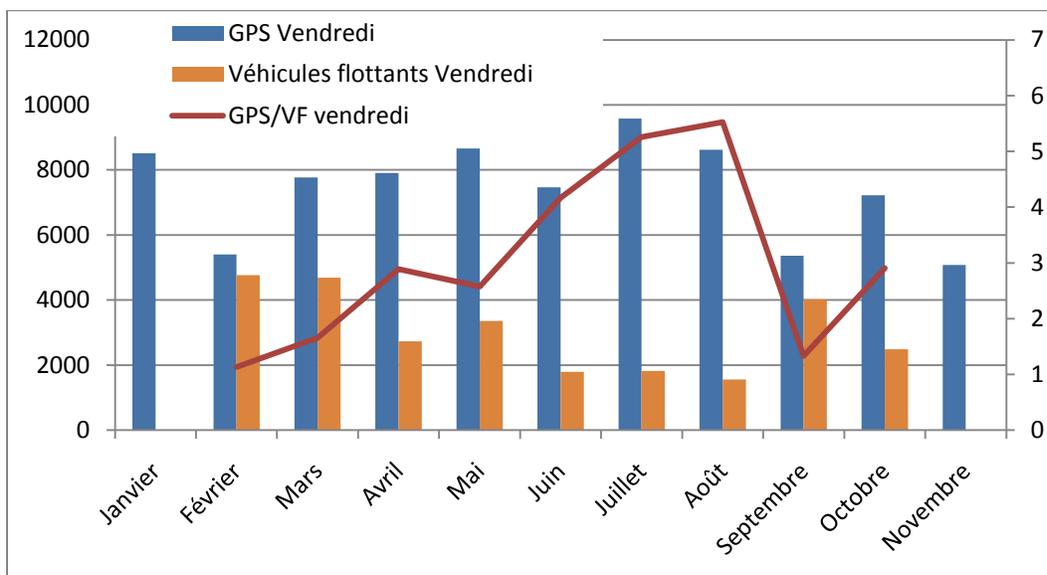


Figure 34. Comparaison des échantillons de données GPS et VF disponibles par mois, le vendredi

Une comparaison des données disponibles selon l'heure de la journée, pour certains mois, a aussi été conduite et est résumée dans les trois prochaines figures (Février : Figure 35, Mai : Figure 36 et Septembre : Figure 37) Tel qu'anticipé, les taux d'échantillonnage offerts par les véhicules de Communauto sont plus faibles en pointe AM qu'en pointe PM et inférieurs à ceux fournis par les véhicules flottants. Par contre, les traces GPS fournissent des données pour toutes les périodes de la journée, ce qui est sans doute en phase avec les tendances actuelles de modélisation qui portent sur des journées entières (modélisation 24 heures) ainsi que sur plusieurs types de jours (ce qui témoigne aussi de l'intérêt de disposer d'échantillonnage lors des jours de fin de semaine).

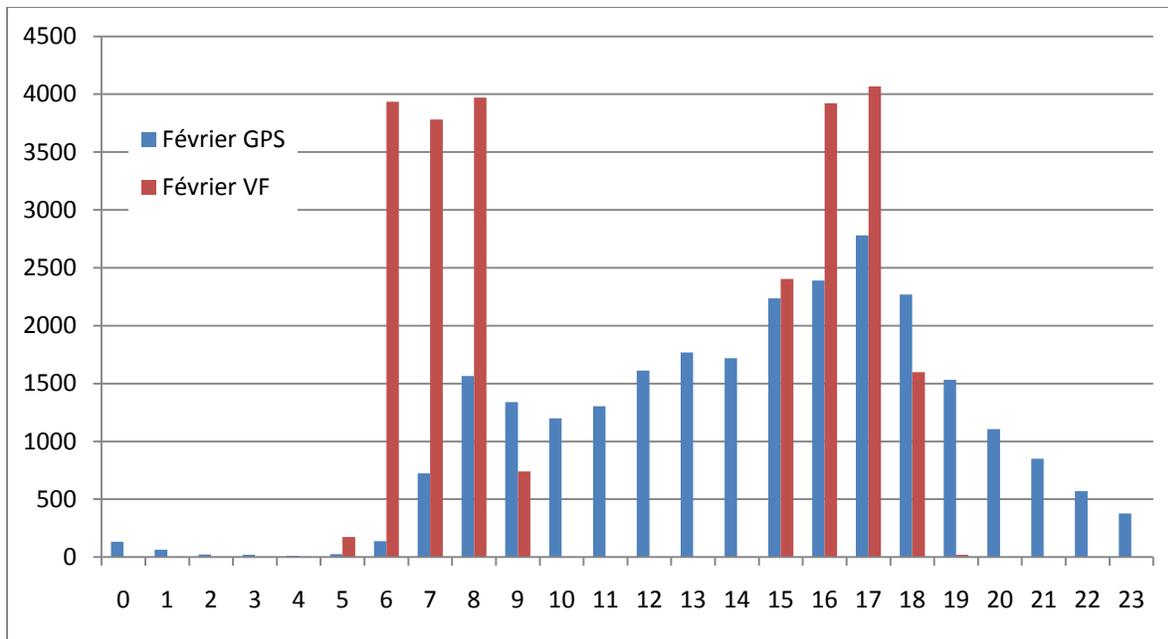


Figure 35. Comparaison des distributions temporelles des échantillons de données (jours de semaine) GPS et VF, mois de février

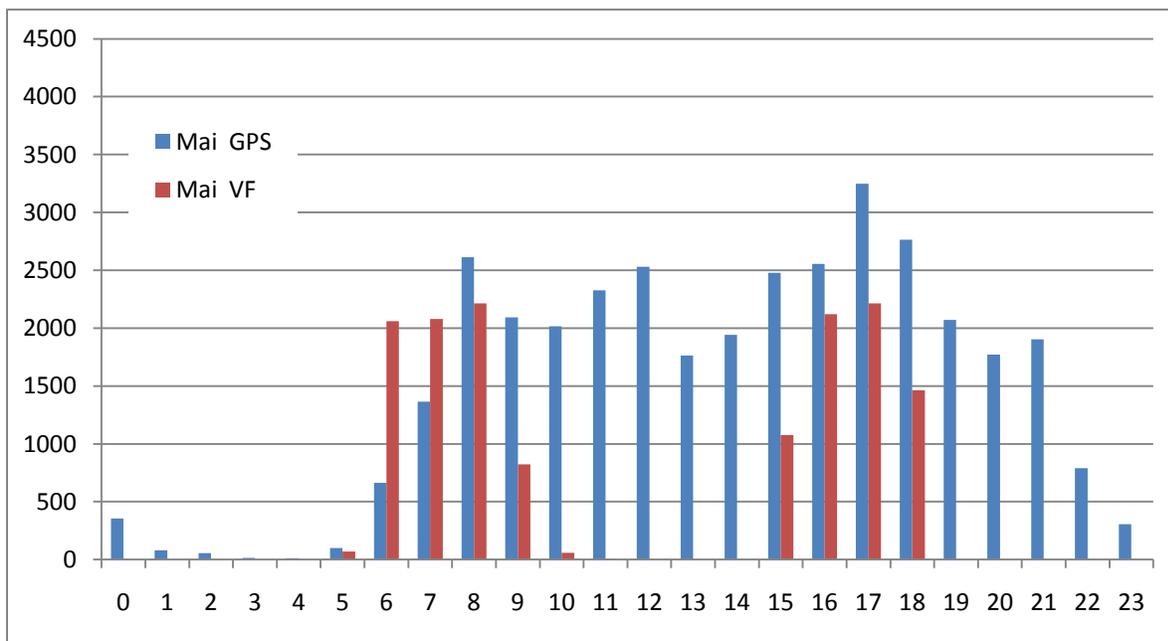
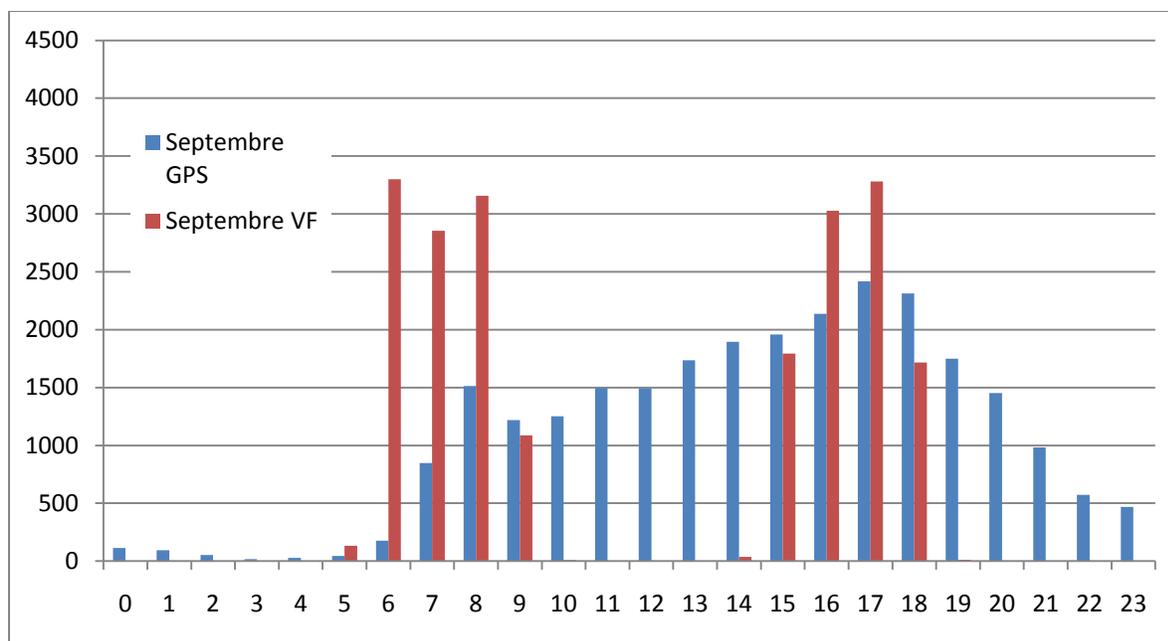


Figure 36. Comparaison des distributions temporelles des échantillons de données (jours de semaine) GPS et VF, mois de mai



**Figure 37. Comparaison des distributions temporelles des échantillons de données (jours de semaine) GPS et VF, mois de septembre**

Une comparaison de la répartition des échantillons GPS et VF, au cours de l'année, est ensuite effectuée. Dans le cas des véhicules flottants, l'échantillonnage de 2004 s'est amorcé à la semaine 3 et s'est étendu jusqu'à la semaine 44, pour une amplitude de 42 semaines. Toutefois, seulement 34 semaines sont incluses dans le plan d'échantillonnage (semaines pour lesquelles au moins un tronçon a été échantillonné). Dans le cas des GPS, l'échantillonnage couvre une période de 49 semaines et toutes ces semaines ont un échantillon quelque part.

Le tableau qui suit résume la répartition de l'échantillon de VF sur les différents tronçons. Il appert donc que près du tiers des tronçons de l'univers étudié ne faisaient pas partie du plan d'échantillonnage de 2004 (1 semaine échantillonnée) et que presque 50% des tronçons ont été échantillonnés sur au moins 18 semaines.

**Tableau 29. Nombre de relevés effectués par véhicules flottants en 2004**

Répartition de l'échantillonnage des VF 2004 sur l'année			
Nombre de semaines avec des données	% des semaines sur l'amplitude maximale (42 semaines)	Nombre de tronçons uniques	% des tronçons
1	2.4%	193	29.5%
2	4.8%	10	1.5%
5	11.9%	136	20.8%
18	42.9%	34	5.2%
30	71.4%	267	40.8%
31	73.8%	15	2.3%

### 3.1.2 Données de vitesses moyennes par segment RealTraffic

---

La société RealTraffic est une nouvelle compagnie de la région de Québec qui développe un logiciel d'analyse en temps réel des données vidéo des caméras de circulation, testé en particulier sur les caméras du réseau du MTQ. Les données produites sont la vitesse et le débit pour une section de route, ainsi que par voie. Dans le cadre de ce projet, nous avons eu accès à ces données pour la région de Montréal pendant le mois d'août. Il est à noter que le calibrage du modèle semble être perfectible pour l'ensemble de données disponible et que, bien évidemment, les résultats s'en ressentent.

#### Description de la base de données

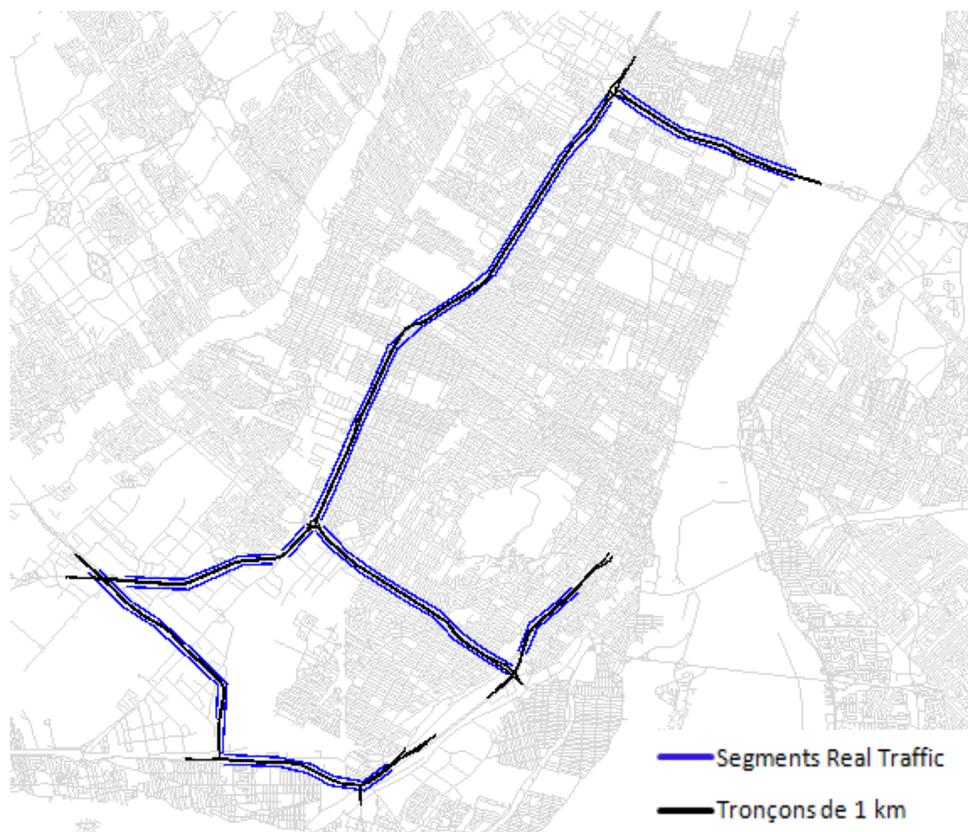
Les données de RealTraffic auxquelles nous avons eu accès portent sur le mois d'août 2010 (du 1er au 31 inclusivement). Le fichier complet de données compte près de 4.5 millions d'enregistrements pour l'ensemble des 96 segments autoroutiers de la région de Montréal. Les données disponibles sont, pour chacun des segments et pour un intervalle de temps d'environ une minute, le débit et la vitesse moyenne de parcours des véhicules, dont le temps de parcours est dérivé. Les autoroutes possédant plusieurs voies de circulation, les données fournies correspondent à la moyenne des voies observées. Pour chacun des segments, nous obtenons 44 570 observations de débit et de vitesse.

#### Méthodologie de couplage avec les tronçons d'un kilomètre

Cette méthodologie a été utilisée systématiquement afin de coupler les segments de RealTraffic avec les tronçons de 1 km.

1. Conversion en MTM (NAD83) des extrémités des segments fournis
2. Calcul des centroïdes des segments RealTraffic et des tronçons de 1 km
3. Calcul de la direction de chacun des segments (azimut)
4. Association par proximité spatiale entre les centroïdes des deux ensembles de segments avec correspondance de la direction des tronçons en respectant 3 critères :
  1. Distance segment RT – segment 1 km < minimale
  2. Distance segment RT – segment 1 km < 1.5 km
  3. Direction segment RT – segment 1 km < 30 degrés

La juxtaposition des segments de RealTraffic et de ceux d'un kilomètre exploités précédemment ne correspond pas. La longueur des 96 segments varie de 550 m à environ 2200 m et ceux-ci peuvent être associés à plusieurs tronçons de 1 km.



**Figure 38. Comparaison des tronçons de 1 km et des segments RealTraffic**

### Synthèse de l'échantillon de données disponible

La méthodologie de couplage se traduit par la disponibilité de l'échantillon de données suivant :

- 96 tronçons routiers sur une période d'un mois (1 au 31 août 2010)
- 4.25 millions d'enregistrements
- Intervalles de temps de 1 minute

### **3.1.3 Données de temps de parcours INRIX**

Les informations sur les données fournies par les services d'INRIX proviennent du guide technique public développé dans le cadre du « I-95 Corridor Coalition Vehicle Probe Project » (Haghani et al. 2009). Ces données de circulation sont le résultat de l'agrégation de différentes sources de données :

- véhicules et appareils équipés de capteurs GPS : environ un million de véhicules flottants, incluant des véhicules particuliers et commerciaux comme des taxis, limousines, navettes d'aéroport, véhicules de livraison, camions de transport longue distance;
- capteurs de circulation (plus de 90 % des capteurs de circulation disponibles aux États-Unis);
- autres sources de données de circulation : téléphones cellulaires, transpondeurs pour péages;

- métadonnées de circulation : événements, temps (phénomènes météorologiques), calendriers scolaires;
- autres données : données variables spatialement et temporellement comme le prix des carburants, les commerces locaux, les nouvelles, la bourse, le temps, les sports, le transport aérien, les temps d'attente aux frontières...

Pour chaque segment routier, correspondant à un code « TMC », les données suivantes sont fournies :

- *speed* : estimation actuelle de la vitesse moyenne spatiale sur le segment considéré (les vitesses sont mesurées en miles par heure);
- *average* : vitesse moyenne historique pour le segment considéré, pour le jour et l'heure de la semaine concernés;
- *reference* : vitesse libre moyenne sur le segment considéré, calculée à partir du 85<sup>ème</sup> centile des données observées (inférieure à 65 mi/h);
- *score* : facteur de confiance simple avec trois valeurs discrètes :
  - « 30 » : la confiance est élevée dans les résultats qui sont basés sur des données temps réel du segment concerné;
  - « 20 » : la confiance est moyenne dans les résultats qui sont basés sur des données temps réel de plusieurs segments et/ou basés sur une combinaison de données attendues et temps réel;
  - « 10 » : la confiance est basse dans les résultats qui sont basés essentiellement sur des données historiques.
- *travelTimeMinutes* : estimation actuelle du temps nécessaire pour parcourir le segment considéré (mesuré en minutes).

### Description de la base de données

Nous avons eu accès à près de 3 mois de données produites par la compagnie INRIX pour le réseau autoroutier de l'île de Montréal (du 13 août au 3 novembre 2010).

### Méthodologie de couplage avec les tronçons d'un kilomètre

Une méthodologie identique que celle utilisée précédemment pour RealTraffic a été systématiquement utilisée pour coupler les données d'INRIX. Comprenant les deux directions, les données fournies par INRIX contiennent 239 tronçons qui varient de 10 m à un peu plus de 3500 m.

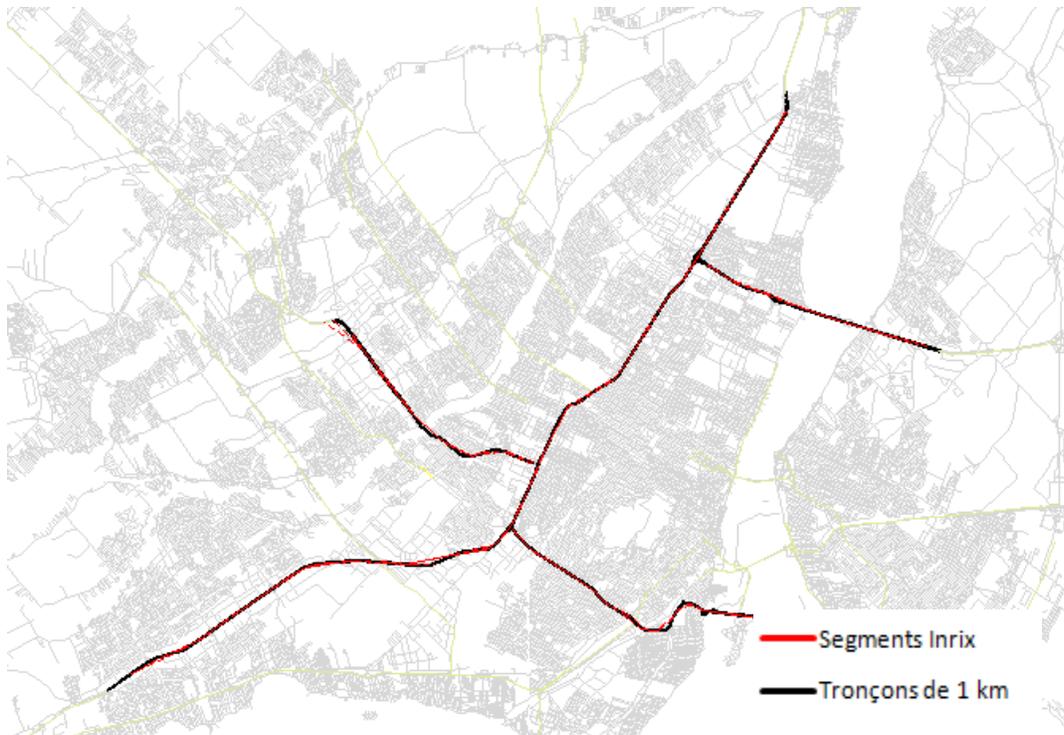


Figure 39. Comparaison des tronçons de 1 km et des segments INRIX

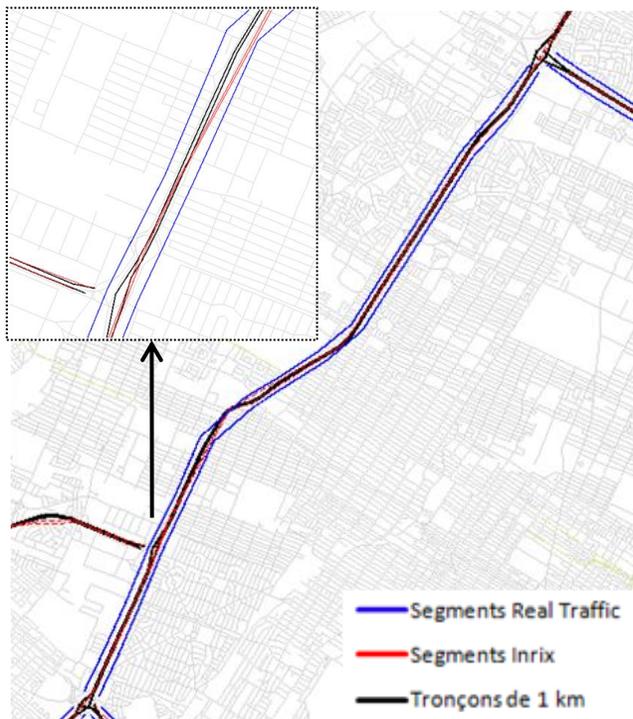


Figure 40. Comparaison des segments routiers pour l'autoroute métropolitaine (A40)

## 3.2 Mesures comparatives des conditions de circulation

Les trois sources de données permettent d'évaluer et décrire les conditions de circulation sur le réseau (ou certains tronçons) montréalais.

### 3.2.1 Indicateurs

---

Les indicateurs utilisés pour procéder à la comparaison sont :

- La vitesse moyenne (erreur moyenne et erreur absolue moyenne) et le coefficient de variation / le temps de parcours moyen par kilomètre et le coefficient de variation;
- Le % d'observations avec des vitesses inférieures à X km/h;
- Le TRI (« Travel rate index »).

### 3.2.2 Données GPS

---

Les données GPS sont maintenant exploitées afin d'évaluer leur capacité à révéler les conditions de circulation. Une autoroute est d'abord caractérisée puis l'ensemble des tronçons à l'étude.

#### État de la circulation sur l'A40 - Direction EST

Trois graphiques reflétant les vitesses moyennes sur les différents tronçons d'un kilomètre de l'A40 Est sont présentés :

- La Figure 41 présente la variation de la vitesse moyenne par tronçon lors des différents mois de l'année 2010 (uniquement les observations recueillies lors des jours de la semaine, excluant samedi et dimanche). D'abord, on constate que les données révèlent bien les zones bien connues de congestion, notamment entre l'A13 et l'A25 avec la vitesse la moindre observée à l'approche de l'A15 vers le SUD (Décarie). De grandes variabilités sont observées pour le mois de septembre qui est bien connu pour présenter des fluctuations plus importantes de la circulation.
- La Figure 42 présente la variation de la vitesse moyenne par période de la journée sur ces mêmes tronçons (uniquement les observations recueillies lors de jours de semaine, excluant samedi et dimanche). On constate que, dans la partie est du circuit, les vitesses sont inférieures lors de la période de pointe du soir (ce qui correspond à l'orientation des mouvements pendulaires) alors qu'elles sont inférieures et plus variables dans la section ouest en pointe AM.
- La Figure 43 présente la variation de la vitesse moyenne par jour de la semaine (toutes périodes et tous mois confondus). Il s'agit de comparaisons rarement conduites puisque les fins de semaine sont souvent omises des processus de collecte de données. On observe sur cette figure que les zones de forte congestion sont aussi problématiques les samedis et dimanches alors que la vitesse moyenne baisse à 40 km/h et 60km/h respectivement à l'approche de l'A15.

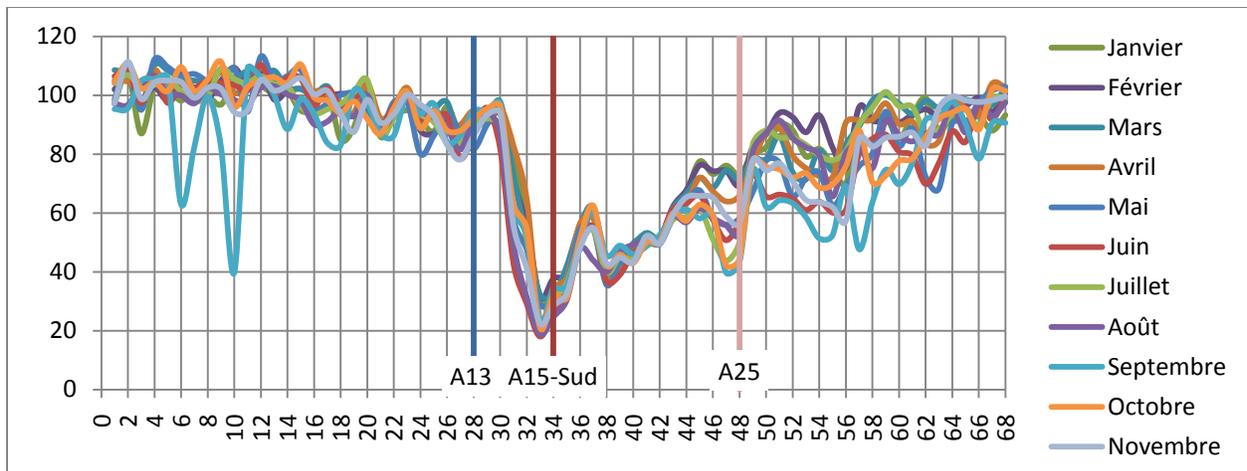


Figure 41. Variation de la vitesse moyenne (tous jours de semaine confondus, excluant samedi et dimanche) sur l'autoroute 40 (direction EST) lors des différents mois.

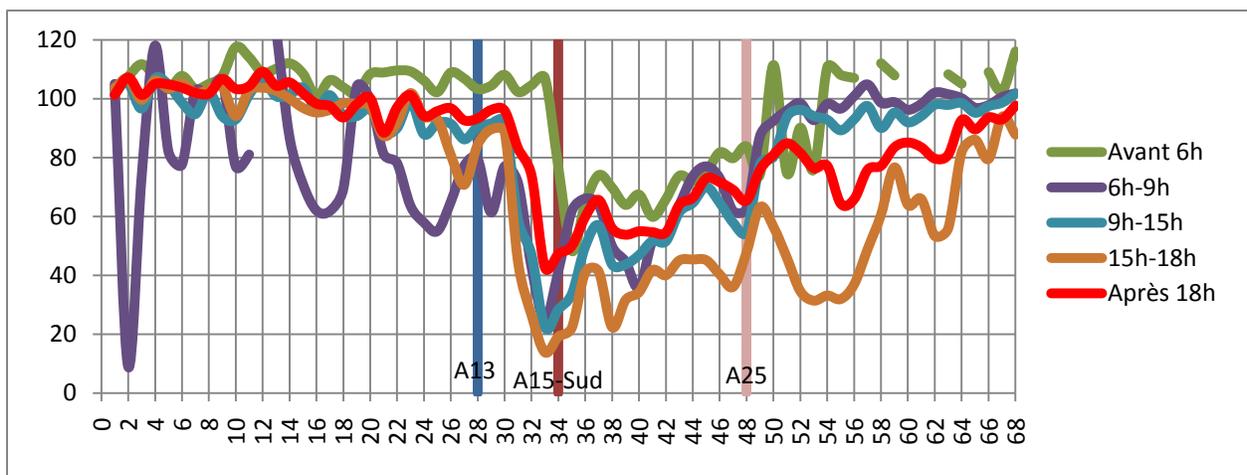


Figure 42. Variation de la vitesse moyenne (tous jours de semaine confondus, tous mois confondus) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée

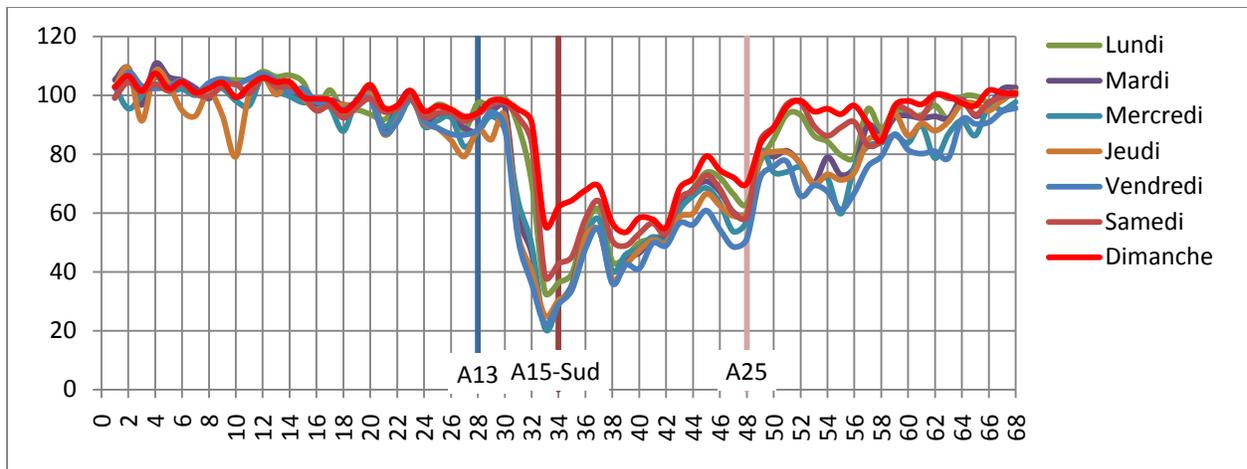


Figure 43. Variation de la vitesse moyenne (tous mois confondus) sur l'autoroute 40 (direction EST) lors des différents jours de la semaine.

Évidemment, avec la taille actuelle de l'échantillon, il n'est pas possible de faire toutes les analyses multivariées souhaitées; ce n'était pas non plus possible avec l'échantillon recueilli avec les véhicules flottants. En outre, comme la taille de l'échantillon pourrait assez rapidement être multipliée, ces limitations pourraient en partie être réduites.

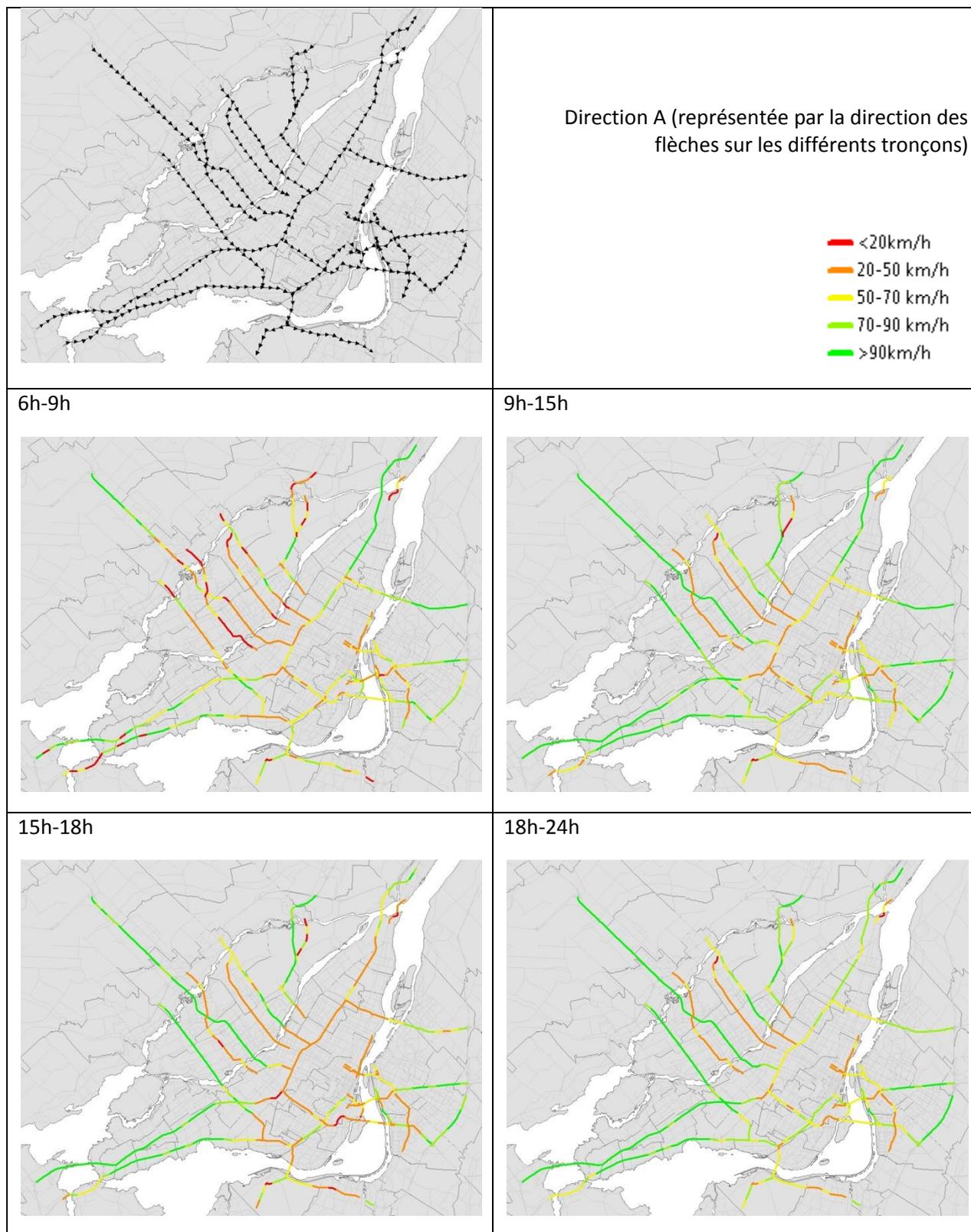


Figure 44. Variation de la vitesse moyenne (tous mois confondus, jours de la semaine excluant samedi et dimanche) sur les différents tronçons du réseau (sous-ensemble de directions, tel que représenté par les flèches), pour différentes périodes

### 3.2.3 Données de vitesses moyennes par segment RealTraffic

Les conditions de circulation sur l'A40 sont maintenant étudiées à l'aide des estimations fournies par RealTraffic. Seul un sous-ensemble des tronçons de l'autoroute peut-être étudié. La variabilité de la vitesse moyenne est d'abord examinée pour différentes périodes de la journée. Il faut rappeler que les statistiques s'appuient uniquement sur des données recueillies en août 2010 (voir Figure 45). On observe déjà que la variabilité générale de la vitesse moyenne estimée à partir des données de RealTraffic est nettement inférieure à celle issue des observations GPS. Entre les tronçons 28 et 49 de l'A-40 est, RealTraffic estime une vitesse moyenne par tronçon qui varie entre 45 et 83 km/h alors que les données GPS indiquent plutôt que la vitesse moyenne fluctue entre 14 et 108 km/h. La différence apparaît énorme. L'analyse par type de jour (Figure 46) confirme aussi la faible variabilité des mesures de vitesses estimées à l'aide de cet ensemble de données. Difficile d'expliquer ces résultats sans plus de détails sur les mécanismes d'extraction de données et d'estimation de vitesses par tronçon.



Figure 45. Variation de la vitesse moyenne (tous jours de semaine confondus, août 2010) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée

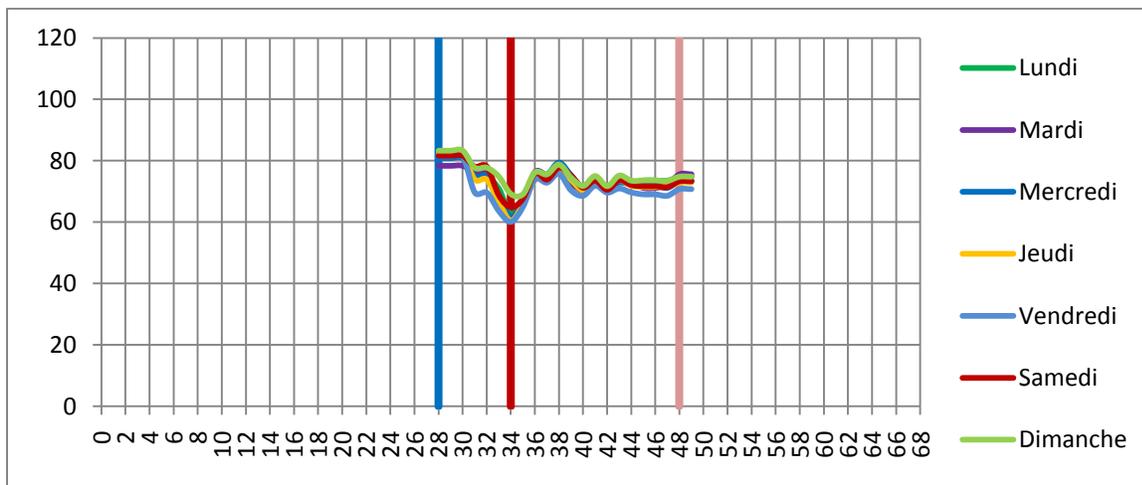
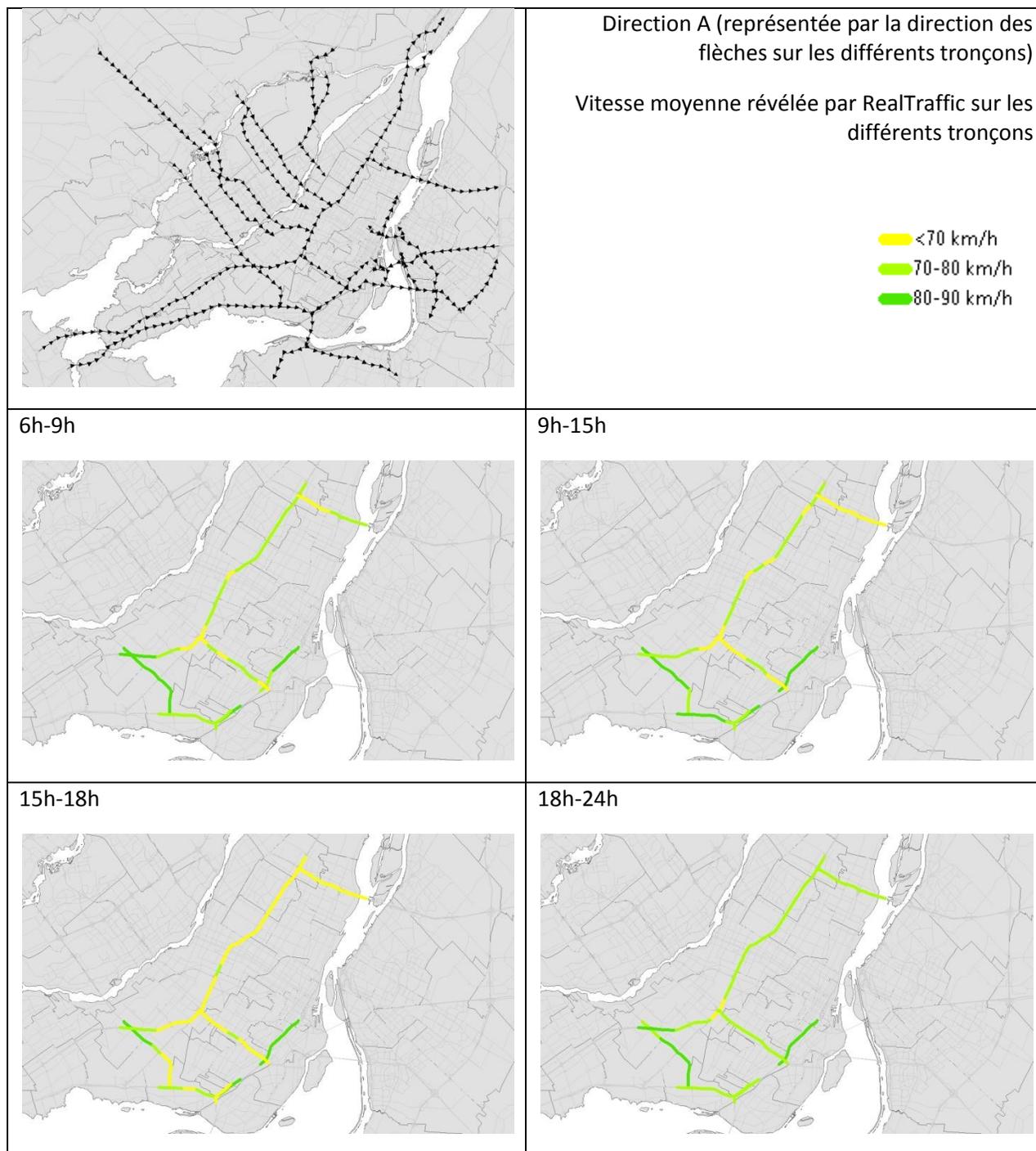


Figure 46. Variation de la vitesse moyenne (tous jours de semaine confondus, tous mois confondus) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée



**Figure 47. Variation de la vitesse moyenne (tous mois confondus, jours de la semaine excluant samedi et dimanche) sur les différents tronçons du réseau (sous-ensemble de directions, tel que représenté par les flèches), pour différentes périodes selon les données de RealTraffic**

### 3.2.4 Données de temps de parcours INRIX

Le même tronçon est étudié avec les données obtenues de la compagnie INRIX. A première vue, les estimations de vitesse obtenues de ces données sont beaucoup plus proches de ce qui est obtenu avec les données GPS que les valeurs précédentes obtenues de RealTraffic. En effet, la variabilité de la vitesse moyenne (entre tronçons) est plus élevée et semble mieux capter les différentes conditions de circulation typiquement vécues par les conducteurs.

- La Figure 48 présente la vitesse moyenne sur les tronçons 16 à 56 pour différentes périodes de la journée. Tel qu'attendu, les vitesses sont inférieures et plus variables pendant les périodes de pointe et fluctuent entre 20 et 105 km/h sur les tronçons étudiés.
- La Figure 49 présente la fluctuation de la vitesse moyenne par kilomètre, pour les différents jours de la semaine. Ces données confirment les vitesses typiquement supérieures pour les jours de fin de semaine ainsi qu'une similarité apparente des vitesses lors des différents jours de la semaine (lundi au vendredi).

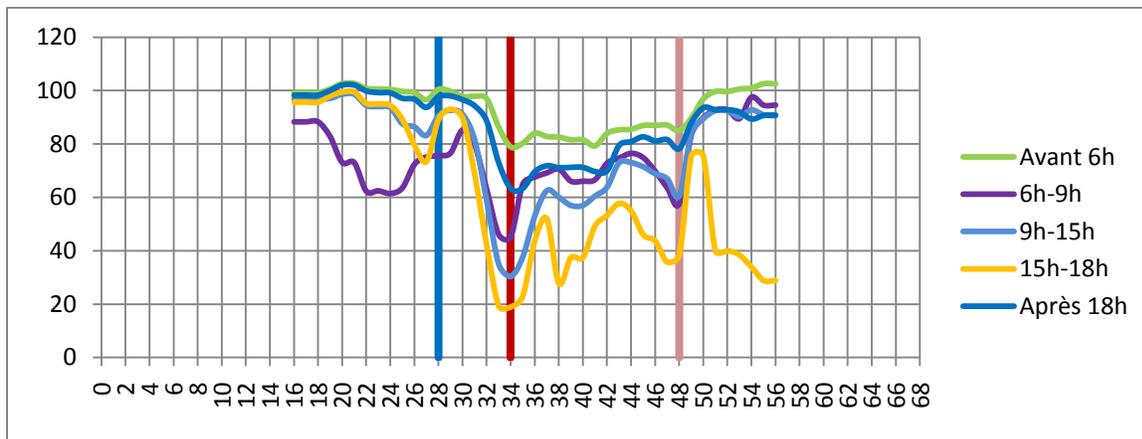


Figure 48. Variation de la vitesse moyenne (tous jours de semaine confondus, août 2010) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée

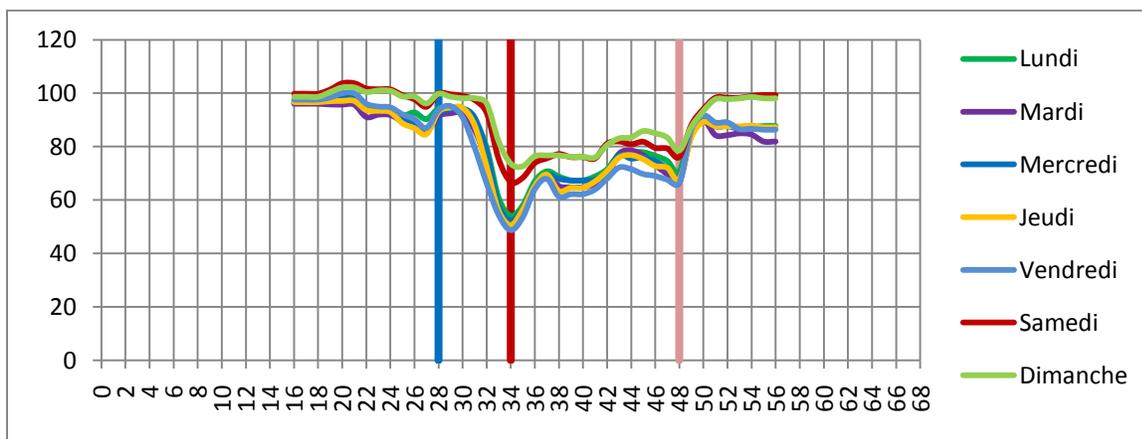
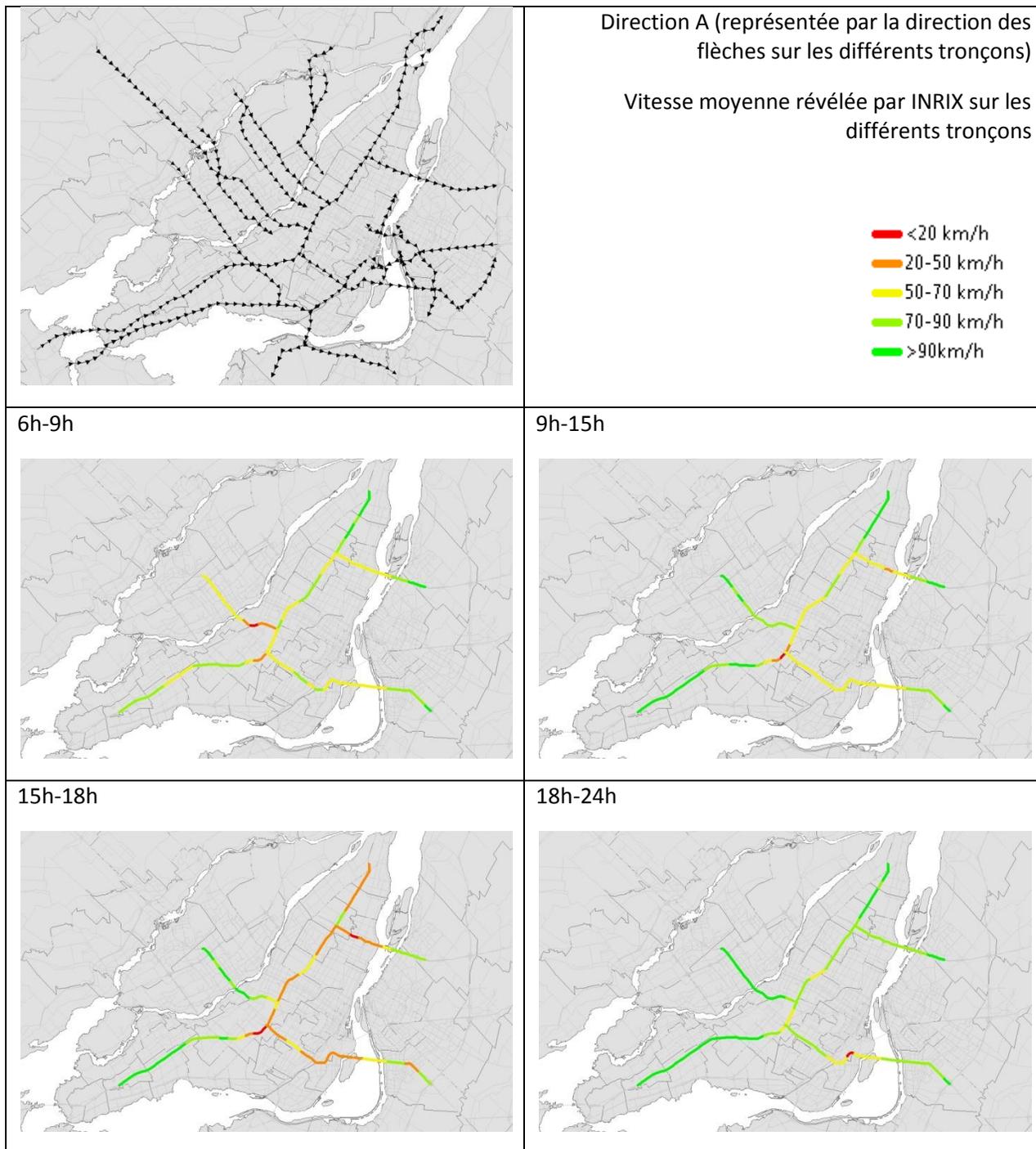


Figure 49. Variation de la vitesse moyenne (tous jours de semaine confondus, tous mois confondus) sur l'autoroute 40 (direction EST) lors des différentes périodes de la journée



### 3.3 Méthodologie d'analyse comparative

Pour fins d'évaluation des différentes sources de données, une analyse comparative est conduite aussi sur les segments communs de l'A40, direction est, pour le mois d'août 2010 et plus précisément sur le tronçon 34 en direction est. Différents indicateurs sont examinés.

#### 3.3.1 Définition des indicateurs

Différents indicateurs sont utilisés pour évaluer la capacité des données à notre disposition à révéler les conditions de circulation sur le réseau.

##### Indicateurs basés sur les vitesses

- Vitesse moyenne sur un tronçon :

$$V_{moyen} = \frac{\sum \text{Vitesse instantanées}}{\text{Nombre d'observations}}$$

- Temps moyen de parcours sur un tronçon :

$$T_{moyen} = \frac{\text{Distance}}{V_{Moyen}}$$

- Indice d'asymétrie de la distribution des vitesses :

$$I_{Asymétrie} = \frac{90^{ième} \text{ Percentile} - 50^{ième} \text{ Percentile}}{50^{ième} \text{ Percentile} - 10^{ième} \text{ Percentile}}$$

Cet indice définit la forme de la distribution de fréquence des temps de parcours. Une valeur de 1 correspond à une distribution normale.

##### Indicateurs basés sur le temps de parcours

- Indice de temps sécuritaire :

$$I_{Temps\ sécuritaire} = \frac{95^{ième} \text{ Percentile} - \text{Temps moyen}}{\text{Temps moyen}} \times 100\%$$

Cet indice correspond au pourcentage de temps supplémentaire nécessaire qui doit être prévu pour ne pas être en retard dans 95% des cas selon les conditions moyennes de circulation à ce moment (FHWA 2005).

- Indice de temps planifié :

$$I_{Temps\ planifié} = \frac{95^{ième} \text{ Percentile}}{\text{Temps en écoulement libre}} \times 100\%$$

Cet indice définit le temps supplémentaire nécessaire pour ne pas être en retard dans 95% des cas selon les conditions en écoulement libre de la circulation.

- Indice de médiocrité :

$$I_{Médiocrité} = \frac{\text{Temps moyen des 5\% les plus lents}}{\text{Temps moyen}} - 1$$

Cet indice mesure le degré de médiocrité du tronçon. C'est la proportion de temps de parcours supplémentaire moyen des 5 % des parcours les plus lents par rapport au temps de parcours moyen.

### 3.3.2 Comparaison des vitesses moyennes

L'objectif ici est de comparer les différents résultats fournis par les sources de données. Pour ce faire, l'ensemble des tronçons communs, entre les trois ensembles de données, est étudié afin de faire révéler leur variabilité.

La Figure 50 représente la distribution cumulée des tronçons en fonction de l'amplitude des vitesses observées sur chacun d'eux. L'amplitude est caractérisée par la différence entre la vitesse maximale et minimale de la journée sur 24 heures. Cet indicateur permet d'évaluer la variabilité journalière d'une source de données par rapport aux autres.

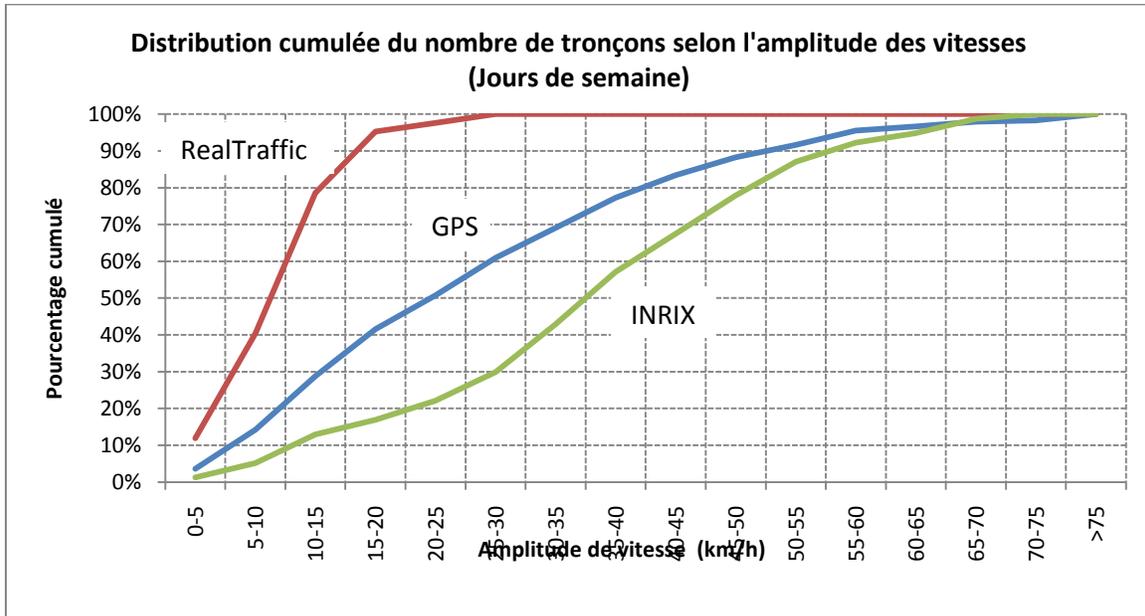


Figure 50. Distribution cumulée des tronçons selon l'amplitude des vitesses observées sur chacun d'eux pendant un jour de semaine

Les résultats nous révèlent qu'il y a peu d'amplitude sur les tronçons de RealTraffic. Il y a donc peu de variabilité de la vitesse à l'intérieur d'un tronçon. 95 % des tronçons ont une amplitude de moins de 20 km/h sur une journée. Les résultats fournis par les données GPS et par INRIX nous montrent une distribution beaucoup étalée. Ce qui signifie qu'il y a plus de variabilité au niveau de la vitesse moyenne, tel qu'on pourrait le supposer lors d'un jour typique sur le réseau routier.

La Figure 51 et la Figure 52 représentent la comparaison des vitesses observées entre les données GPS, INRIX et RealTraffic. En montrant la proportion de tronçons dont la différence de vitesse par groupe d'heures et par tronçon par rapport aux données GPS est inférieure à 5, 10, 15 et 20 km/h

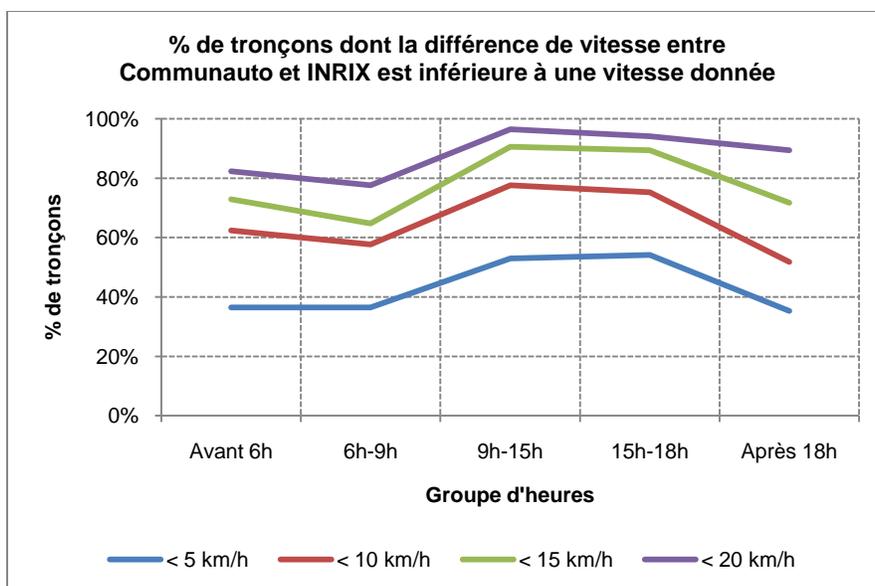


Figure 51. Comparaison des vitesses observées entre les GPS de Communauto et INRIX par tronçon et par groupe d'heures (Jour de semaines, août 2010)

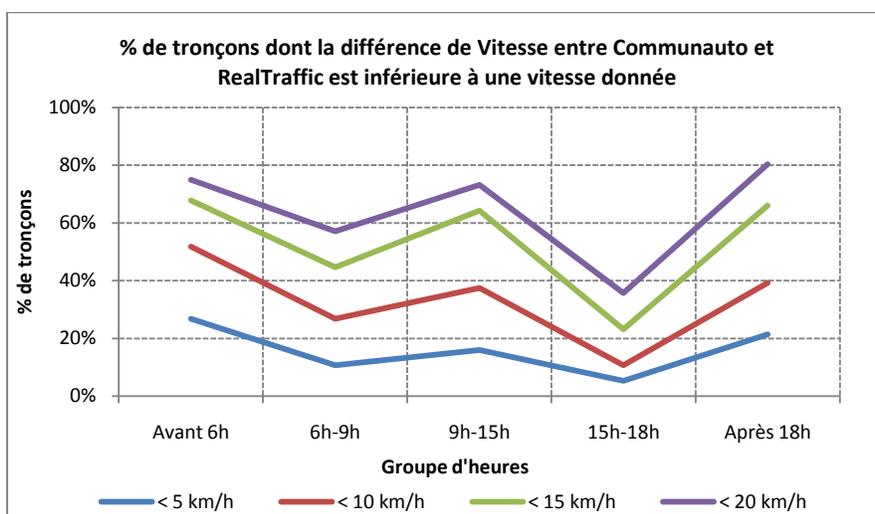


Figure 52. Comparaison des vitesses observées entre les GPS de Communauto et Real Traffic par tronçon et par groupe d'heures (Jour de semaines, août 2010)

Premièrement, il y a une grande différence entre les différentes sources de données. En général, les données GPS semblent mieux correspondre avec celles d'INRIX. Environ la moitié (50%) des tronçons ont une vitesse moyenne par tronçon inférieure à 5 km/h et environ 90% des tronçons ont une différence inférieure à 30 km/h. Ceci est aussi observé en période de pointe de l'après-midi (15h00 à 18h00) où les vitesses GPS et INRIX sont plus similaires que celles estimées via RealTraffic. En effet, pour RealTraffic, seulement 10% des tronçons ont une différence de vitesse inférieure à 5 km/h et 40% pour une différence de 20 km/h tandis que pour INRIX, près de 95% des tronçons ont une différence observée inférieure à 20 km/h.

### 3.3.3 Comparaison des distributions

La Figure 53, la Figure 54 et la Figure 55 présentent les distributions de fréquence des vitesses observées pour l'ensemble des tronçons étudiés. Il est à noter que les données fournies par INRIX et RealTraffic sont des ensembles de données continus dans le temps tandis que les données GPS sont instantanées. Il y a donc beaucoup moins d'observations pour les données GPS que pour les autres sources de données.

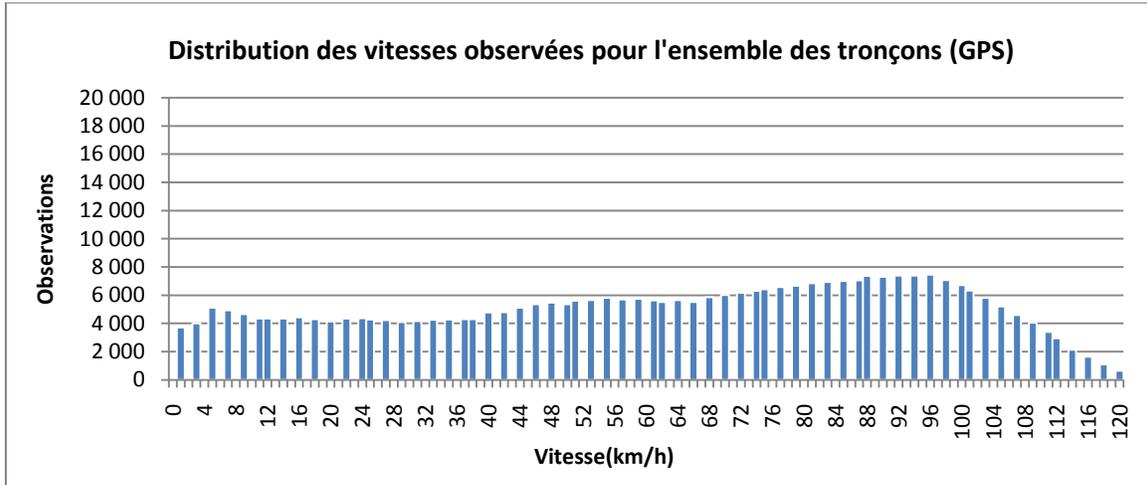


Figure 53. Distribution des observations de vitesses GPS de Communauto (Tous les tronçons échantillonnés, jours de semaine seulement)

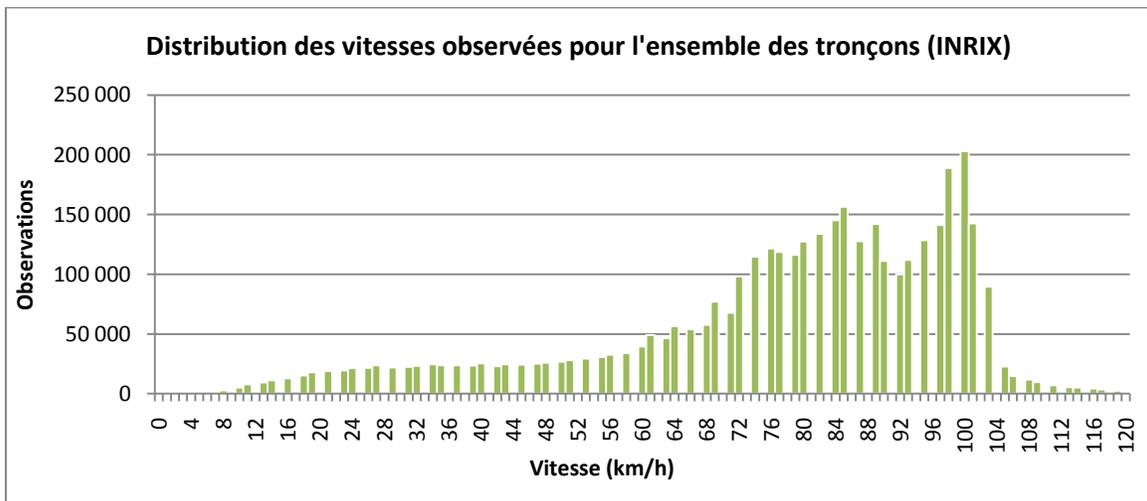


Figure 54. Distribution des observations de vitesses INRIX (Tous les tronçons échantillonnés, jours de semaine seulement)

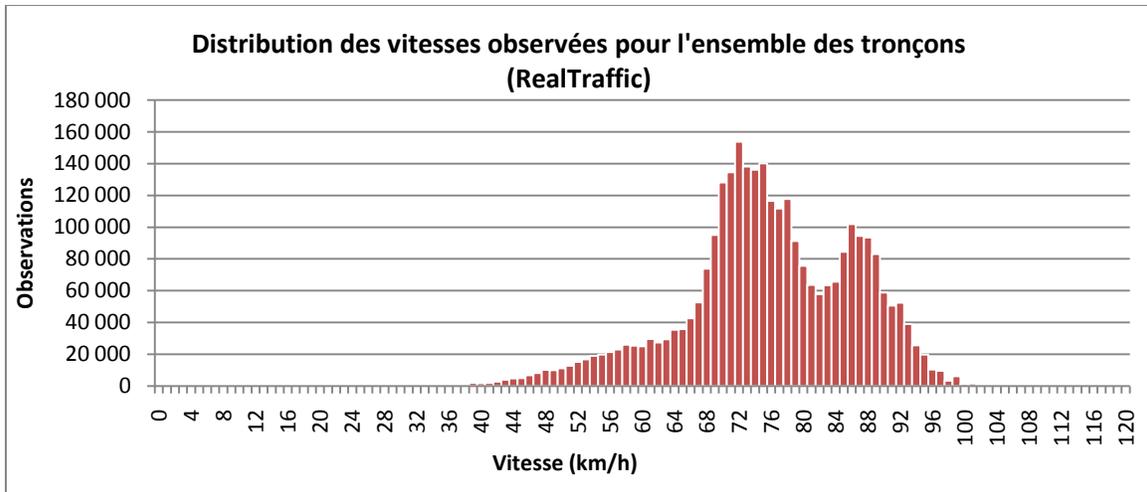


Figure 55. Distribution des observations de vitesses Real Trafic (Tous les tronçons échantillonnés, jours de semaine seulement)

On observe une distribution très étalée pour les données GPS. Ceci est dû, entre autres, à l'échantillonnage qui est non-uniforme pendant la journée. Une autre particularité à noter est l'absence de données pour certaines vitesses particulières dans les résultats d'INRIX et GPS. Sans trop en connaître la source, un peu supposer qu'il s'agit d'un « biais » dû aux technologies utilisées, biais compréhensible dans le cas des traces GPS (une seule technologie) mais plus étonnante pour INRIX étant donnée la multitude de sources potentielles utilisées pour la production des données. La distribution des vitesses pour RealTraffic est beaucoup moins étalée. La vitesse maximale observée est de 100 km/h et la vitesse minimale est de 40 km/h.

### 3.3.4 Étude d'un tronçon particulier

Les prochaines figures se basent sur l'étude d'un tronçon unique de 1 km, se situant sur l'autoroute 40 en direction Est à la jonction de l'Autoroute 15 sud. La position géographique de ce tronçon est montrée à la Figure 56. Ce lien routier est caractérisé par une grande variabilité de la vitesse moyenne pendant la journée. L'objectif ici est de présenter une méthodologie permettant de comparer les distributions de différentes sources de données. Cette méthode pourrait être appliquée à l'ensemble des tronçons.

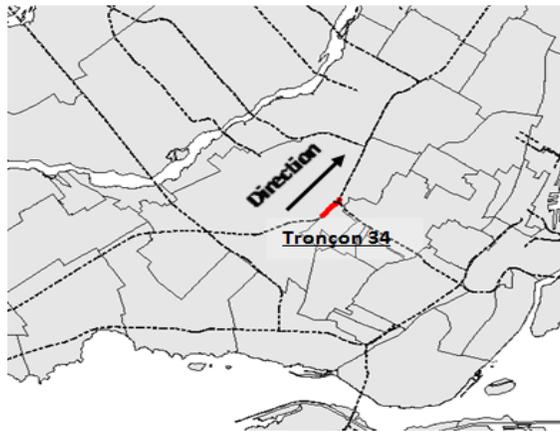


Figure 56. Position géographique du tronçon 34 de l'autoroute 40 en direction est.

La Figure 57 et la Figure 58 présentent la distribution horaire des vitesses moyennes observées pour le mois d'août 2010.

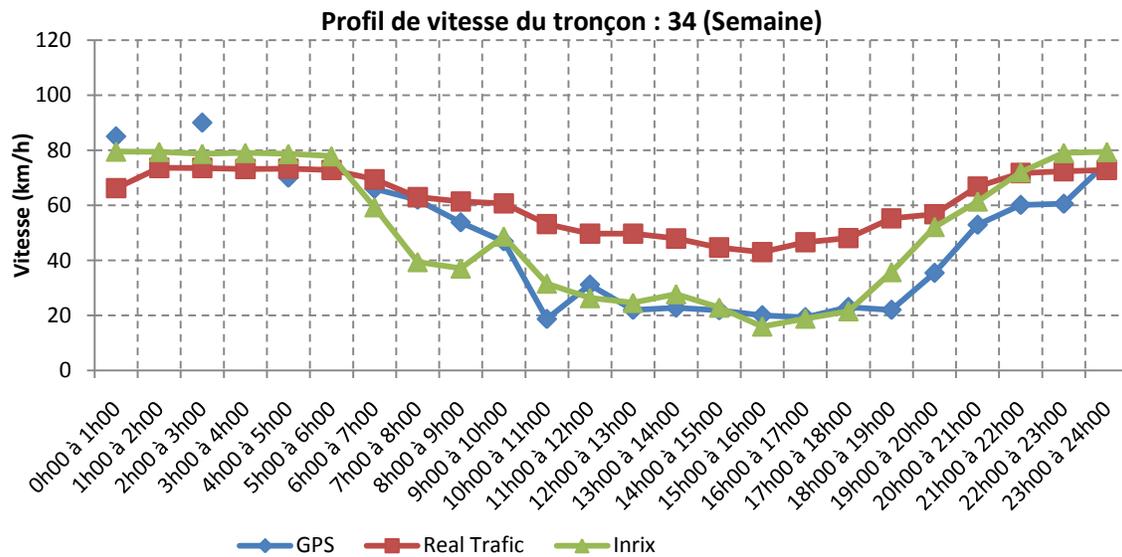


Figure 57. Distribution horaire des vitesses moyennes par heure sur le tronçon 34 en direction est (Jours de semaine, août 2010)

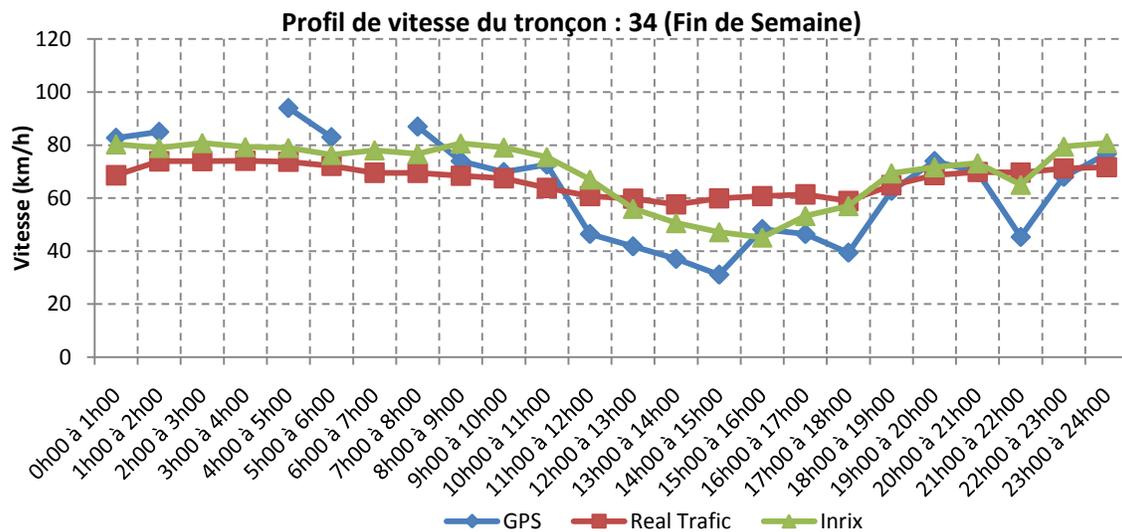


Figure 58. Distribution horaire des vitesses moyennes par heure sur le tronçon 34 en direction est (Jours de fin de semaine, août 2010)

La distribution horaire des données GPS et d'INRIX est semblable. Cependant, on remarque que celle de RealTraffic est très différente des autres. Elle est très peu variable pendant la journée, ce qui rejoint les observations faites dans les sections précédentes. On remarque aussi une plus grande variabilité pendant la nuit et le début de la matinée pour les données GPS. Ceci est dû à la taille de l'échantillon qui est beaucoup plus petite pendant ces périodes.

La Figure 59, la Figure 60 et la Figure 61 présentent les distributions de vitesses en période de pointe de l'après-midi pour les trois ensembles de données.

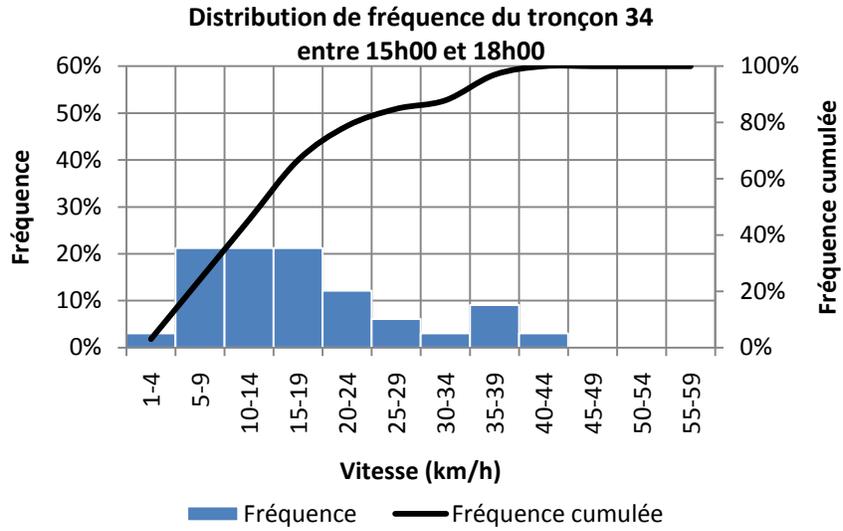


Figure 59. Distribution de fréquence des vitesses observées par les GPS de Communauto du tronçon 34 en direction Est en pointe de l'après-midi (tous jours de semaine, août 2010)

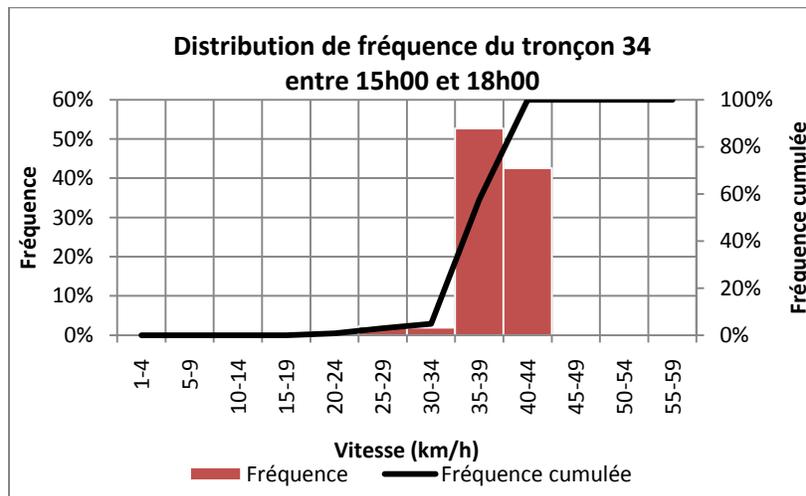


Figure 60. Distribution de fréquence des vitesses observées par RealTraffic du tronçon 34 en direction Est en pointe de l'après-midi (tous jours de semaine, août 2010)

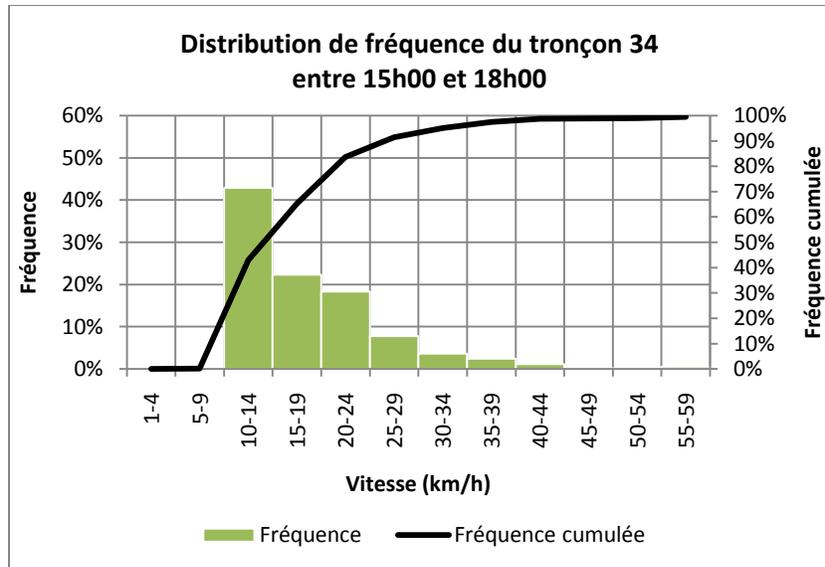


Figure 61. Distribution de fréquence des vitesses observées par INRIX du tronçon 34 en direction Est en pointe de l'après-midi (tous jours de semaine, août 2010)

On remarque que la distribution est très étalée pour les données GPS. Pour les résultats de RealTraffic, la distribution est très peu étalée. Il y a donc très peu de variabilité dans les mesures, constat qui rejoint les observations faites précédemment. La distribution de vitesse d'INRIX est assez étalée et on remarque une asymétrie vers la droite.

Le Tableau 30 présente les indicateurs décrits précédemment est estimés pour le tronçon 34 à l'aide des trois ensembles de données.

Tableau 30. Caractérisation des distributions pour le tronçon 34 en pointe de l'après-midi

Indice	Résultats		
	GPS	Real Trafic	INRIX
Asymétrie	2.2	1.0	1.7
Sécuritaire	182 %	9.9 %	71.8 %
Planifié	1130 %	200 %	335 %
Médiocrité	3.6	0.21	0.75

Ces différents indices présentent différents constats. Premièrement, les résultats nous montrent que les données GPS ont une plus grande variabilité que les autres sources de données. L'indice de temps sécuritaire signifie qu'il faut prévoir 1.8, 0.1 et 0.7 fois, respectivement pour les données GPS, RealTraffic et INRIX, le temps moyen de parcours pour s'assurer d'arriver à temps dans 95% des trajets. L'indice de temps planifié nous donne une indication semblable mais, en fonction du temps de parcours en écoulement libre. L'indice de médiocrité nous donne une idée sur l'étendue de la distribution. Plus cet indice est élevé, plus la différence est importante entre les véhicules les plus lents et la moyenne et donc, plus la distribution est étalée.

### 3.4 Discussion

Cette section a d'abord proposé une synthèse des différentes sources de données disponibles à savoir :

- Un ensemble de points représentant des vitesses instantanées recueillies à l'aide de GPS localisés dans quelque 400 véhicules de la compagnie Communauto;
- Un ensemble de données provenant de la compagnie RealTraffic et reflétant les vitesses moyennes observées par tronçon telles qu'estimées par une extraction d'information depuis des bandes vidéo;
- Un ensemble de données de la compagnie INRIX qui fournit des estimés de vitesses par tronçon, avec différents niveaux de qualité.

### 3.4.1 Limitations

---

A priori, les trois sources de données étudiées ne sont pas compatibles au niveau spatial. En effet, les données GPS ne sont couplées à aucune donnée de réseau, les données de RealTraffic sont disponibles pour des tronçons routiers de type schématique et les données INRIX disponibles aussi pour des tronçons routiers de taille variable. La procédure de couplage des données utilisée pour mettre les trois sources de données sur une définition uniforme de réseau basé sur des tronçons à un kilomètre implique une perte d'information et des imprécisions au niveau de la localisation spatiale de la mesure (les tronçons des différentes sources ne sont évidemment pas parfaitement compatibles).

### 3.4.2 Synthèse

---

Les analyses ont permis d'observer les différences au niveau de l'échantillonnage (par rapport à l'échantillon de données typiquement obtenu de relevés avec véhicules flottants) ainsi que des distributions de vitesse globalement, pour une route ainsi que pour un tronçon particulier. À la lumière de ces analyses, les vitesses produites par RealTraffic ne sont pas représentatives des comportements typiques. Il semble que le modèle ne soit pas correctement calibré ce qui expliquerait ces résultats. En contrepartie, les estimations de vitesses obtenues des GPS ainsi que d'INRIX sont proches et révèlent correctement la variabilité attendue des conditions de circulation. Les données d'INRIX qui nous ont été fournies semblent donc fournir des informations correctes sans toutefois qu'il soit possible d'en évaluer la provenance et la qualité de l'échantillonnage. Ces informations ont toutefois un prix élevé. Les données GPS méritent d'être examinées et valorisées de façon plus approfondies, tel qu'il sera exposé dans les paragraphes qui suivent.

### 3.4.3 Recommandations

---

À la lumière des différents résultats et de l'analyse comparative des échantillons disponibles, il appert que les points GPS extraits des véhicules de Communauto offrent un potentiel très intéressant pour le MTQ en termes de monitoring des conditions de circulation sur son réseau autoroutier. Bien que d'autres outils soient aussi intéressants et fournissent une information pertinente, il semble qu'ils soient moins accessibles et moins souples. Dans ce contexte, nous recommandons au MTQ d'amorcer des discussions avec les représentants de Communauto afin de discuter des possibilités d'accès en continu aux données GPS issues des véhicules partagés. Les éléments suivants devraient être discutés :

- Rythme possible de production des données (délais requis pour accéder aux ensembles de données, possibilités d'échange direct avec le fournisseur des équipements, etc.);
- Possibilité d'augmenter le niveau de résolution temporel des données aux minutes ou 30 secondes afin d'améliorer l'échantillonnage tout en assurant une forme d'indépendance des observations utilisées pour caractériser les tronçons autoroutiers;

- Possibilité d'augmenter le nombre de véhicules équipés de capteurs GPS afin d'améliorer l'échantillonnage, tant au niveau spatial que temporel.

Dans la mesure où un partenariat est fait avec la compagnie Communauto, nous pourrions envisager le développement d'un outil de monitoring stratégique de l'évolution de la circulation sur les principales autoroutes de la région basé sur les principes suivants :

- accessible sur plate-forme web;
- intégrant des fonctionnalités de traitement systématique de lots de données (lots mensuels par exemple), fonctionnalités basées sur la méthodologie de couplage déjà développée, qui pourra être raffinée pour profiter pleinement de la richesse des bases de données, surtout sur leur niveau de résolution temporelle augmente;
- permettant de visualiser l'état de la circulation (vitesses moyennes et coefficient de variation) sur des tronçons autoroutiers selon différentes conditions (neige, pluie) ou circonstances (jour de la semaine, période de la journée, mois). Un découpage plus raffiné du réseau autoroutier serait à envisager (500m ou 250m avec délimitations respectant davantage les attributs physiques);
- fournissant des statistiques évolutives des conditions de circulation avec le traitement de nouveaux lots.

Ce type d'outil pourrait d'abord être utilisé pour fins de planification stratégique mais pourrait aussi être enrichi notamment d'un outil d'information aux usagers qui permettrait aux voyageurs d'obtenir un temps plausible de déplacement sur différents trajets (agrégation de tronçons autoroutiers) selon différentes conditions et circonstances.

## 4. Analyse des segments routiers en fonction de leurs caractéristiques physiques

Cette section décrit une expérimentation effectuée sur la caractérisation physique des segments du réseau routier supérieur de la grande région de Montréal. Suite aux caractérisations et groupements effectués à l'aide des données de véhicules flottants, l'idée est ici de déterminer s'il est possible de prédire l'appartenance à l'un de ces groupes (que nous appellerons groupes « temps de parcours », ou TP) d'un segment n'ayant pas été échantillonné par des véhicules flottants. Le cas échéant, il serait envisageable, à partir des caractéristiques physiques d'un segment routier (nombre de voies, présence d'accotement, d'intersections, etc.), de déterminer une distribution fréquentielle plausible de temps de parcours.

La méthodologie utilisée compte 6 étapes qui seront détaillées dans les sections suivantes (Figure 62). Les premières étapes touchent le montage de la base de données (1) et l'analyse du réseau routier sur la base des facteurs physiques (2). Après avoir effectué une analyse factorielle des correspondances multiples (3), le groupement des segments selon leurs facteurs physiques est présenté (4), suivi d'une confrontation de ces groupements dits « FP » aux groupements « TP » (5). Enfin, une méthode de prédiction d'appartenance des segments à des groupes typiques, pouvant être généralisée, est discutée (6). Les travaux présentés ici sont plus détaillés dans Kalboussi (2011).

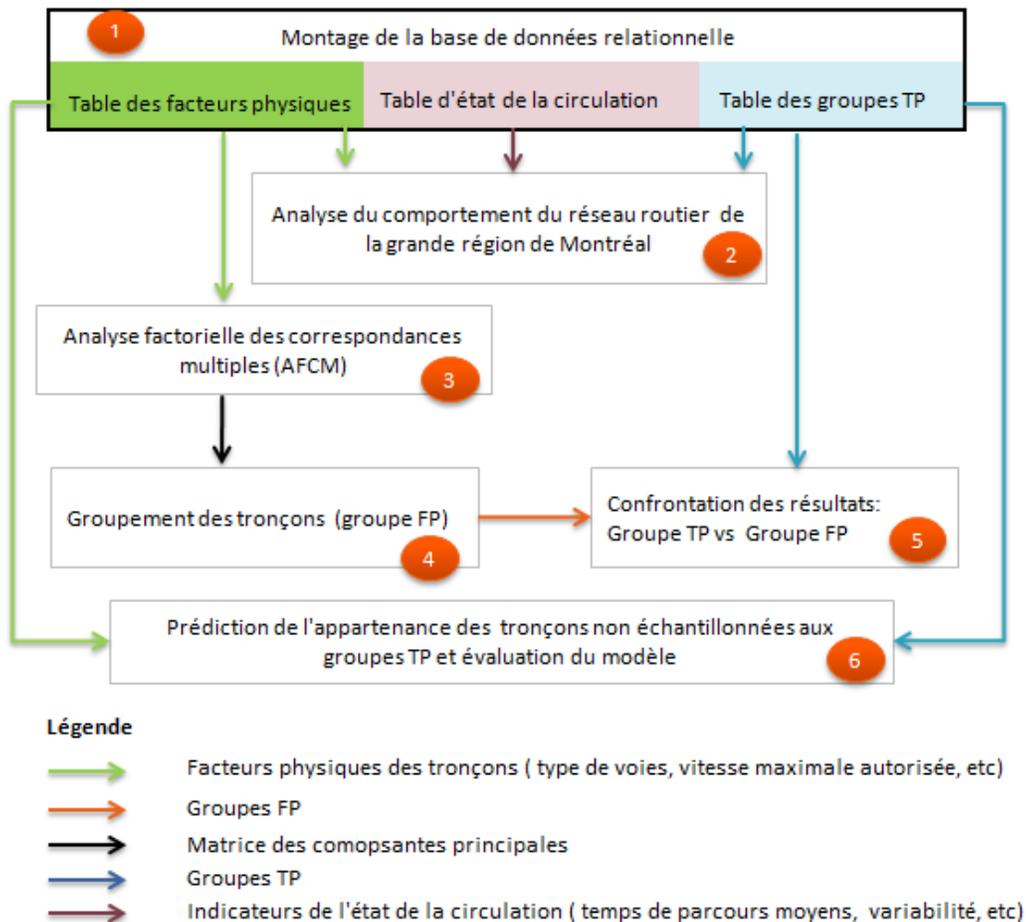


Figure 62. Méthodologie de caractérisation des segments routiers selon leurs caractéristiques physiques

### 4.1.1 Montage de la base de données

La base de données constituée pour l'analyse compte deux tables majeures. La première table contient de l'information sur les groupements effectués lors de l'analyse des temps de parcours (groupes TP). Nous allons présenter ici le contenu de la deuxième table, destinée à l'analyse des facteurs physiques (Tableau 31). Ces informations ont été colligées principalement par l'examen des segments, un à un, via le site web de Google StreetView.

Tableau 31. Contenu de la table de données sur les segments

Nom du Champ	Exemple	Description
TRUNI	21	Identifiant du tronçon unique
ID_FP	21	Identifiant du tronçon caractérisé
Nbr_sorties_sc	1	Nombre de sorties (sans contrôle)
Type_sorties_sc	TS=AU1_BV/AV0_CH/MO0	Type de sortie (sans contrôle)
Nbr_entrées_sc	0	Nombre d'entrées (sans contrôle)
Type_entrées_sc	TE=AU0_BV/AV0_CH/MO0	Type d'entrée (sans contrôle)
Nbr_intersections_av	0	Nombre d'intersections (avec contrôle)
Type_intersection_av	TI=AU0_BV/AV0_CH/MO0	Type d'intersection (avec contrôle)
sens_HC/VC	VM	Sens de la circulation
Type_voie	AU	Type de voie
Acc_D	Oui	Présence d'accotement à droite
Type_Barr_D	0	Code de type de barrière à droite
Acc_G	Non	Présence d'accotement à gauche
Type_Barr_G	3	Code de type de barrière à gauche
sur_pont	0	Nombre de ponts
Tunnel	0	Nombre de tunnels
Nbr_voies	2	Nombre de voies
Vitesse_aut	100	Vitesse autorisée en km/h

Les sorties et les entrées font référence aux bretelles d'autoroutes et aux croisements à niveau. Les intersections constituent la somme des entrées et sorties. Les intersections avec contrôle sont celles où il y a des feux de circulation. Pour chaque segment, un sens « vers » ou « depuis » le centre-ville a été attribué (les segments étant unidirectionnels). La présence d'accotement à droite ou à gauche (largeur suffisante pour un véhicule) est notée. Les barrières sont également indiquées. La table recense également la présence de ponts (viaducs) et de tunnels. Le dernier champ contient la vitesse autorisée sur le segment (dans certains cas, la vitesse recommandée, comme par exemple pour les courbes, a été substituée à la limite de vitesse affichée). À noter que la table et l'analyse associée font abstraction des volumes enregistrés sur les segments.

### 4.1.2 Analyse descriptive du réseau routier

L'analyse du réseau routier est faite ici en fonction de leur groupement TP en période AM. Nous rappelons dans le Tableau 32 les caractéristiques de ces groupes.

Tableau 32. Valeurs caractéristiques des groupes AM

AM	Temps de parcours moyen (s)	Coefficient de Variation	Vitesse (km/h)	Pourcentage de tronçons	% observations > 360 sec
1	58.6	58.2%	61.4	10.9%	0.2%
2	39.2	63.8%	91.9	32.6%	0.2%
3	48.2	46.2%	74.7	21.3%	0.4%
4	156.2	114.8%	23.0	3.8%	11.0%
5	109.1	74.2%	33.0	6.5%	3.0%
6	79.8	44.2%	45.1	15.8%	0.3%
7	119.4	46.7%	30.1	6.9%	0.9%
8	189.7	45.7%	19.0	2.2%	9.0%

La Figure 63 présente la distribution des segments selon leur vitesse autorisée, par groupe TP. On y remarque que pour le groupe TP2 (plus bas temps de parcours moyen), il y a une grande proportion de segments ayant une limite de 100 km/h. Ce résultat intuitif est contrecarré dans le cas du TP4, ayant également beaucoup de segments à 100 km/h, mais où le temps de parcours est beaucoup plus élevé.

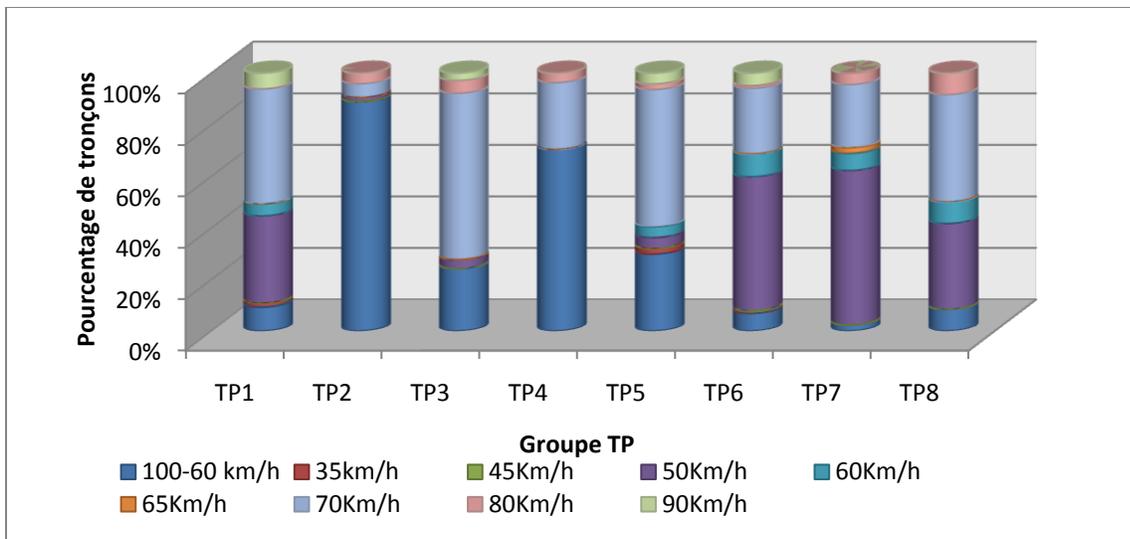


Figure 63. Vitesses autorisées par segment, selon leur groupe TP

La Figure 65 présente le nombre moyen d'intersections par groupe TP. Le groupe TP4, assez fortement congestionné, possède peu d'intersections (réseau autoroutier), tandis que les groupes 6, 7 et 8 comptent plus d'intersections, et se trouvent possiblement en milieu artériel urbain.

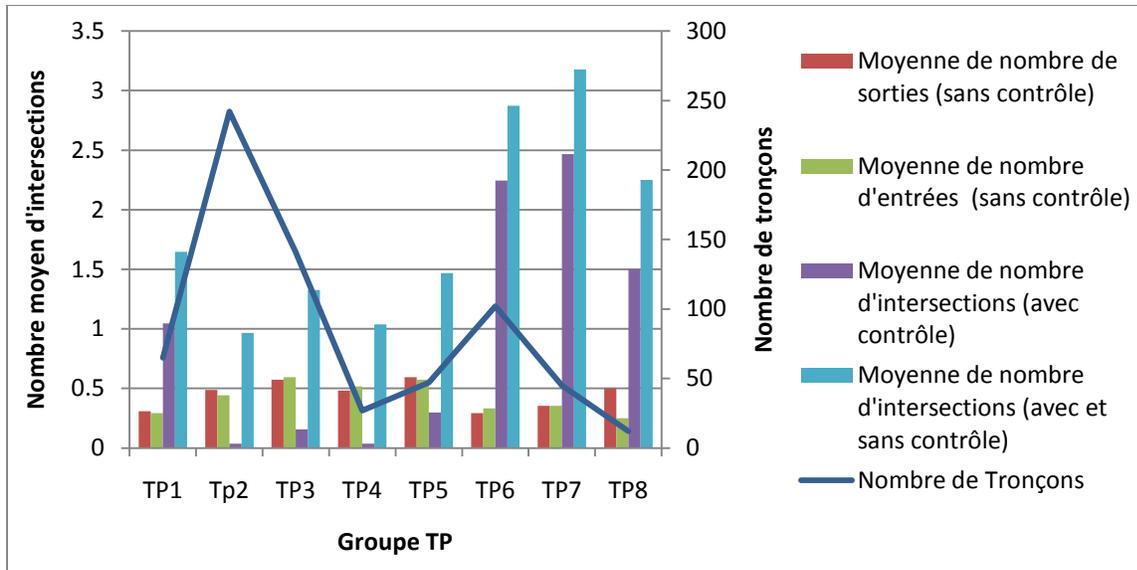


Figure 64. Nombre d'intersections par segment, selon leur groupe TP

La Figure 65 présente la répartition des segments en fonction du nombre de voies par direction, et selon le groupe TP. On y voit également l'appartenance autoroutière du TP4 (aucun segment à 1 voie), tandis que les TP2 et TP3 comptent plus de segments à 4 voies.

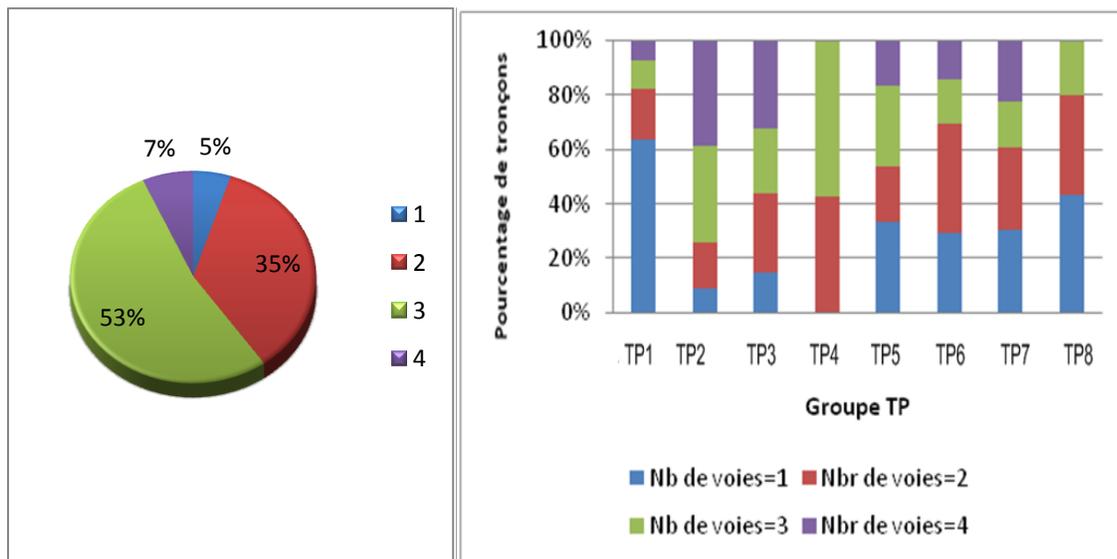


Figure 65. Nombre de voies par segment, selon leur groupe TP

#### 4.1.3 Analyse factorielle de correspondances multiples

La méthode de l'analyse factorielle des correspondances multiples (AFCM) a été utilisée dans cette étude parce que le nombre de modalités des facteurs physiques des tronçons est élevé. Cette méthode vise à étudier simultanément les relations entre les différentes modalités et permet d'agréger les données et de les représenter dans un plan multidimensionnel.

L'analyse des correspondances multiples a permis d'identifier 28 axes résumant les informations fournies par la base de données des facteurs physiques. La figure 5.1 illustre la répartition de la valeur propre totale sur les axes factoriels et présente également le pourcentage d'inertie cumulé. La question qui se pose est la détermination du nombre d'axes à retenir pour procéder au groupement des tronçons. Afin que le plan factoriel soit plus représentatif de la table de données des facteurs physiques et la répartition des tronçons soit par la suite plus significative, on a fixé un seuil d'information expliquée (pourcentage d'inertie cumulé), soit de 70 %. Ainsi, les treize premiers axes ont été retenus, avec 71% de l'inertie totale.

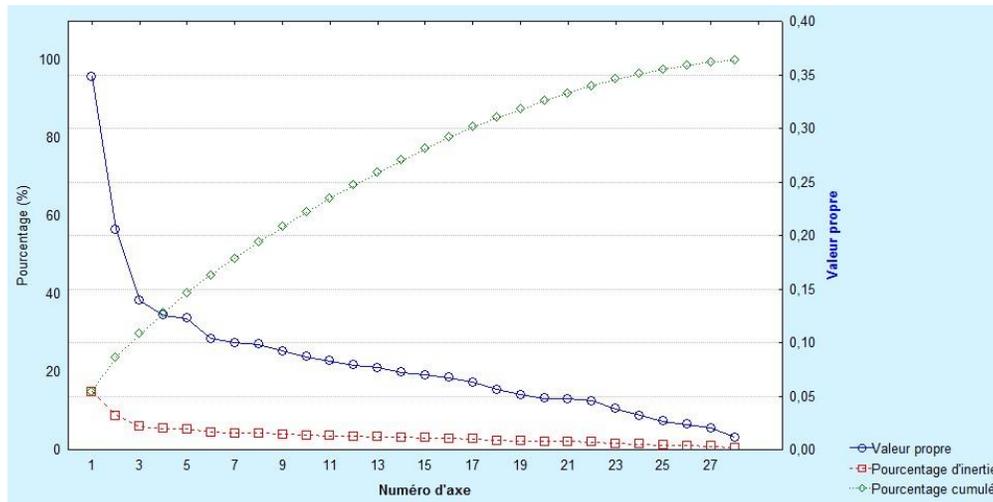


Figure 66. Valeur propre, pourcentage d'inertie, pourcentage cumulé expliquée par les axes factoriels

Dans un but exploratoire, on a projeté les modalités sur le plan factoriel qui est formé par les deux premiers axes factoriels (Figure 67). De cette figure, on tire certaines observations :

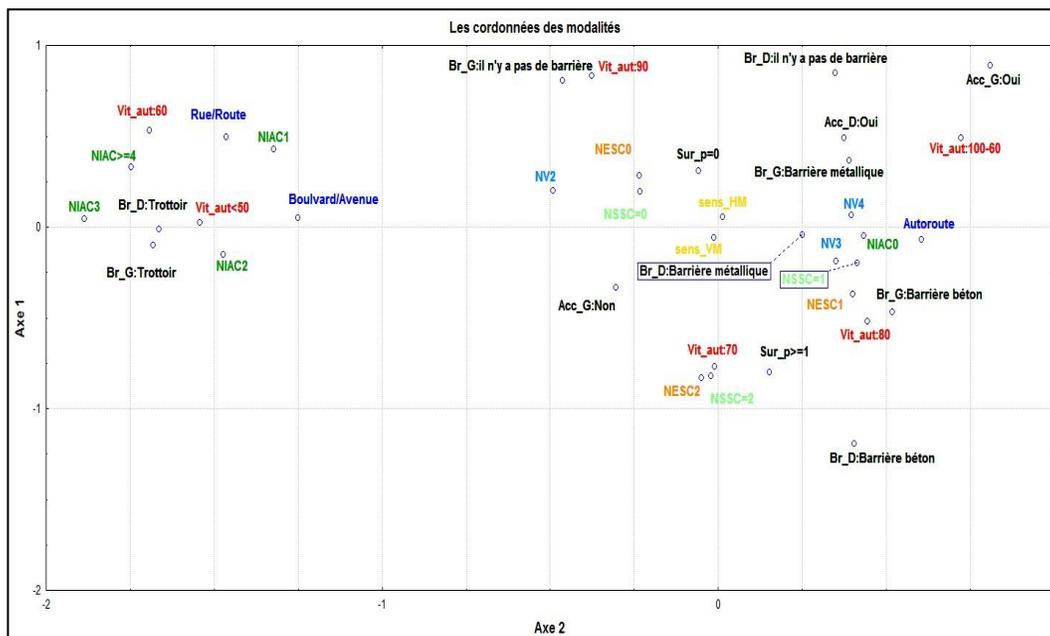


Figure 67. Coordonnées des modalités des facteurs physiques

- La plupart des modalités sont proches du centre du repère (axes 1, axes2), notamment les deux modalités du sens de circulation ayant les codes « VM » (Vers Montréal) et « HM » (Hors Montréal), hormis la modalité « vitesse autorisée égale 60 km/h » est la plus éloignée. Ceci se justifie par son poids faible dans la base de données.
- L'écart entre les deux modalités, la vitesse autorisée « 70 km/h » et le passage sur un pont, est faible. Ceci s'explique simplement par le fait que la vitesse maximale autorisée sur un pont de type « autoroute » diminue de 100 à 70 km/h.
- Il y a une certaine analogie entre le nombre et le type de voies, et la vitesse autorisée. En effet, si le nombre de voies d'un tronçon de type « autoroute » varie entre 3 et 4, alors la vitesse maximale autorisée varie entre 100 et 80 km/h. Toutefois, s'il n'y a qu'une seule voie, alors le tronçon est de type « rue ou route ».

#### 4.1.4 Groupements FP

Après la simplification des données par la méthode d'analyse factorielle des correspondances multiples (AFCM), un groupement des segments a été effectué à l'aide de la méthode des k-moyennes. Cet algorithme permet de grouper les observations selon les composantes principales issues de la méthode AFCM, soit les coordonnées des tronçons sur les axes du plan factoriel. Un nombre de 8 groupes a été constitué. Ce nombre a été choisi afin de minimiser les dispersions intragroupe et intergroupe, tout en respectant un nombre minimal de 25 segments par groupe (Figure 68).

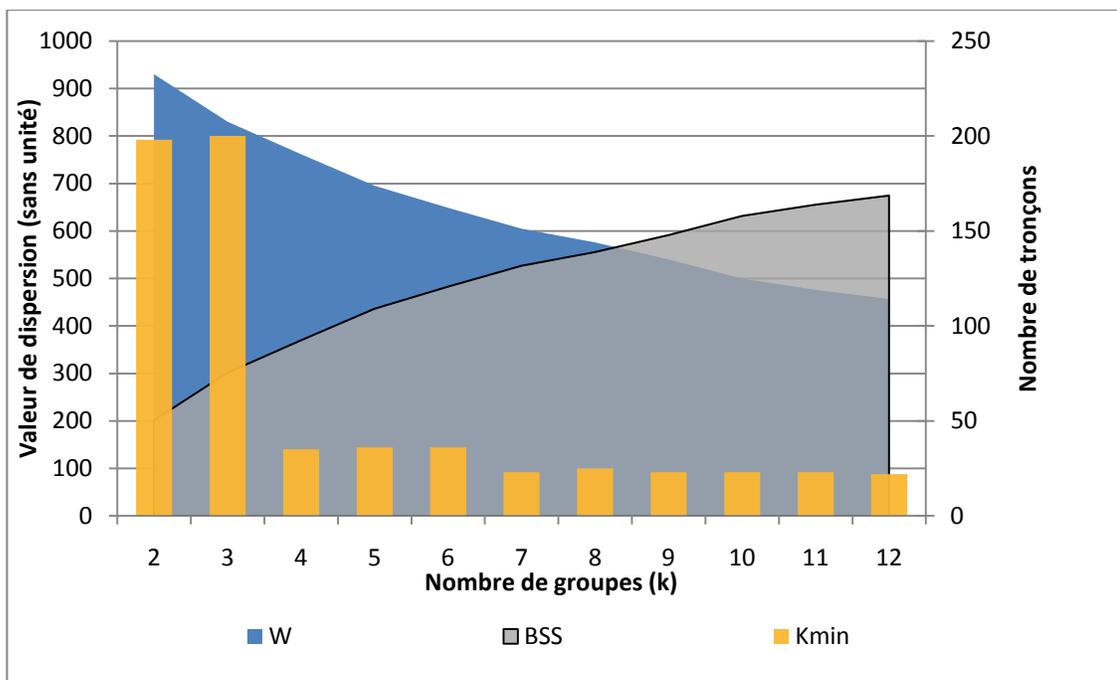


Figure 68. Dispersions intragroupe (w) et intergroupe (BSS)

#### 4.1.5 Confrontation des résultats (FP vs TP)

Le Tableau 33 présente la répartition de chaque groupe TP en fonction de leur appartenance au groupe FP. On y remarque en premier lieu qu'il n'y a pas de correspondance évidente entre les deux groupements.

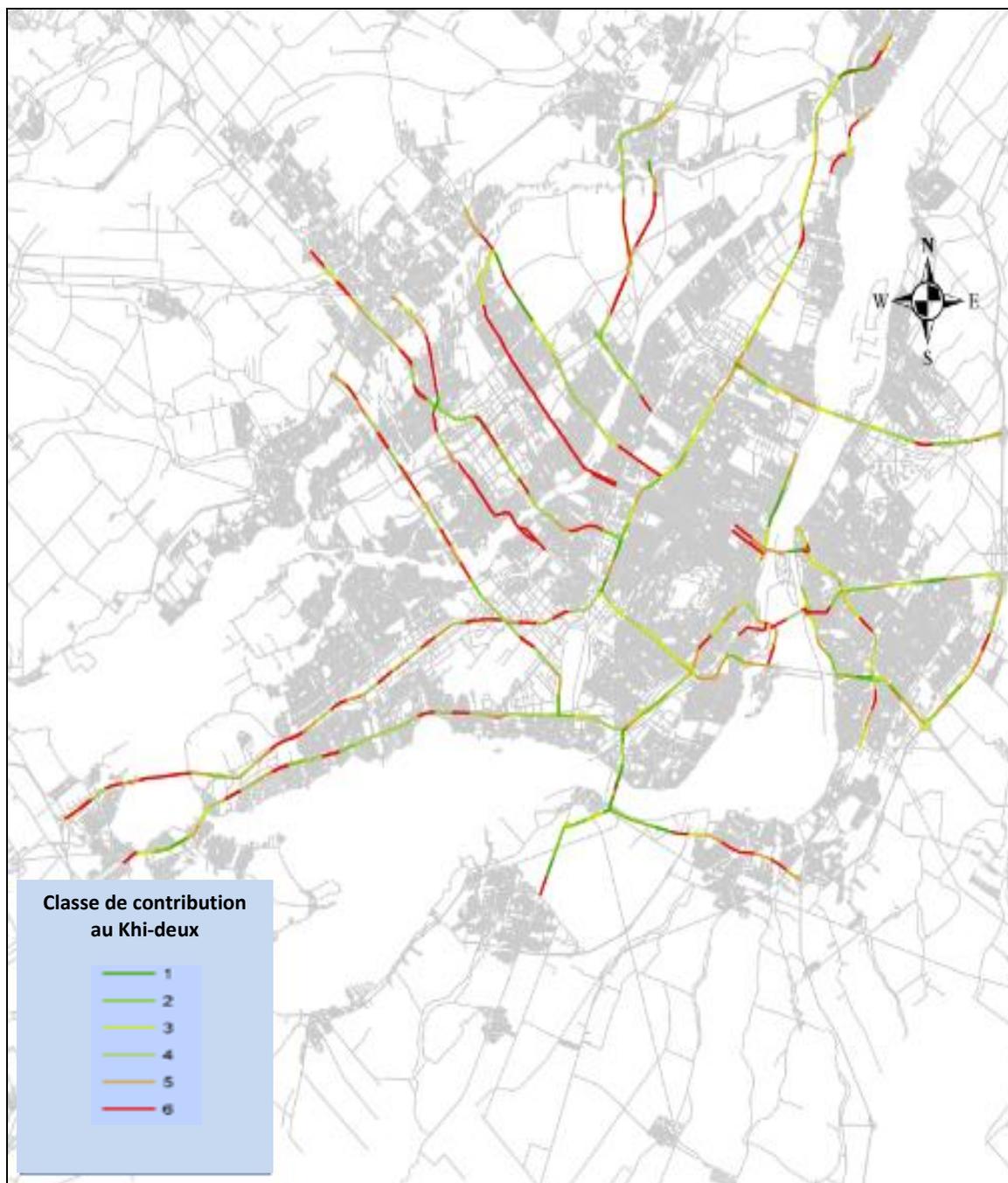
**Tableau 33. Répartition des tronçons de chaque groupe FP selon les groupes TP**

	FP1	FP2	FP3	FP4	FP5	FP6	FP7	FP8	Total
TP1	9%	16%	40%	3%	20%	8%	5%	15%	10%
TP2	26%	1%	7%	66%	4%	0%	50%	3%	36%
TP3	39%	1%	13%	16%	0%	20%	24%	6%	21%
TP4	3%	0%	0%	7%	0%	4%	5%	0%	4%
TP5	10%	1%	7%	5%	8%	8%	9%	3%	7%
TP6	7%	53%	23%	2%	44%	36%	3%	50%	15%
TP7	3%	23%	10%	0%	20%	20%	2%	18%	7%
TP8	2%	3%	0%	0%	4%	4%	3%	6%	2%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%

**Tableau 34. Cartogramme des contributions au khi-deux**

	TP1	TP2	TP3	TP4	TP5	TP6	TP7	TP8
FP1	1	3	5	1	2	4	2	1
FP2	3	5	4	2	3	6	5	1
FP3	5	3	2	2	1	2	1	1
FP4	4	5	2	3	2	4	4	3
FP5	2	3	3	2	1	4	3	2
FP6	1	4	1	1	1	3	3	2
FP7	2	3	2	1	1	4	3	1
FP8	2	4	3	2	2	5	3	3

Le cartogramme du Tableau 34 présente la contribution de chaque paire FP-TP établie à l'aide d'un test de khi-deux. Une discrétisation de la variable « contribution au khi-deux », bâtie sur la synthèse statistique obtenue, a été faite. Six classes des contributions au khi-deux ont été retenues : [Min; décile1[,..., [décile9; Max]. À la lumière de ce tableau, on remarque que les groupes TP2 et TP6 se caractérisent par une forte contribution au khi-deux. On pourrait donc, en théorie, expliquer leur « comportement » de temps de parcours en fonction de leurs facteurs physiques. Par contre, les groupes TP5 et TP8 contribuent au khi-deux moins que les autres groupes TP. Ceci s'explique, conjointement avec le cartogramme de différence entre les proportions observées et les proportions théoriques, par le fait que la différence (surévaluation ou sous-estimation) entre la situation réelle et la situation théorique est plus faible que les autres groupes TP. Les contributions au khi-deux sont également représentées par une carte thématique (figure 5.14), permettant par la suite de ressortir les différentes particularités des groupes TP.



**Figure 69. Carte thématique représentant la contribution au khi-deux**

Cette carte thématique représente le degré de contribution de chaque tronçon au khi-deux. En d'autres termes, il s'agit de présenter les liens entre les groupes TP et FP. À la lumière de cette carte, on peut noter que la partie est de l'autoroute 40 présente peu de corrélations entre les groupes TP et FP (faible khi-deux). Facteurs physiques et temps de parcours ne sont donc pas corrélés. Par contre, l'explication est meilleure dans la partie ouest, puis pour certains segments particuliers comme ceux de la route 335 à Laval. Ces résultats laissent déjà voir que, tel qu'attendu, les facteurs physiques sont corrélés aux

distributions de temps de parcours précédemment modélisés mais que le volume de circulation sur les routes est aussi déterminant (différence entre corrélation observée en périphérie vs sur l'A40).

#### 4.1.6 Prédiction de l'appartenance

Les résultats de ce processus d'analyse ne sont pas particulièrement concluants, car ils ne permettent pas de caractériser les temps de parcours de tous les segments à partir uniquement de leurs caractéristiques physiques. En d'autres mots, les caractéristiques physiques ne sont pas les seuls éléments explicatifs. Si on persiste à vouloir prédire l'appartenance d'un segment à un groupe TP à partir de ses caractéristiques physiques, on peut utiliser l'arbre de décision de la Figure 70, qui nous permettra de déterminer le groupe TP le plus probable.

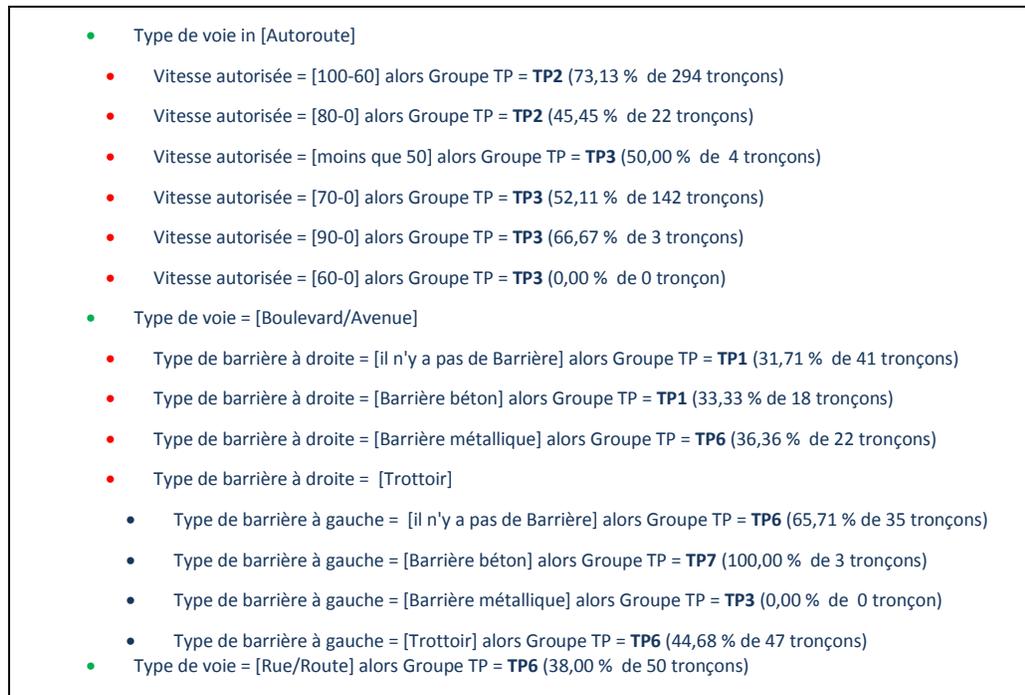


Figure 70. Arbre de décision C4.5 complet

## 4.2 Synthèse

En vue d'améliorer la compréhension des distributions de temps de parcours observées sur les différents tronçons routiers, une base de données de facteurs physiques a été constituée. Différents outils ont été exploités pour ce faire. Un des objectifs était d'identifier le niveau de corrélation entre les facteurs physiques décrivant chaque tronçon et les distributions fréquentielles typiques de temps de parcours (groupements). L'étude se traduit d'abord par la disponibilité d'une banque de données décrivant les tronçons. Elle confirme par ailleurs que les facteurs physiques seulement ne permettent pas de prédire le type de distribution fréquentielle plausible qui sera observée sur un tronçon, tâche souhaitable en l'absence d'observations. Des données complémentaires, le volume sur le tronçon par exemple, permettraient sans doute de raffiner l'approche.

## 5. Exploration des données d'incidents

Une banque de données d'incidents a aussi été rendu disponible dans le cadre de l'étude des temps de parcours. Il semble évident que l'occurrence d'incidents a un effet sur les conditions de circulation. Jusqu'à présent, ces données n'ont jamais été exploitées pour fins d'analyse ou modélisation. Cette section vise donc à illustrer certaines de leurs potentialités. Évidemment, dans leur forme actuelle, les bases de données doivent a priori être traitées avant d'être exploitables, ce qui a été fait pour le premier ensemble de données auquel nous avons eu accès.

Les sections qui suivent présentent d'abord la structure et le contenu de la base de données.

### 5.1 Base de données

#### 5.1.1 Concepts

Les principaux termes techniques utilisés sont :

- Incident : ensemble des incidents de la base de donnée; il existe 62 types d'incidents différents dans la base dont 5 types majeurs (accident, débris, panne, piéton et véhicule);
- Accident : sous-ensemble des incidents répertoriés comme accident dans la base de données;
- Tronçon : les mêmes tronçons d'un kilomètre sont utilisés pour l'analyse.

#### 5.1.2 Variables

Les premiers fichiers de données reçus ont été intégrés sous forme de base de données rassemblant les incidents survenus entre juillet 2001 et décembre 2004 sur les routes montréalaises. Les incidents ont d'abord été géolocalisés à partir d'une déclaration spatiale parfois floue (autoroute sur laquelle est survenu l'incident et route transversale la plus proche). Ceux localisés sur les tronçons d'un kilomètre ont été retenus pour fins d'analyse.

On peut catégoriser les variables de la base de données en 3 groupes : les variables spatiales, les variables d'identification et les variables temporelles. Le tableau qui suit présente les champs de la base de données étudiée.

Tableau 35. Contenu de la base de données

Variable	Libellé	Type
<i>IdLieux</i>	Numéro d'identification de lieux	Spatiale
<i>Autoroute</i>	Nom de l'autoroute	
<i>blvd</i>	Nom du boulevard perpendiculaire le plus proche	
<i>Numrecord</i>	Numéro d'enregistrement de l'incident	Identificatrice
<i>Type</i>	Type de l'incident (issue de la typologie du ministère)	

<i>Datedeb</i>	Date de début d'incident	Temporelle
<i>Datefin</i>	Date de fin d'incident	
<i>Annebdd</i>	Anne de l'incident	
<i>Hredeb – Mindeb – Secdeb</i>	Heures, minutes et secondes du début de l'incident	
<i>Hrefin – Minfin - Secfin</i>	Heures, minutes et secondes de la fin de l'incident	
<i>Durée</i>	Durée de l'incident (en min)	
<i>catdurée</i>	Catégorie de durée de l'incident, découpée en 6 modalités :  <ul style="list-style-type: none"> <li>- <math>x &lt; 5\text{min}</math>,</li> <li>- <math>5\text{min} &lt; x &lt; 15\text{min}</math></li> <li>- <math>15\text{min} &lt; x &lt; 30\text{min}</math></li> <li>- <math>30\text{min} &lt; x &lt; 60\text{min}</math></li> <li>- <math>1\text{h} &lt; x &lt; 2\text{h}</math></li> <li>- <math>x &gt; 2\text{h}</math></li> </ul>	
<i>Période</i>	Indicateur de période de la journée : 4 modalités :  PPAM : 7h à 9h59 (Période de pointe AM)  PHPJ : 10h à 15h59 (Pointe creuse)  PPPM : 16h à 18h59 (Pointe PM)  SetN : 19h à 6h59 (Soit et Nuit)	

## 5.2 Analyse descriptive

Différentes analyses ont été conduites pour comprendre le contenu de la base de données et développer une compréhension du phénomène d'occurrence des incidents sur les tronçons autoroutiers.

### 5.2.1 Distribution des incidents

La Figure 71 présente la répartition annuelle des incidents pour les années 2002, 2003 et 2004. On peut relever, qu'entre janvier et juillet, ainsi qu'en novembre et décembre, le nombre d'incidents disponibles dans la base de données a connu une évolution croissante chaque année, tandis que les mois d'août, septembre et octobre connaissent une tendance inverse. Le mois d'août 2004 a un nombre étonnamment bas d'incidents.

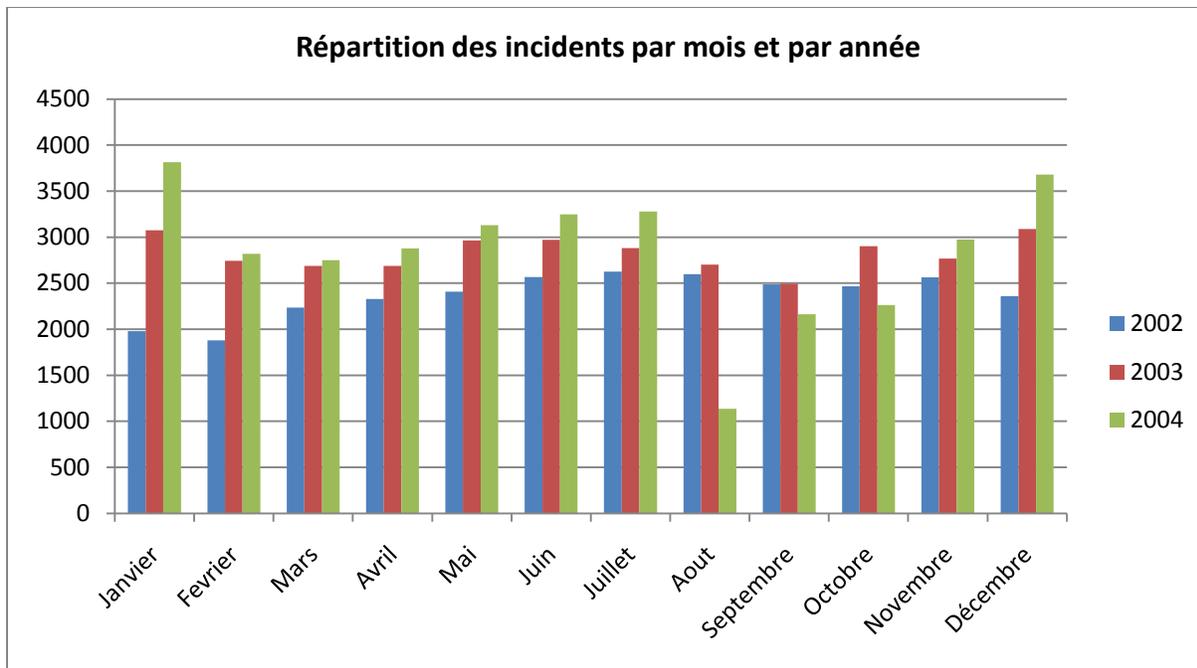


Figure 71. Distribution de l'ensemble des incidents entre 2002 et 2004, par mois de l'année

La Figure 72 présente la répartition des incidents par jour de la semaine, toutes années et tous mois confondus. Grosso modo, on observe que les quantités quotidiennes d'incidents sont moindres lors des fins de semaines et que le vendredi présente plus d'incidents que les autres jours de la semaine. Ceci est conforme à notre connaissance qualitative du phénomène.

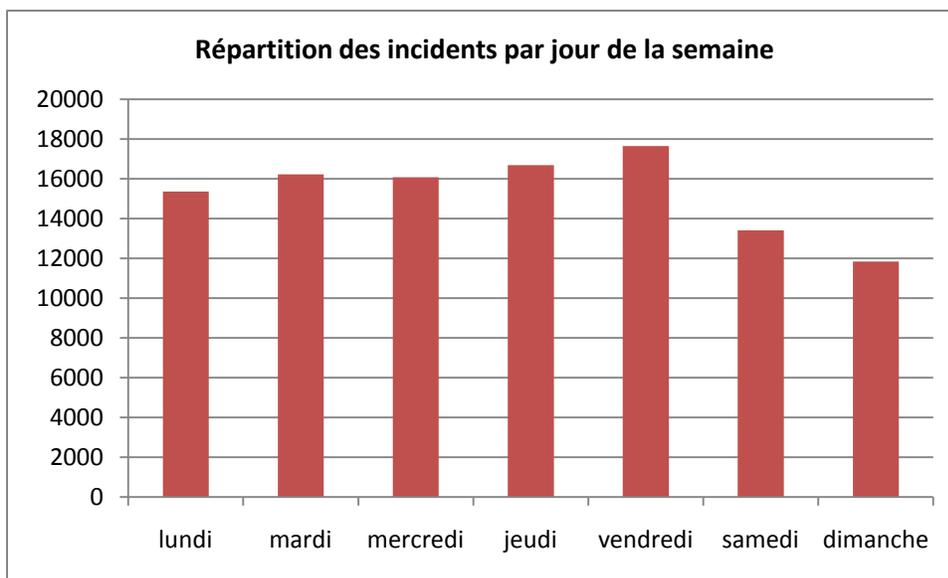


Figure 72. Répartition des incidents selon le jour de la semaine (données 2001-2004)

La Figure 73 présente la distribution temporelle des incidents, toutes années et tous mois confondus, par heure de la journée. Tel qu'attendu, le nombre d'incidents répertoriés est supérieur pendant les

périodes de pointe, celle du soir étant plus importante que celle du matin. Il y donc une corrélation évidente entre le volume de circulation et l'occurrence d'incidents.

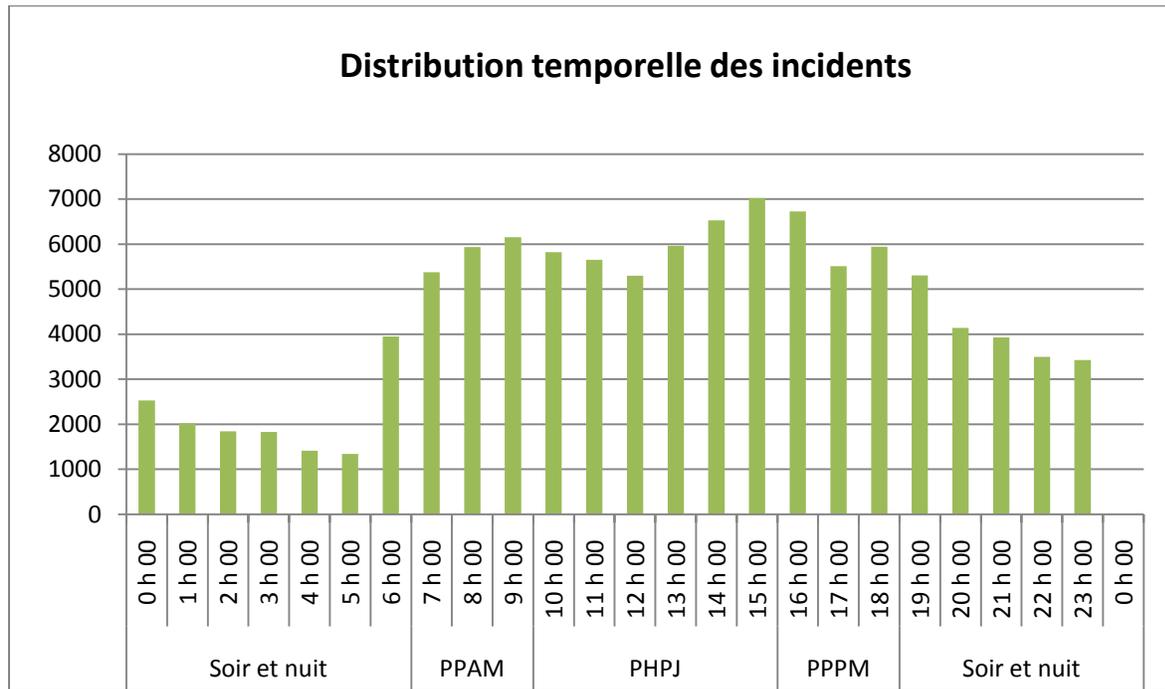


Figure 73. Répartition des incidents selon l'heure de la journée

La Figure 74 présente la répartition des incidents par grand axe autoroutier (qui sont de longueurs différentes). Sans surprise, c'est l'axe de l'autoroute 40 qui compte le plus d'incidents inventoriés dans la base de données. L'étude de quelques grands axes par tronçon d'un kilomètre permet une comparaison plus juste des quantités d'incidents. À titre d'exemple, la Figure 75 présente la répartition des incidents sur les différents tronçons de l'A-40, dans les deux directions. On observe en effet que le nombre d'incidents répertoriés varie en fonction des tronçons.

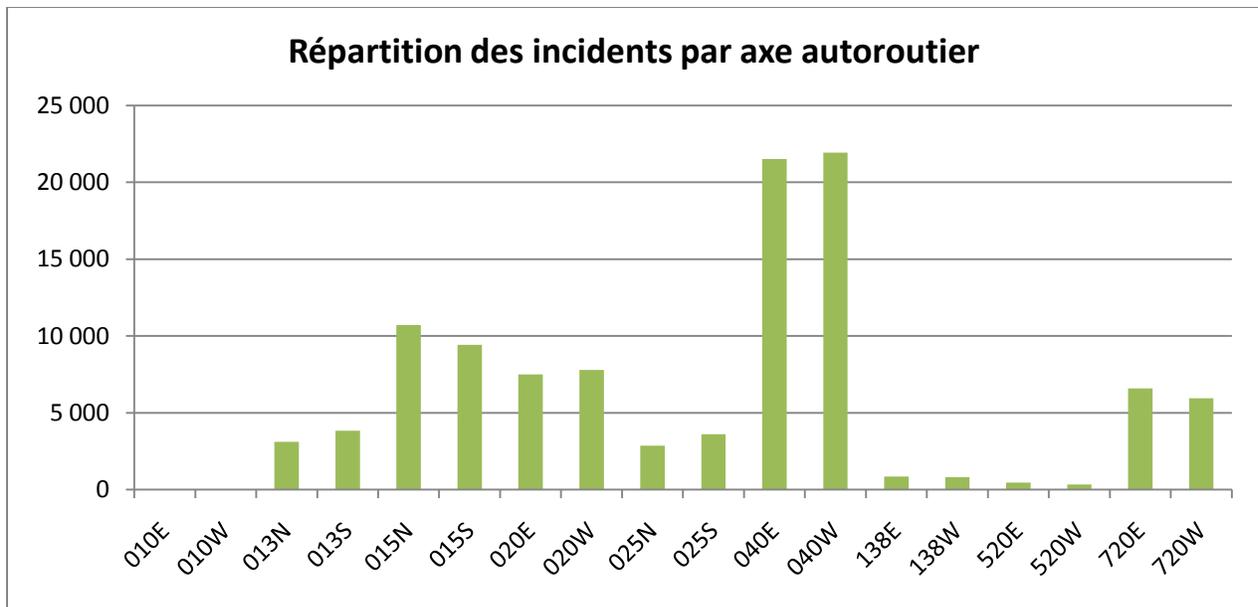


Figure 74. Répartition des incidents par axe routier (2001 à 2004)

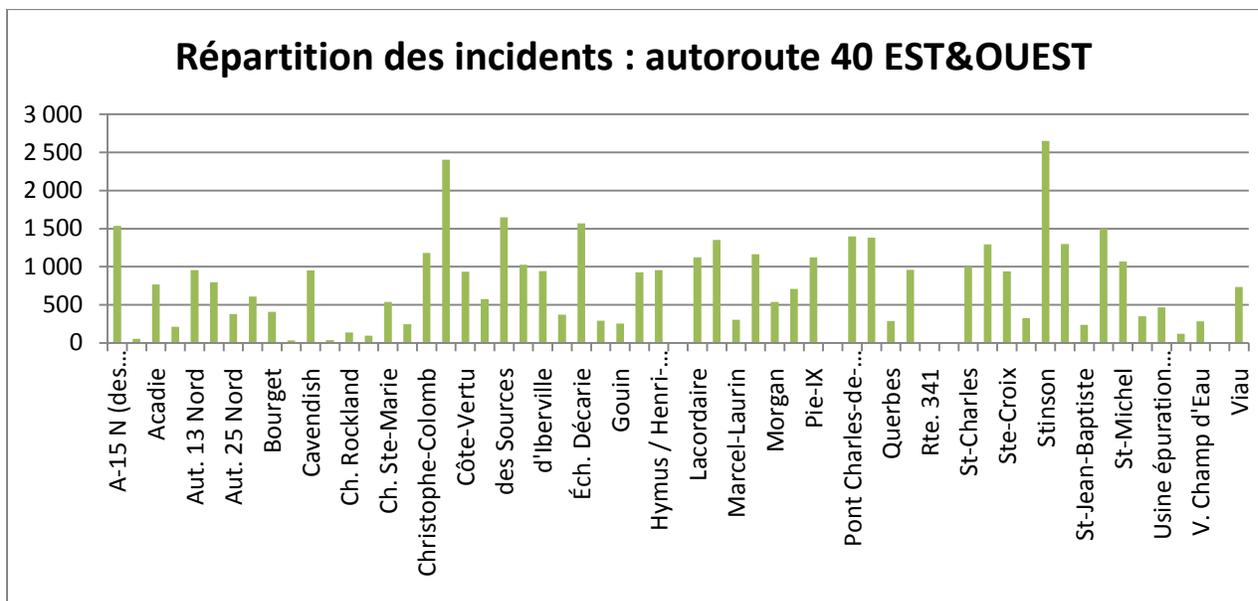


Figure 75. Répartition des incidents : autoroute 40

La courbe de Lorenz, souvent utilisée pour mesurer des iniquités dans des distributions, est utilisée pour qualifier la dispersion des incidents sur les différents tronçons autoroutiers du réseau. La Figure 76 présente cette courbe qui met en relation le pourcentage cumulé des incidents par rapport au pourcentage cumulé des tronçons. Cette courbe permet d'observer que presque 60% des incidents sont répertoriés sur 20% des tronçons et que 80% des incidents sont répertoriés sur 40% des tronçons. C'est donc dire que la répartition spatiale des incidents n'est pas uniforme et que certains tronçons sont plus susceptibles d'être le lieu d'un incident que d'autres.

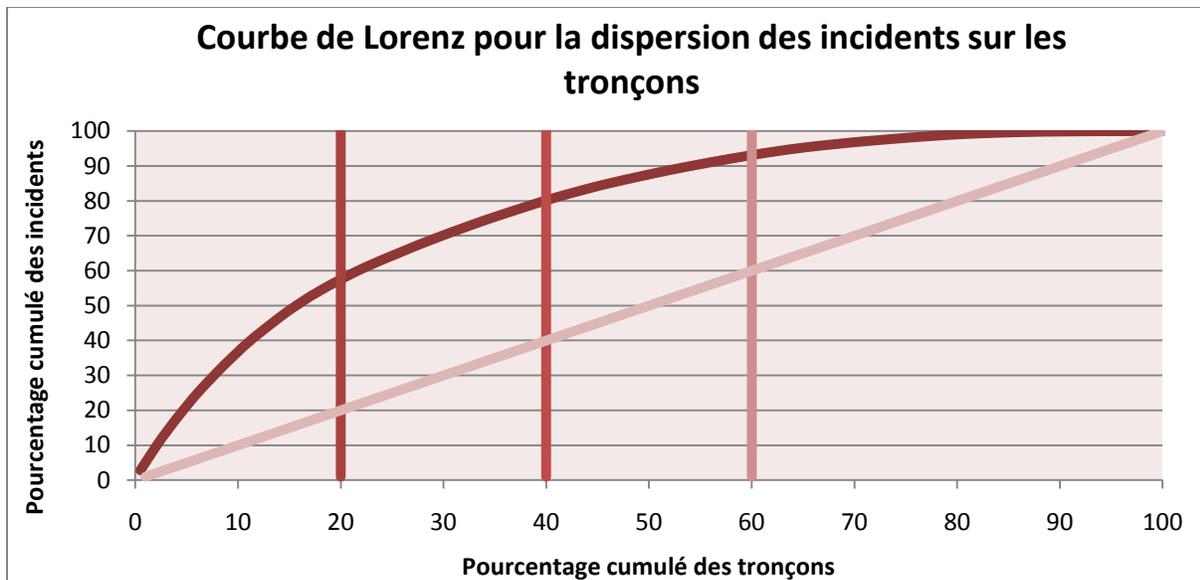


Figure 76. Dispersion des incidents sur le réseau : courbe de Lorenz

### 5.2.2 Analyse de la durée d'incident

Les analyses portent maintenant sur les durées d'incidents. La Figure 77 présente la distribution fréquentielle des incidents selon leur durée (par classe de durée). La durée moyenne d'un incident est de 26.7 minutes. En outre, 46% des incidents ont une durée de moins de 15 minutes, 78% de moins de 30 minutes et 95% de moins de 60 minutes. La qualité de cette information est difficile à évaluer car peu de détails sur le processus de collecte de données sont disponibles.

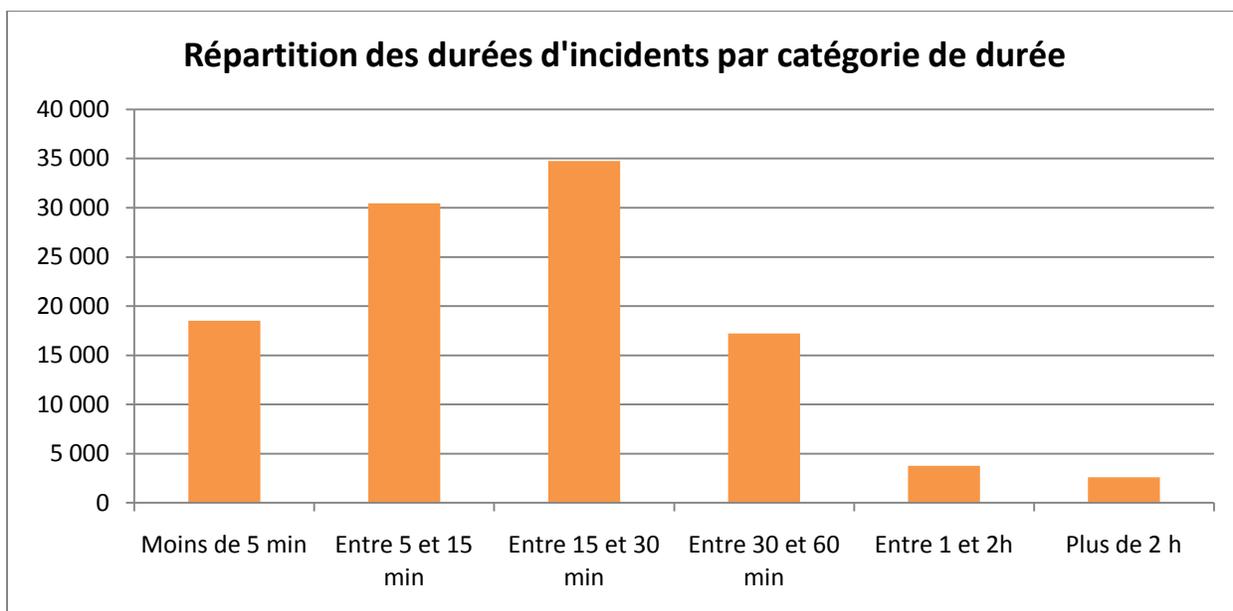


Figure 77. Distribution fréquentielle des incidents selon leur durée

### 5.3 Analyse classifiée

La base de données mise à disposition comprenait une variable indicatrice du type d'incident. Dans cette base de données on retrouve 60 types d'incidents différents.

La Figure 78 présente l'occurrence des cinq types d'incidents les plus fréquents dans la base de données, qui cumulent 89% des incidents observés. À elles seules, les pannes rassemblent 53% des incidents.

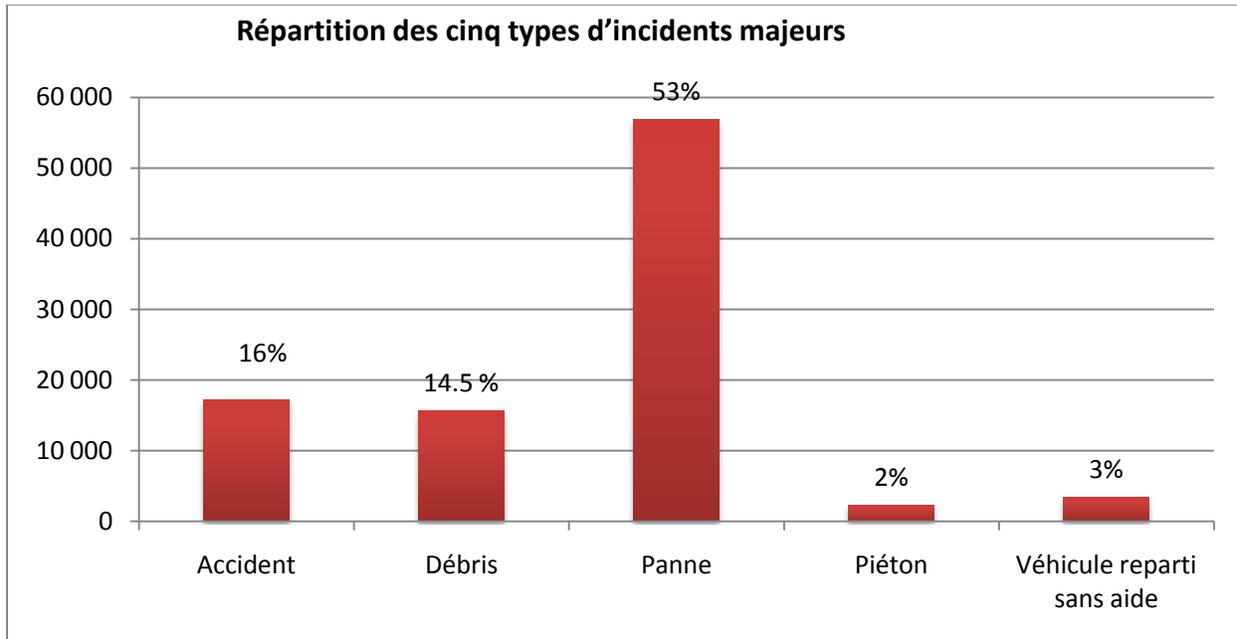


Figure 78. Occurrence des cinq types d'incidents les plus fréquents

La Figure 79 présente, pour les cinq types d'incidents les plus fréquents, l'évolution de leur fréquence ainsi que de leur durée moyenne pour les années observées (la quantité n'est pas montrée pour 2001 puisque seule une partie des données a été rendu disponibles). On peut remarquer une certaine fluctuation des durées d'incidents pour les types « véhicule reparti sans aide », « piéton » et « débris » contrairement aux types « panne » et « accident ». Les accidents se distinguent quant à eux avec une durée moyenne élevée et un nombre plutôt constant dans le temps.

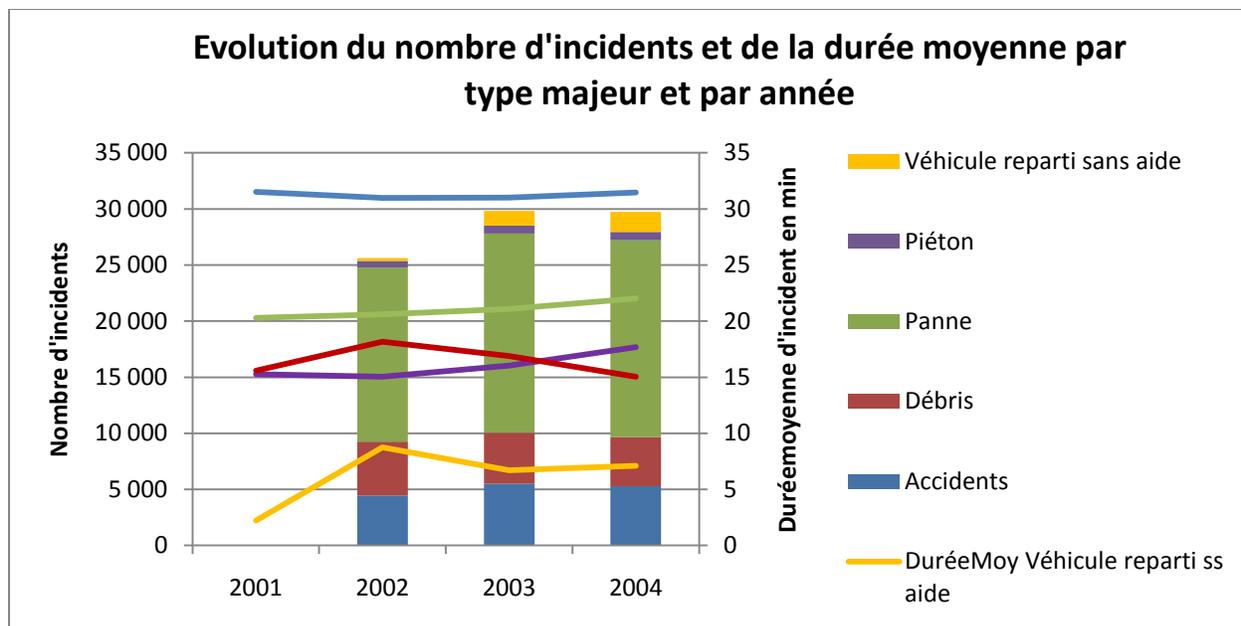


Figure 79. Évolution du nombre d'incidents par année et de la durée d'incident moyenne

Nous avons jusqu'à présent traité des données d'incidents. On rappelle que l'on entend par incident l'ensemble des types d'incidents présents dans la base de données (panne, débris, accidents...)

Nous allons donc maintenant nous intéresser uniquement aux données concernant les accidents.

## 5.4 Étude des incidents de type accident

### 5.4.1 État des connaissances

Après une consultation des articles de modélisation statistiques des incidents de la route, il paraît que la grande majorité de ceux-ci visent à produire des modèles de développement d'accidents. La typologie des différents incidents n'est pas prise en compte et on ne s'intéresse exclusivement qu'aux accidents. Dans la plupart des études ayant pour but la modélisation des accidents routiers on retrouve des recherches basées sur la modélisation par le biais de la loi de Poisson-Gamma (Abbess et al. 1981 ; Persaud and Dzibik 1993 ; Miaou and Song 2005). C'est la méthode statistique la plus fréquemment utilisée.

Certaines études tentent par ailleurs de modéliser les accidents de la route par le biais des méthodologies liées aux *processus aléatoires* (Miaou and Lord, 2003), ou encore en faisant appel aux *estimations de modèles linéaires généralisés* – GLM – (Lord and Persaud, 2000, Abdel-Aty and Addella, 2004). D'autres études se concentrent sur des modèles de prédiction de risques d'accident en fonction de plusieurs variables, avec comme variables dominantes le trafic et certaines caractéristiques des routes (Wichert S. et Cardozo J. 2007, Stefan C. 2006).

Dans le cadre de cette recherche et au vue des attributs de la base de données disponible, il serait intéressant de tenter une tout autre méthode.

En effet en considérant dans une étude les 100 tronçons les plus représentés en matière d'accidents, et en formant une nouvelle base de données « chronologique » prenant en compte, pour chacun des 100 tronçons, les épisodes d'incidents survenus et en calculant la durée de non-incident, grâce au modèle de

Cox, il est possible de calculer les risques relatifs en fonction des caractéristiques temporelles des incidents et des caractéristiques physiques des tronçons.

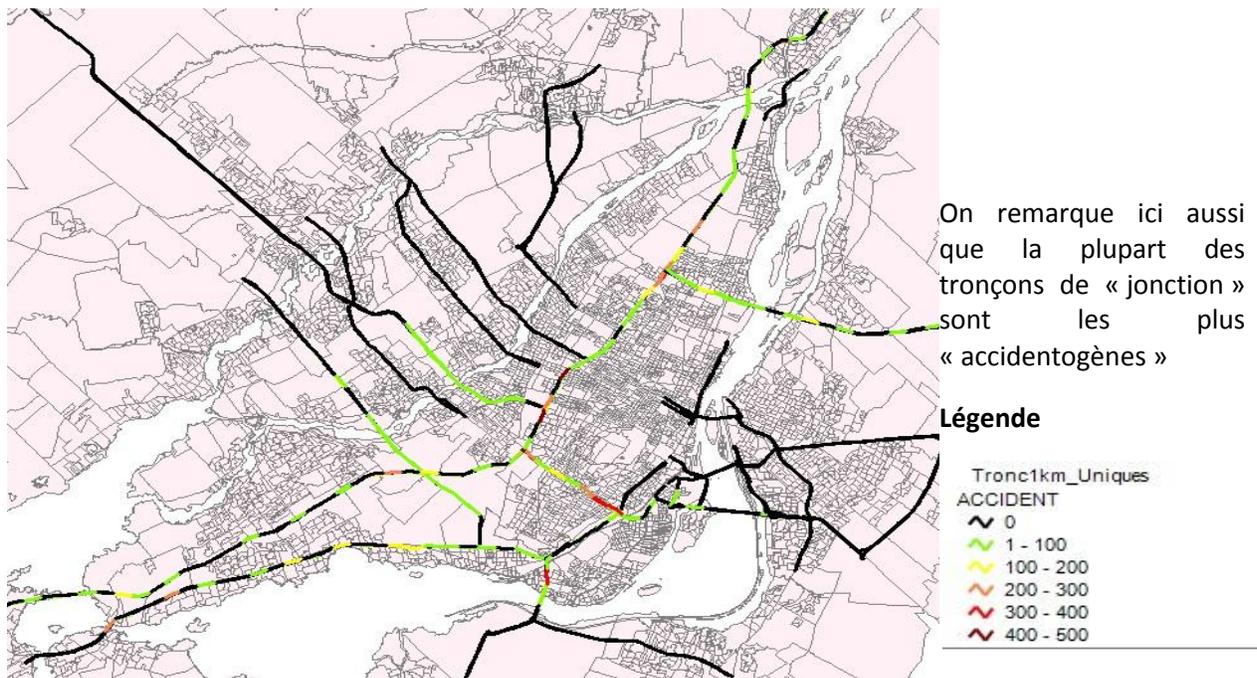
## 5.5 Analyse des données d'accidents

Les faits saillants des analyses conduites sur la base de données d'accidents sont présentés.

D'abord, une comparaison des durées moyennes des incidents, selon leur type, est effectuée au Tableau 36. La Figure 80 donne un aperçu des tronçons les plus accidentogènes par une cartographie du nombre d'accidents répertoriés par tronçon de 2002 à 2004.

**Tableau 36. Durée moyenne des incidents par type**

Type d'incident	Durée moyenne (minutes)
<b>Accident</b>	<b>31,20</b>
Débris	16,60
Panne	21,15
Piéton	16,22
Véhicule reparti sans aide	7,07



**Figure 80. Cartographie des accidents sur le réseau autoroutier de Montréal entre janvier 2002 et décembre 2004**

### 5.5.1 Dispersion des accidents : courbe de Lorenz

Le même concept de courbe de Lorenz est utilisé pour caractériser la dispersion des accidents sur les différents tronçons. La courbe a une allure très semblable à celle réalisée pour l'ensemble des incidents : 20% des tronçons subissent 60% des incidents et 40% des tronçons en subissent 80%.

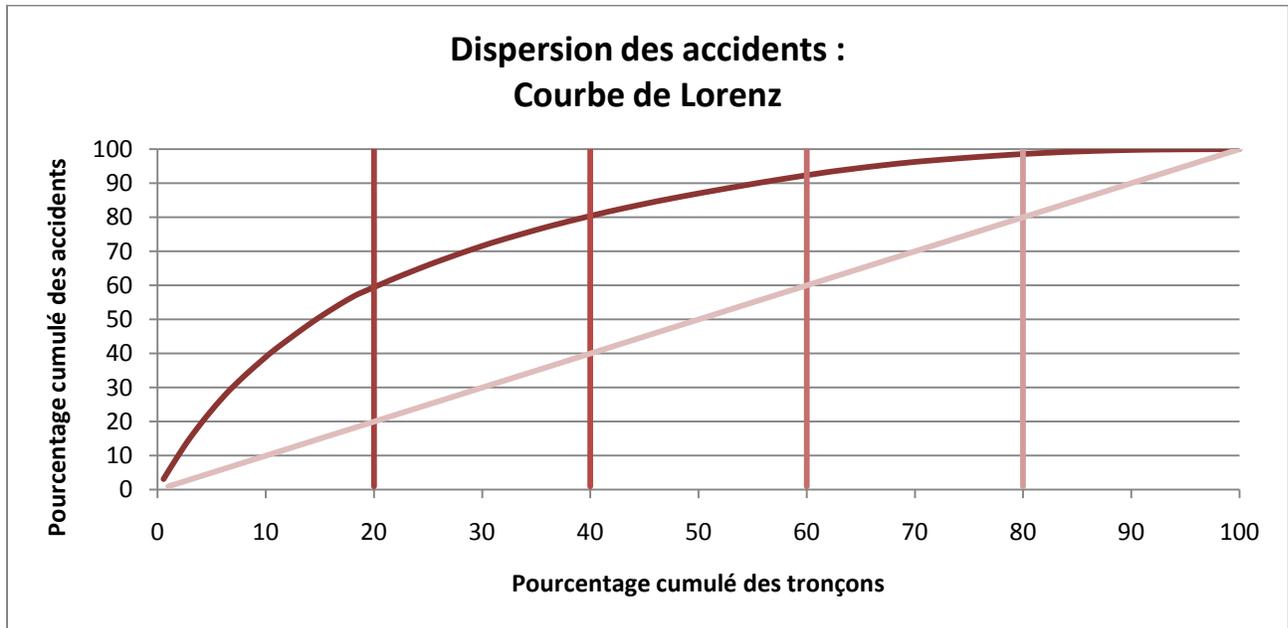


Figure 81. Courbe de Lorenz pour les accidents

## 5.6 Modélisation de survie d'un accident

L'étude de la base de données et les différents résultats obtenus par l'analyse descriptive nous incitent à explorer ces données d'incidents à l'aide d'une modélisation de survie des accidents.

En effet, considérer les tronçons comme des individus et transformer la base de données afin d'obtenir une base chronologique des accidents par tronçon nous permet de créer une variable  $\Delta T$ , qui représente la durée sans accident sur un tronçon. C'est cette variable qui sera à la base de notre modélisation.

De plus, une certaine caractérisation physique des tronçons a été développée. Ainsi la modélisation de Cox nous permettrait d'éditer des rapports de risques par rapport à différentes variables de caractérisation physique des tronçons.

### 5.6.1 Filtrage et transformation de la base de données

Pour la création de la base de données chronologique, des choix méthodologiques ont été faits. En effet, nous avons choisi de filtrer les tronçons et de ne garder que les quelques 100 tronçons les plus accidentogènes (en réalité, 97 tronçons dont le nombre d'accidents dépasse la barre symbolique de 50 sont examinés). Ce choix est justifié par le fait que 80% des accidents surviennent sur 40% des tronçons comme nous l'avons établi précédemment.

De la base initiale, on obtient donc une base de données regroupant 97 tronçons dépassant 50 accidents.

La base a été ensuite triée en deux étapes. Une première étape a permis de trier par tronçon (individu) puis dans une seconde étape, la sous-variable de tri par date de survenue des accidents. On obtient donc une base de données avec 97 tronçons et pour chaque tronçon tous les accidents par ordre chronologique.

Pour fins d'expérimentation, les données de 2001 sont retirées afin de ne garder que les années complètes (2002-2003-2004). Pour modéliser la durée sans accident, une variable *jourdepuis* est créée, celle-ci comptabilisant le nombre de jours écoulés entre la date d'initialisation que nous avons choisie (01/01/2002) et la date de l'accident.

La variable à modéliser est ensuite construite (intervalle de temps sans accident). On a donc obtenu une variable comptabilisant le nombre de jours écoulés entre deux accidents sur un même tronçon, c'est notre variable à expliquer : *tps*.

Différentes variables de caractérisation physique des tronçons ont été incluses à la base de données. L'incorporation de ces données s'est faite dans le but d'éditer des rapports de risque et d'étudier les possibles effets de chacune des modalités de ces variables dans la survie d'un accident.

Les variables de caractérisation sont présentées ci-dessous :

**Tableau 37. Variables de caractérisation physique des tronçons**

Variables	Type	Libellé
<i>Sens_VC</i>	« dummy »	Indique le sens du tronçon (Vers centre Hors-centre)
<i>Nssc_0 - Nssc_1 - Nssc_2</i>	« dummy »	Nombre de sorties sans contrôle
<i>Essc_0 - Essc_1 - Essc_2</i>	« dummy »	Nombre d'entrées sans contrôle
<i>Inter_ctrl_0 - Inter_ctrl_1 - Inter_ctrl_2</i>	« dummy »	Nombre d'intersection avec contrôle
<i>Nb_voie_2 - Nb_voie_3 - Nb_voie_4</i>	« dummy »	Nombre de voies
<i>V_Max_50 - V_Max_60 - V_Max_70 - V_Max_100</i>	« dummy »	Vitesse maximale autorisée

Les archives météo des chutes de neige et de pluie étant disponibles sur internet, nous avons décidé de les incorporer également en variable dichotomique. Les deux variables créées se présentent comme suit :

Tableau 38. Variables décrivant les conditions météorologiques

Variable	Critère	Modalités
neige	Chute de neige dépassant 10 cm	1 si oui 0 sinon
pluie	Niveau de pluie dépassant 10 mm	1 si oui 0 sinon

Enfin une variable relative au sens du tronçon (vers centre-ville ou hors centre-ville a été ajoutée (*Sens\_VC* 1 si vers centre 0 sinon)

### 5.6.2 Analyse deltaT à travers l'effet d'une variable

Nous allons présenter ici un graphique que nous avons tracé pour tenter de représenter les deltaT à travers la chronologie de l'étude et par des exemples des différentes modalités d'une variable donnée. La Figure 21 permet d'observer la fréquence des accidents en fonction du temps, pour 9 tronçons différents représentant trois classes de nombre de voies. Une croix sur le graphique indique qu'un accident a été répertorié sur le tronçon.

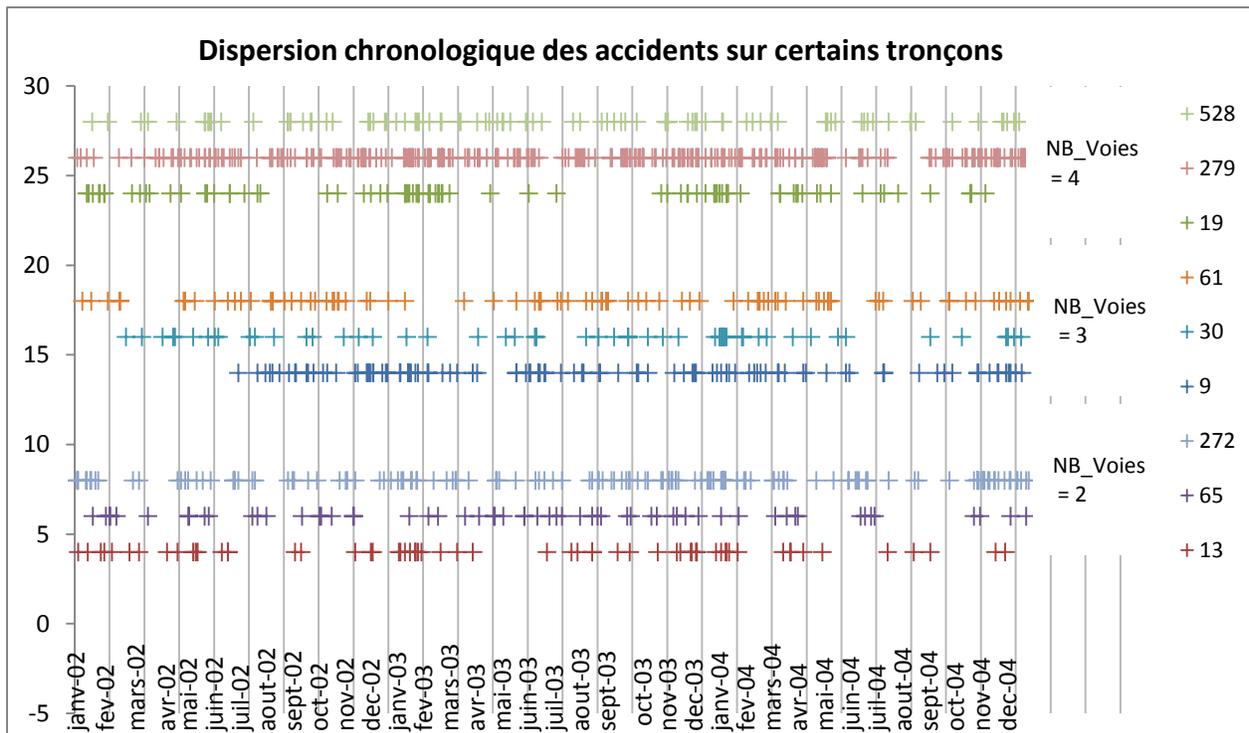


Figure 21. Dispersion des relevés d'accidents par tronçons entre 01/01/2002 et 31/12/2004

On a donc ici les relevés des accidents sur 9 tronçons différents. Trois groupes de tronçons dont la seule variable qui varie est le nombre de voies de ce tronçon. On remarque à la lecture du graphique que pour

les traces des tronçons à 4 voies, notamment le tronçon 279, les croix sont très rapprochées contrairement aux traces des trois tronçons à 2 voies.

C'est justement l'écart entre chaque croix (chaque accident) qui sera modélisé. En effet le temps écoulé entre chaque accident sur un même tronçon est représenté en jours par notre variable *tps*. Le modèle de Cox nous permettra de quantifier statistiquement les hypothèses que nous pouvons suggérer au regard de ce graphique.

### 5.6.3 Modèle de Kaplan et estimation des « probabilités de survie » d'accidents à travers le temps

Le modèle de Kaplan-Meier est la première partie de l'analyse de durée. Cette méthode de modélisation caractérisée par le processus SAS « proc lifetest » permet d'éditer les courbes de survie de la variable numérique deltaT qui représente, on le rappelle, la durée sans accident sur un tronçon.

A priori, cette méthode ne tient pas en compte initialement les variables du tronçon. En effet, la formulation mathématique tient compte des fluctuations de *x* qui est la durée à estimer au cours du temps. Il est possible toutefois d'étudier ponctuellement l'effet d'une variable par des méthodes de stratification que nous verrons ci-après.

Le tracé de la courbe de Kaplan se fait conjointement sous SAS, par l'édition en parallèle du tableau de l'évolution des probabilités pour les différentes valeurs de *x*. Ci-dessous une séquence de ce tableau.

**Tableau 39. Tableau des estimations de survie de Kaplan-Meier**

Estimations de survie de Kaplan-Meier						
tps		Survie	Défaillance	Erreur type de survie	Nombre ayant échoué	Nombre restant
<b>0.000</b>		0.8819	0.1181	0.00301	1356	10126
<b>1.000</b>		0.7274	0.2726	0.00416	3130	8352
<b>4.000</b>		0.4651	0.5349	0.00465	6142	5340
<b>11.000</b>		0.2136	0.7864	0.00383	9029	2453

La variable *tps* représente le temps écoulé entre deux incidents sur un même tronçon. On peut remarquer qu'au bout de *tps*=4 (4 jours) la probabilité de n'avoir pas terminé cette période d'incident, que l'on peut reformuler en probabilité que le tronçon n'ait pas encore enregistré un accident quatre jours après en avoir connu un, est de 0.4651. On remarque une décroissance logique de la probabilité en fonction de *tps*. Ce tableau est traduit par le tracé de la courbe de survie (Figure 82) qui estime la valeur de la probabilité de n'avoir pas enregistré un incident, pour un tronçon, en fonction du temps.

## Estimation de Kaplan-Meier : durée sans accident

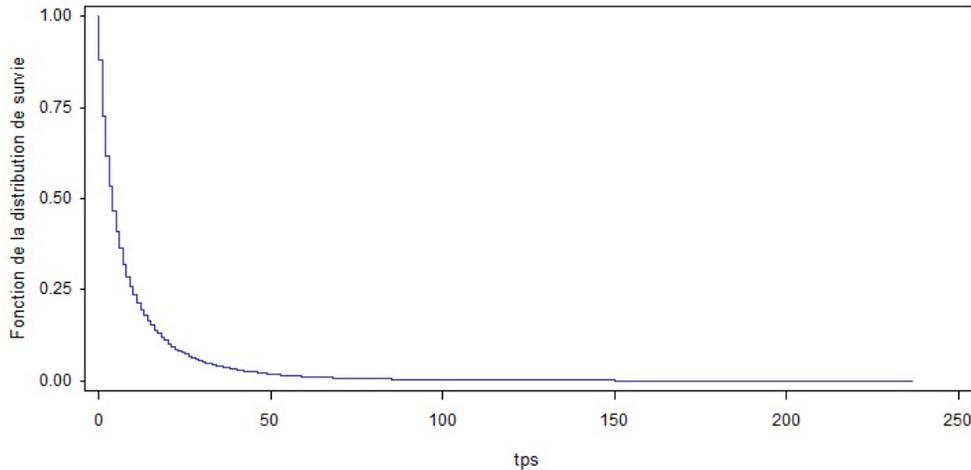


Figure 82. Courbe de survie pour la durée sans accident

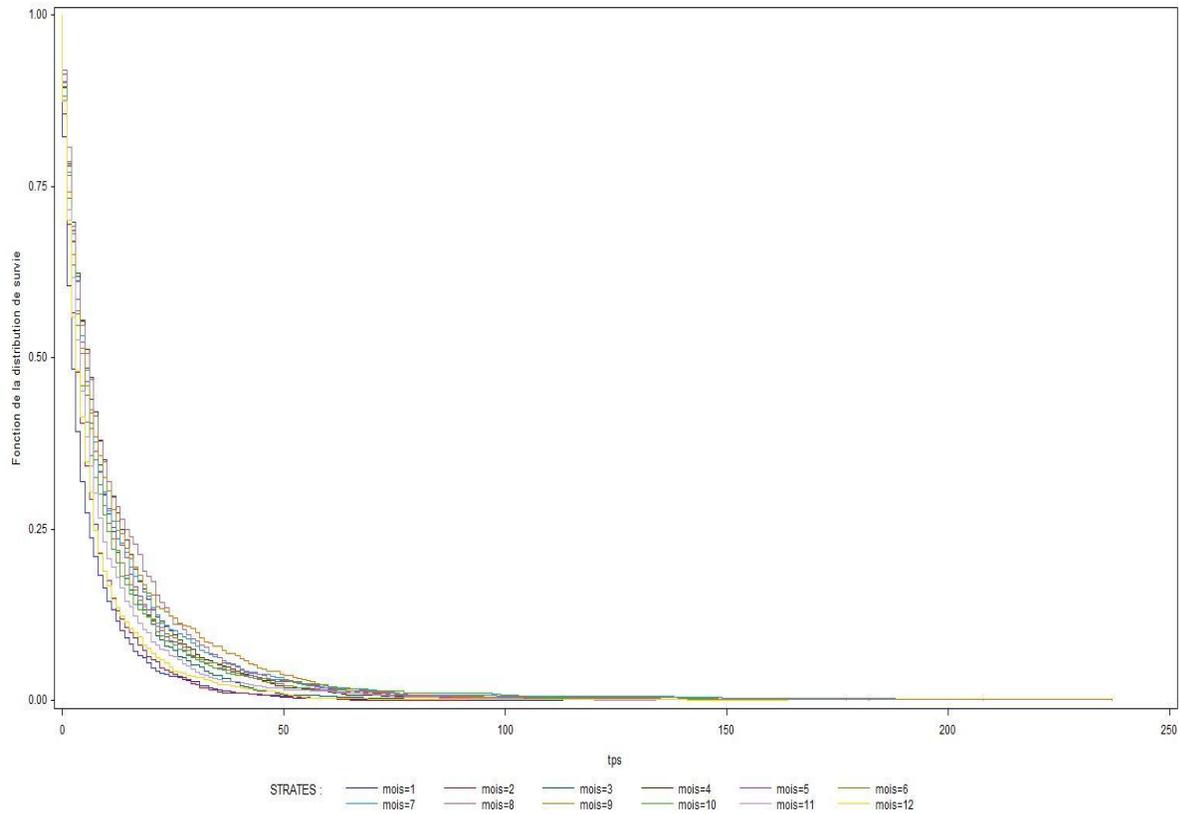
On remarque que cette courbe est fortement décroissante entre  $t=0$  et  $t=30$  puis se stabilise. Ainsi la probabilité de voir une période sans accident durer plusieurs jours est minime; en outre, la probabilité de dépasser le cap des 50 jours sans accident sur un même tronçon est pratiquement nulle.

Cependant, ces courbes, étudiées par strate, peuvent nous renseigner sur l'effet des variables sur la probabilité qu'un accident survienne dans le temps.

Nous avons en effet tracé 12 courbes de survie correspondant chacune aux données de chaque mois de l'année. Ces courbes sont présentées à la Figure 83. On remarque que les trois courbes qui se distinguent par leur décroissance plus rapide correspondent aux mois de décembre, janvier et février soit les trois principaux mois d'hiver.

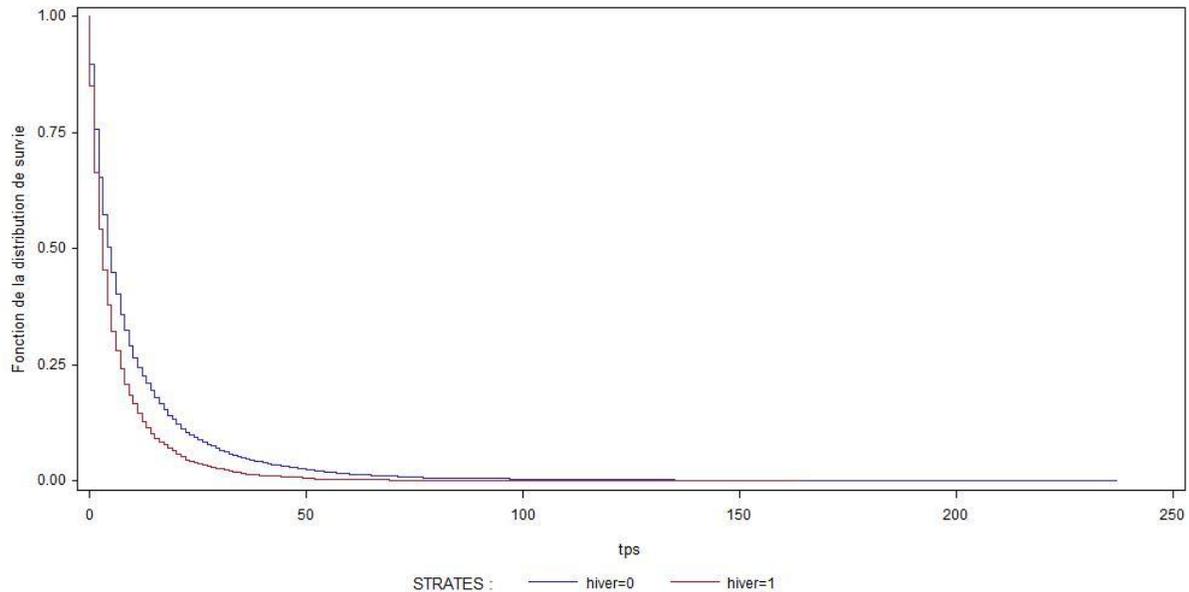
En regroupant de manière plus grossière les 9 autres courbes, on remarque un certain effet de saison par le fait que les courbes des mois de la même saison ont des comportements quasi similaires.

## DUREE SANS ACCIDENTS



**Figure 83. Courbe de survie pour la durée sans accident, stratifiée par mois de l'année**

Afin d'observer plus facilement le décalage des comportements de cette durée sans accident entre les périodes hivernales et non hivernales, nous avons tracé les deux courbes regroupant les données de chaque période (Figure 84).



**Figure 84. Courbe de survie pour la durée sans accident, stratifiée par hiver/autres saisons**

L'hiver concerne la période du 15 décembre au 15 mars (soit la date légale d'obligation de pneus d'hiver sur les véhicules circulant au Québec). La deuxième courbe (bleu) regroupe l'ensemble des données entre le 16 mars et le 14 décembre.

On remarque ici une nette différence pour des durées comprises entre 10 et 40 jours. En hiver, la multiplication des accidents fait chuter la probabilité d'avoir plusieurs jours sans accident sur un même tronçon. L'effet de l'hiver est à relativiser car la définition de l'hiver n'est pas ici météorologique mais temporelle. Les données météorologiques seront étudiées par le biais du modèle de Cox.

#### 5.6.4 Modèle de Cox: rapports de risques entre les caractéristiques physiques des tronçons

Les modèles de Cox ont été obtenus sous SAS grâce au processus « Proc Phreg ». Cette procédure permet de comparer les effets des variables de chaque tronçon dans la survie d'un accident.

Cette méthode nécessite l'insertion des variables sous forme « dummy ». Les choix des variables de référence a été fait arbitrairement :

- *Sens* : le sens Hors centre a été pris comme référence
- Nombre de sortie sans contrôle : ref= 1
- Nombre d'entrées sans contrôle : ref = 1
- Nombre d'intersection sans contrôle : ref = 0
- Nombre de voie : ref = 2
- Vitesse maximale autorisée sur le tronçon : ref = 50km/h

### Test d'ajustement du modèle

Avant de procéder au calcul des rapports de risque, il est utile de présenter le test d'ajustement du modèle :

→ Test :

- ✗  $H_0 : \beta_1 = 0, \dots, \beta_p = 0$
- ✗  $H_1 : \text{Au moins un des coefficients est non nul}$

Tableau 40. Sortie SAS – test d'hypothèse BETA=0

Test de l'hypothèse nulle globale : BETA=0			
Test	Khi-2	DDL	Pr > Khi-2
Likelihood Ratio	630.0201	12	<.0001
Score	628.9554	12	<.0001
Wald	617.7185	12	<.0001

On peut donc rejeter l'hypothèse  $H_0$  à 99%. Ainsi on rejette l'hypothèse de nullité simultanée des coefficients, au moins une des variables semble être déterminante dans la survie d'un accident.

### Rapports de risques

On peut donc maintenant s'intéresser aux rapports de risques édités par cette procédure.

Tableau 41. Rapports de risques issus de la procédure

Variabiles	Modalités	Estimation	Hazard Ratio
Sens	VC	-0.09696*	0.908
	HC	REF	REF
Nssc	0 sortie SC	0.02061	1.021
	1 sortie SC	REF	REF
	2 Sortie SC	-0.0008059	0.999
Essc	0 entrée SC	-0.02100	0.979
	1 entrée SC	REF	REF
	2 entrées SC	0.13945**	1.150
Intet_ctrl	0 intersection AC	REF	REF
	2 intersections AC	-0.81036**	0.445
	4 intersections AC	0.28775**	1.333
Nb_voie	2 voies	REF	REF
	3 voies	-0.02315	0.977
	4 voies	-0.31220**	0.732
Vitesse Max	50	REF	REF
	60	-0.26591**	0.767
	70	0.02267	1.023
	100	-0.40346**	0.668

Les rapports présentant un rapport significatif ont été surlignés (\*\*1% et \* : 5%). Parmi ces rapports de risques, la valeur du « hazard ratio » de la modalité correspondant à 4 intersections sur les tronçons montre que, toute chose égale par ailleurs, par rapport à un tronçon à zéro intersection, la probabilité d'enregistrer un accident est augmentée de près de 33% pour le tronçon à quatre intersections.

En ce qui concerne le nombre de voies, par rapport à un tronçon à deux voies, la probabilité de subir un accident sur un tronçon à quatre voies est diminuée de près de 25%.

Enfin au sujet des vitesses maximales autorisées : par rapport aux tronçons limités à 50km/h, la probabilité de subir un accident sur un tronçon limité à 100 km/h est diminuée de près de 34%.

### 5.6.5 Conclusion et perspectives

---

Le mandat d'exploration de données d'incidents que nous avons présenté dans ce rapport n'est qu'un aperçu de ce qu'il est possible d'établir à partir de ces données.

L'analyse descriptive spatiale de ces données nous a permis surtout d'établir une cartographie de ces incidents à travers l'identification des points chauds en matière d'accidents et plus globalement d'incidents.

L'analyse descriptive temporelle nous a permis de retrouver le lien de corrélation évoqué dans la littérature entre accident et densité de trafic.

Les analyses de durées nous ont permis là aussi d'approcher de plus près les hypothèses littéraires concernant la congestion et les effets que peuvent avoir les incidents (plus précisément les 5 types majeurs d'incidents) sur les temps de parcours.

La modélisation de survie a permis à travers la caractérisation physique des tronçons d'établir des rapports de risques entre les différents tronçons en fonction donc des caractéristiques liées aux nombre de voies, au sens du tronçon, à la vitesse maximale autorisée, etc.

Maintenant, il faut reconnaître qu'après une exploration de ces données et une tentative de valorisation de ces données il apparaît aujourd'hui qu'il serait judicieux de continuer les recherches possibles sur ces tronçons, d'affiner les analyses déjà exposées.

Il serait aussi possible de tenter une liaison entre le projet d'estimation des temps de parcours et les incidences des accidents et autres types d'incidents sur ces temps de parcours.

Notons aussi que beaucoup d'articles traitent des séries chronologiques et des processus aléatoires pour expliquer la survie des incidents, c'est une piste à explorer.

## 6. Conclusion générale

---

Ce projet de recherche, qui visait initialement à évaluer les enjeux analytiques et les potentialités des données de temps de parcours (recueillies par la méthode des véhicules flottants) pour mesurer la fiabilité du réseau routier montréalais, a finalement permis de procéder à une série d'expérimentations sur la mesure des temps et vitesses sur les réseaux routiers ainsi qu'à une réflexion plus générale sur la modélisation de ces mesures.

### 6.1 Modélisation des temps de parcours

D'abord, un choix méthodologique important a été fait en début de projet, celui de segmenter systématiquement les relevés de véhicules flottants en tronçons d'un kilomètre afin notamment de mieux valoriser les données disponibles; en effet, cette approche a permis de mieux valoriser l'échantillon de données qui s'avérait parfois insuffisant à certains endroits et lors de certaines périodes. La construction d'un ensemble de tronçons routiers de taille uniforme a permis d'exploiter des fonctionnalités d'agrégation par similarité sur la base d'un indicateur extrait des relevés soit le temps de parcours et sa distribution fréquentielle sur chaque tronçon. Ainsi, suite à une validation assistée des données (par le biais d'un outil de consultation interactif), des temps de parcours par kilomètre ont été calculés. Différentes analyses statistiques ont permis de valider la significativité de facteurs tels que l'année, le mois, la période de la journée. De plus, des techniques de classification automatique des tronçons en fonction de la similarité des distributions fréquentielles des temps de parcours ont résulté en la classification des tronçons en 16 groupes typiques (8 par période de la journée : avant-midi et après-midi). Les distributions fréquentielles de temps de parcours de chaque groupe type de tronçons ont été modélisées à l'aide de modèles basés sur trois lois log-normales.

Aussi, deux indicateurs ont été proposés afin de caractériser les conditions de circulation sur les tronçons : un indicateur témoignant du temps de parcours moyen et de la variabilité de ce temps moyen (qui se résume typiquement en quatre quadrants) ainsi qu'un indicateur de dysfonctionnement du réseau qui témoigne de l'occurrence de temps de parcours très élevés sur le tronçon.

### 6.2 Comparaison des technologies de collecte de données

Le projet a aussi donné lieu à différents exercices de comparaison de données issues de différentes sources et technologies. L'équipe de recherche a profité du projet pour examiner les potentialités de différentes technologies pour fournir de l'information de planification sur les vitesses/temps de parcours. En complémentarité de cette étude, des données issues de quelques unes de ces technologies ont été traitées puis comparées. Les sources de données examinées plus spécifiquement sont :

- Données GPS issues de la compagnie Communauto inc. qui fournissent des valeurs de vitesse instantanée en différents points du réseau routier (même au-delà de la région montréalaise);
- Un extrait de données reçu de la compagnie INRIX, couvrant certains segments de grandes autoroutes montréalaises;
- Des données de vitesses moyennes extraites de bandes vidéo par la compagnie RealTraffic;
- Des temps de parcours estimés par la reconnaissance successive d'identifiants uniques par des capteurs Bluetooth (expérience conduite par le MTQ);
- Des vitesses moyennes extraites des bandes vidéo provenant des caméras du MTQ.

En plus de mieux comprendre les fonctionnalités et potentialités de ces différentes sources, l'analyse des banques de données disponibles a permis d'évaluer l'intérêt de différentes technologies et leur capacité à documenter, pour des fins de planification et analyse, les performances du réseau routier.

À la lumière des expérimentations réalisées, l'équipe de recherche propose d'examiner plus sérieusement la possibilité d'établir un partenariat avec Communauto afin de pouvoir utiliser les données systématiquement pour suivre l'évolution des vitesses sur le réseau.

### **6.3 Facteurs physiques**

En vue de mieux comprendre les distributions de temps de parcours observées sur les tronçons routiers, une banque de données sur les facteurs physiques des différents tronçons routiers étudiés a été développée. En effet, l'hypothèse selon laquelle certains attributs physiques sont corrélés aux distributions fréquentielles de temps de parcours méritait d'être examinée plus en profondeur. En plus de permettre la constitution d'une banque de données pertinente sur les facteurs physiques des tronçons, l'analyse a permis de confirmer différentes relations intuitives.

### **6.4 Banque de données d'incidents**

Afin de mieux comprendre l'occurrence de certains temps de parcours très élevés, une banque de données d'incidents a été examinée. Les données brutes, obtenues du MTQ, ont été géolocalisées et étudiées à l'aide de diverses statistiques descriptive. En guise d'exploration, certains modèles ont été testés et permis de confirmer que des travaux plus approfondis sur la modélisation des incidents sont souhaitables et permettraient certainement de raffiner les capacités de modélisation des temps de parcours. Notre équipe de recherche est convaincu qu'un projet de valorisation des banques de données d'incidents devrait être élaboré.

## 7. Références complémentaires

---

- Abdel-Aty, M., Adella, M.F., 2004. Linking roadway geometrics and real time traffic characteristics to model daytime freeway crashes: generalised estimating equations for correlated data. *Transport. Res. Rec* 1897, 106-115
- Abess, C., Jarett, D, C.C., 1981. Accidents at blackspots : estimating the effectiveness of remedial treatment. With special reference to the « regression-to-mean » effect. *Traffic Eng. Control* 22 (10). 535-542
- Kalboussi, A. (2011). Méthode d'analyse et de classification des segments du réseau routier supérieur de la région de Montréal, M.Sc.A., Département de mathématiques et génie industriel, École Polytechnique de Montréal, 87 pages.
- Lord., D., Persaud, B.N ., 2000. Accident prediction models with and without trend : application of the generalized estimating equations procedure. *Transport. Res. Rec* 1717, 102-108
- Miaou, S-P., Lord. D., 2003. Modeling traffic crash-flow relationships for intersections dispersion parameter functional form, and Bayes versus empirical Bayes. *Transport. Res. Rec.* 1840, 31-40
- Miaou, S-P., Song, J.J., 2005. Bayesian ranking of sites for engineering safety improvements: Decision parameter, treatability concept, statistical criterion and spatial dependance. *Accident Anal. Prev.* 37 (4), 699-720.
- Persaud, B.N., Dzbik, L.,1993. Accident prediction models for freeways. *Transport. Res. Rec.* 1401, 55-60
- Stefan, C., 2006. Predictive model injury accidents on Austrian motorways. KfV. Vienna
- Wichert S., Cardoso J., Accident Prediction Models for Portuguese Single Carriageway Roads. LNEC, Lisbon

## 8. Annexes

---

### 8.1 Analyse statistique du circuit 1 au moyen de Statistica

Cette première partie a pour but de donner quelques résultats généraux quant à la distribution des temps de parcours observés sur le circuit 1.

#### 8.1.1 Statistiques descriptives

---

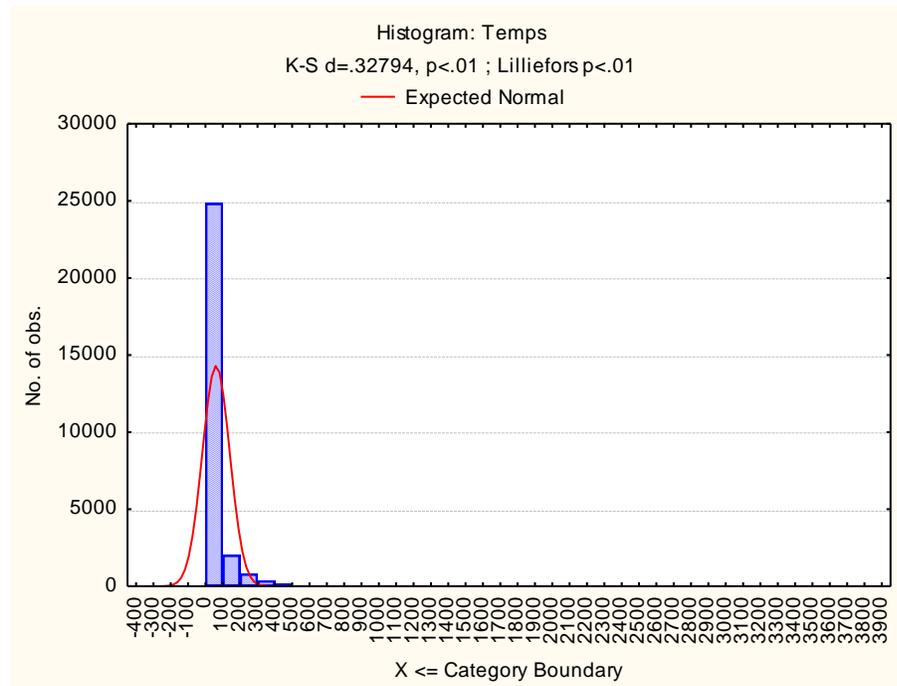
Tableau 42. Statistiques descriptives: temps de parcours circuit 1

Statistiques descriptives							
	Nombre de relevés	Moyenne	Minimum	Maximum	Percentile 10	Percentile 90	Ecart-type
<b>temps de parcours</b>	28296	61.28	4.72	3854.2	30.72	118.68	78.93

On observe un très grand écart entre les valeurs minimum et maximum de cette distribution puisque le maximum vaut 3854.2s, ce qui correspond à presque 1h30 de temps passé sur un seul tronçon de 1 km, soit une vitesse de parcours de 0.93 km/h (cette observation pourrait être acceptée dans le cas où le tronçon sur lequel découle cette expérience est un tronçon sans sortie et ayant subi un incident très important, malheureusement d'autres observations précédentes ou suivantes sur la même journée sur des tronçons identiques semblent prouver l'erreur expérimentale, à moins en dernier lieu que ce soit justement le véhicule flottant qui soit tombé en panne), alors que le minimum correspond à une vitesse aberrante de 762 km/h.

Néanmoins, la moyenne ainsi que les percentiles semblent confirmer l'idée générale d'une plage de temps de parcours raisonnable entre 30 et 120 s, avec une moyenne de 61,27s soit une vitesse de parcours égale à 59km/h. Par conséquent observons l'histogramme des temps de parcours.

### Histogramme des temps de parcours



**Figure 85. Histogramme des temps de parcours**

L’histogramme porte plusieurs conclusions pertinentes : premièrement, les observations semblent toutes être comprises dans un intervalle réduit. Deuxièmement, le test de Kolmogorov-Smirnoff donne une distance de 0.32794 ce qui implique que l’hypothèse de normalité est à rejeter.

### 8.1.2 Tables de fréquence

**Tableau 43. Fréquence des observations par année**

Tableau de fréquence : Année K-S d=.29592, p<.01				
	Nombre	Cumulative	Pourcentage	Cumulative %
<b>1998</b>	3950	3950	13.96	13.96
<b>1999</b>	1817	5767	6.42	20.38
<b>2000</b>	4596	10363	16.24	36.62
<b>2001</b>	2222	12585	7.85	44.48
<b>2002</b>	2595	15180	9.17	53.65
<b>2003</b>	0	15180	0	53.65
<b>2004</b>	13116	28296	46.35	100
<b>Manquant</b>	0	28296	0	

**Tableau 44. Fréquence des observations par mois**

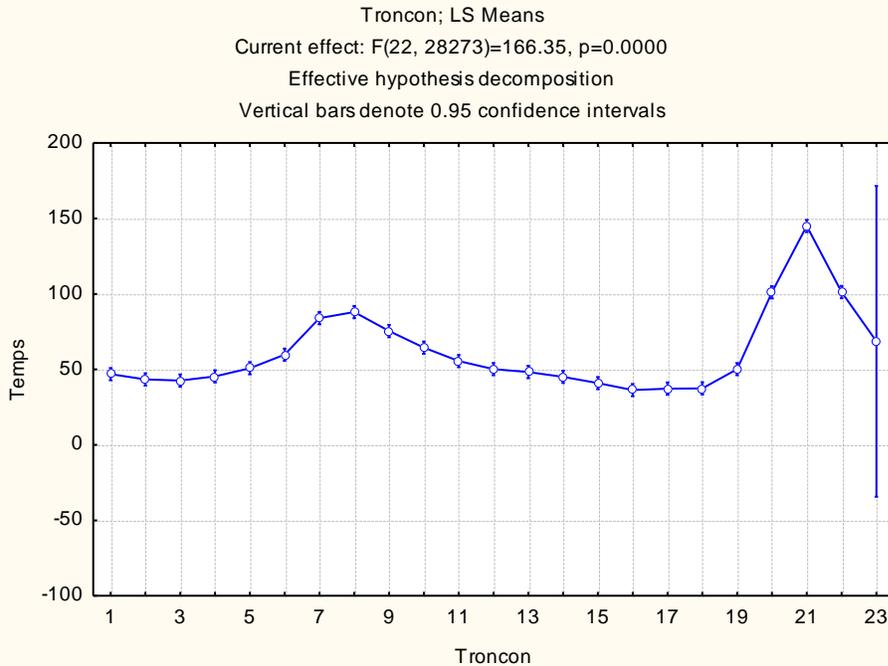
Tableau de fréquence : Mois K-S d=.13950, p<.01				
	<b>Nombre</b>	<b>Cumulative</b>	<b>Pourcentage</b>	<b>Cumulative %</b>
<b>janvier</b>	1248	1248	4.41	4.41
<b>février</b>	3622	4870	12.8	17.21
<b>mars</b>	3359	8229	11.87	29.1
<b>avril</b>	2267	10496	8.01	37.09
<b>mai</b>	1914	12410	6.76	43.86
<b>juin</b>	1868	14278	6.6	50.46
<b>juillet</b>	2285	16563	8.07	58.53
<b>aout</b>	1826	18389	6.45	64.99
<b>septembre</b>	3561	21950	12.58	77.57
<b>octobre</b>	3011	24961	10.64	88.21
<b>novembre</b>	2213	27174	7.82	96.03
<b>décembre</b>	1122	28296	3.96	100
<b>Manquant</b>	0	28296	0	

Ces tableaux dont l'utilité est certes réduite nous informent simplement du fait que sur ce circuit, le nombre d'observations est très largement suffisant pour effectuer les tests statistiques désirés.

### **8.1.3 Moyennes par tronçon**

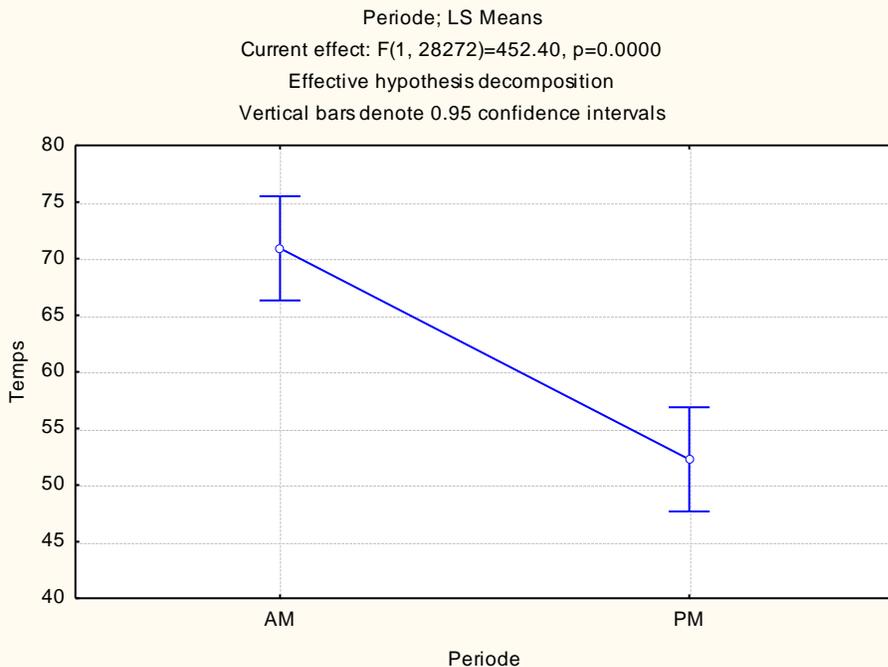
---

Observons désormais les différentes moyennes de tes temps de parcours selon le tronçon (sont rajoutés les intervalles de confiance à 95% de ces moyennes).



**Figure 86. Moyennes des temps de parcours selon le tronçon**

On constate donc une nette augmentation du temps de parcours au niveau du tronçon 7, avec une dispersion au niveau du 10, alors que la fin du tronçon est problématique. On peut d'ores et déjà en donner la raison. Au tronçon 7, appartient un échangeur et la fin du circuit correspond à l'arrivée au niveau de Montréal.



**Figure 87. Moyenne des temps de parcours selon la période**

De plus, il y a une différence significative entre les relevés matinaux et de l'après-midi. En effet, pour la période AM, la moyenne des temps de parcours est de 71s alors qu'elle n'est que de 53s pour l'après-midi.

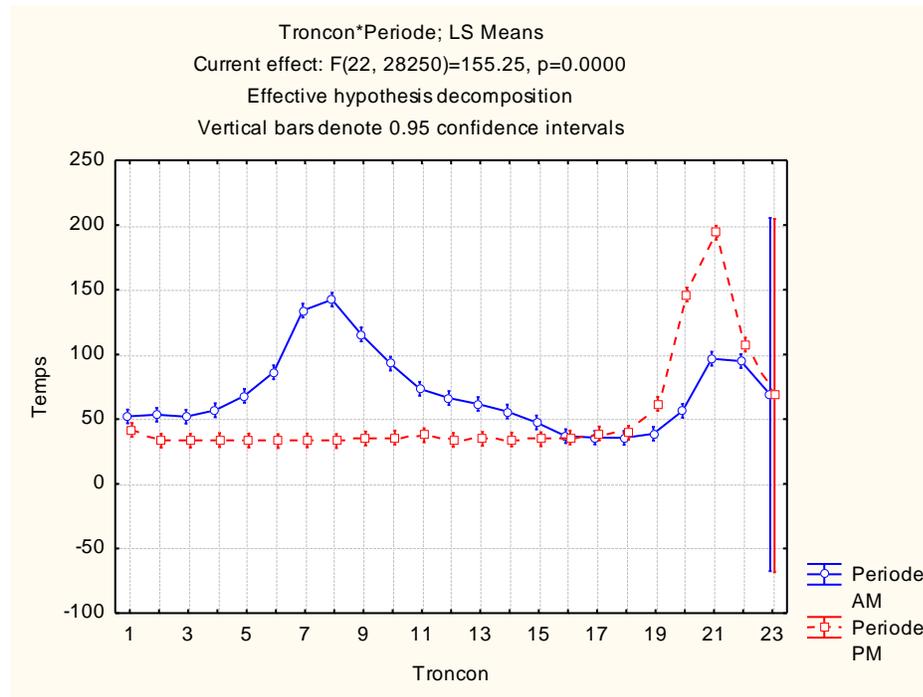


Figure 88. Moyennes des temps de parcours selon le tronçon et la période

L'observation des temps de parcours selon le tronçon et la période, confirme cette constatation. En effet, on peut émettre l'hypothèse que cette différence – en particulier au niveau du tronçon 7 – correspond au départ des personnes vivant en zone urbaine périphérique et travaillant à Montréal.

Tableau 45. ANOVA pour le tronçon et la période plus effet d'interaction

Tableau ANOVA					
	Deg de liberté	Temps SS	Temps MS	Temps F-ratio	Temps p-value
<b>Constante</b>	1	1.06E+08	1.06E+08	21902	0
<b>Tronçon</b>	21	2.03E+07	9.66E+05	199	0
<b>Période</b>	1	2.44E+06	2.44E+06	502	0
<b>Tronçon*Période</b>	21	1.66E+06	7.89E+05	163	0
<b>Erreur</b>	28250	1.37E+08	4851		
<b>Total</b>	28293	1.76E+08			

L'ANOVA confirme la signification des deux facteurs catégoriques tronçon et période ainsi que leur interaction.

### 8.1.4 Analyse de la variance selon l'année

Dans la perspective de dégager les évolutions des temps de parcours, nous choisissons de nous attarder en particulier sur les moyennes des temps de parcours selon l'année, facteur catégorique à 6 niveaux (1998-1999-2000-2001-2002-2004).

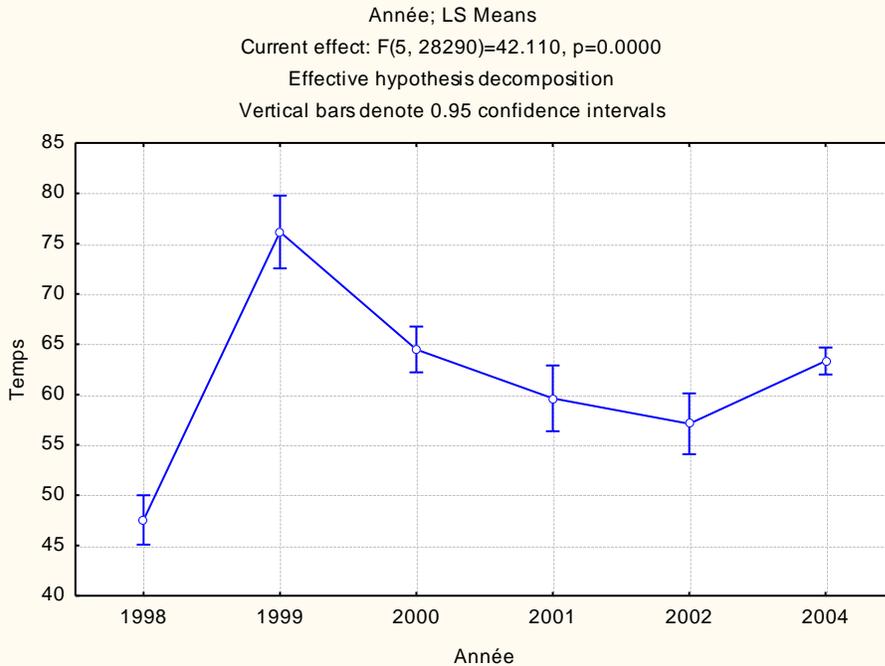
**Tableau 46. ANOVA selon l'année**

Tableau ANOVA					
	Deg de liberté	Temps SS	Temps MS	Temps F-ratio	Temps p-value
<b>Constante</b>	1	7.02E+07	7.02E+07	11342	0
<b>Année</b>	5	1.30E+06	2.60E+05	42	0
<b>Erreur</b>	28290	1.73E+08	6185		
<b>Total</b>	28295	1.76E+08			

Le tableau d'analyse de la variance implique donc que le facteur année est très largement significatif.

**Tableau 47. Coefficients de la régression selon l'année**

Estimation des coefficients						
	Niveau	Colonne	Temps Coeff	Temps Std. Er.	Temps t	Temps p
<b>Constante</b>		1	61.37	0.58	106.5	0
<b>Année</b>	1998	2	-13.83	1.17	-11.79	0
<b>Année</b>	1999	3	14.79	1.613	9.17	0
<b>Année</b>	2000	4	3.11	1.11	2.8	0.005
<b>Année</b>	2001	5	-1.75	1.48	-1.18	0.24
<b>Année</b>	2002	6	-4.28	1.39	-3.09	0.002



**Figure 89. Moyennes des temps de parcours selon l'année**

On constate donc une amélioration de la moyenne des temps de parcours selon l'année. On ne doit toutefois guère s'avancer sur cette conclusion. En effet, tous les mois ne sont pas échantillonnés selon les années.

**Tableau 48. Moyennes et intervalles de confiance selon l'année**

Statistiques Descriptives							
	Niveau	N	Moyenne des relevés	Ecart-type	Erreur Std	Temps +95%	Temps - 95%
Total		28296	61.28	78.93	0.47	60.36	62.2
Année	1998	3950	47.54	40.13	0.64	46.29	48.79
Année	1999	1817	76.16	85.53	2.01	72.23	80.1
Année	2000	4596	64.48	78.35	1.16	62.22	66.75
Année	2001	2222	59.62	65.44	1.39	56.9	62.35
Année	2002	2595	57.09	63.38	1.24	54.66	59.53
Année	2004	13116	63.33	90.41	0.79	61.79	64.88

Tableau 49. Calcul des coefficients de signification

Coefficient de signification			
	R	R <sup>2</sup>	R <sup>2</sup> ajusté
<b>Temps</b>	0.09	0.01	0.01

On trouve de plus un R ajusté très faible.

Tableau 50. Test de LSD selon l'année

Test LSD selon l'année (différence des moyennes entre années)							
	Année	{1}	{2}	{3}	{4}	{5}	{6}
1	1998		0.00	0.00	0.00	0.00	0.00
2	1999	0.00		0.00	0.00	0.00	0.00
3	2000	0.00	0.00		0.02	0.00	0.39
4	2001	0.00	0.00	0.02		0.27	0.04
5	2002	0.00	0.00	0.00	0.27		0.00
6	2004	0.00	0.00	0.39	0.04	0.00	

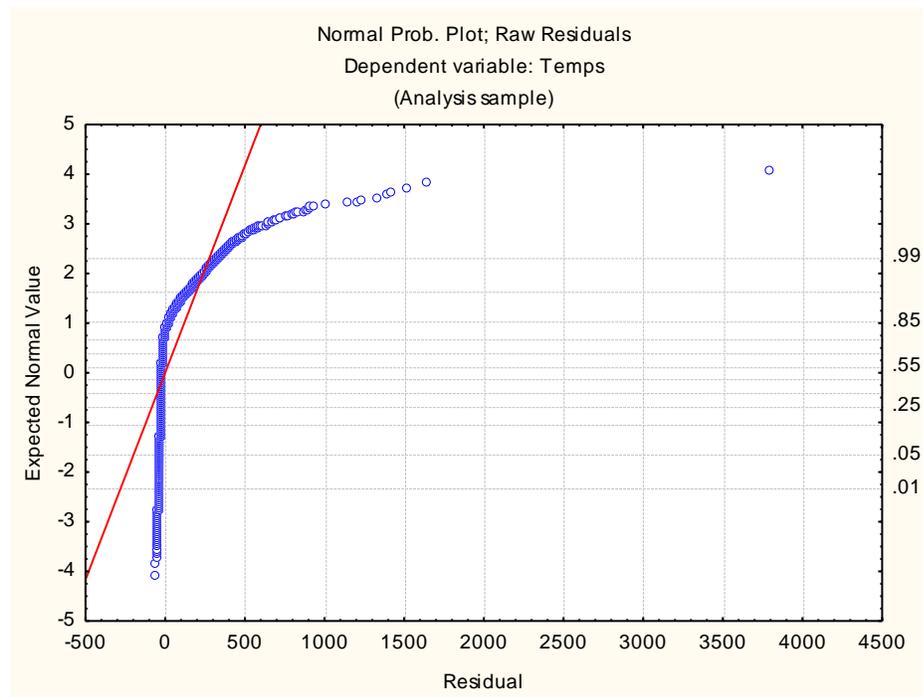


Figure 90. Résidus sur échelle normale

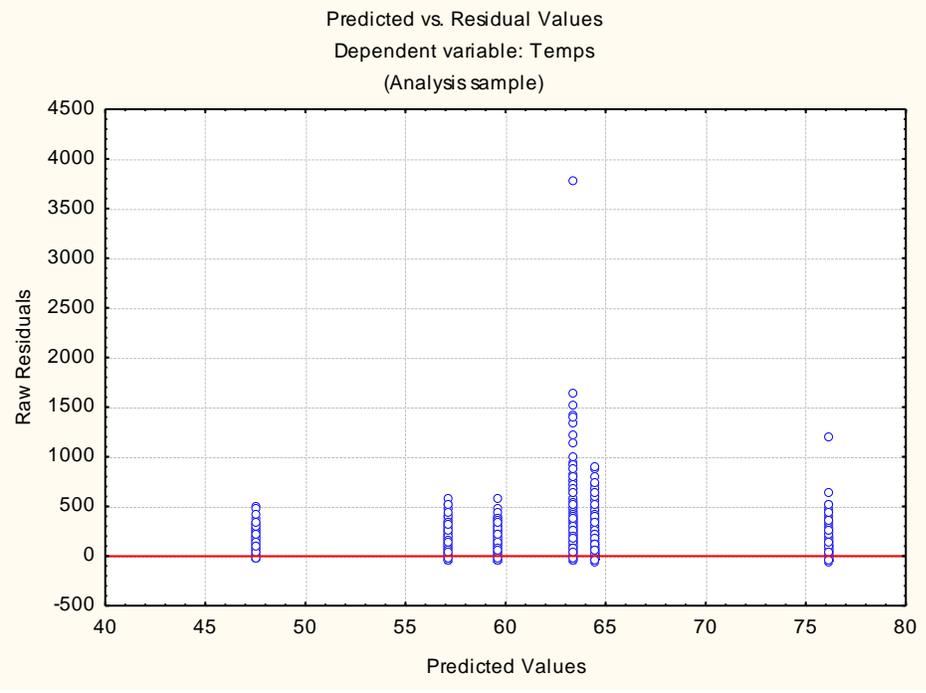


Figure 91. Valeurs prédites vs résidus

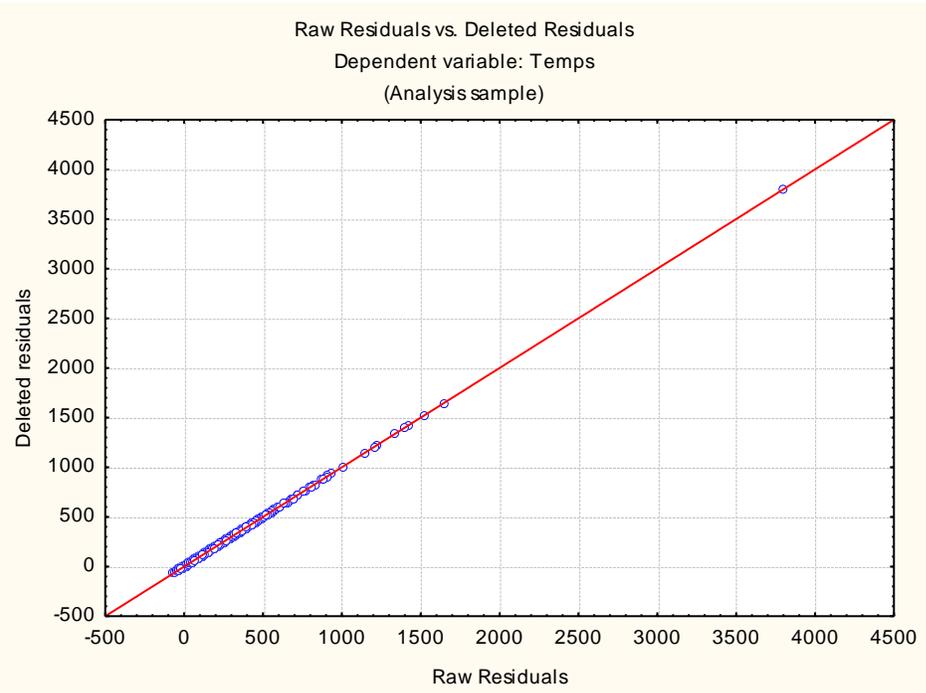


Figure 92. Résidus vs Résidus exclus

**Tableau 51. Coefficients de corrélation**

Tableau de Corrélations des relevés par année										
	Niveau	Colonne	Effet	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7
<b>Constantes</b>		1	Fixe							
<b>Année</b>	1998	2	Fixe		1.00	0.78	0.71	0.76	0.75	-0.05
<b>Année</b>	1999	3	Fixe		0.78	1.00	0.77	0.81	0.80	0.00
<b>Année</b>	2000	4	Fixe		0.71	0.77	1.00	0.76	0.74	-0.01
<b>Année</b>	2001	5	Fixe		0.76	0.81	0.76	1.00	0.79	-0.02
<b>Année</b>	2002	6	Fixe		0.75	0.80	0.74	0.79	1.00	-0.03
<b>Total</b>		7			-0.05	0.00	-0.01	-0.02	-0.03	1.00

Le tableau des corrélations montrent que les observations annuelles sont relativement bien corrélées ce qui pourrait mettre en doute le choix de ce modèle.

Dans un deuxième temps nous effectuons une ANOVA selon l'année et la période.

**Tableau 52. ANOVA selon la période et l'année avec effet d'interaction**

Tableau ANOVA					
	SS	Degré de liberté	MS	F	p
<b>Constante</b>	70085431	1	70085431	11533.57	0
<b>Période</b>	1325251	1	1325251	218.09	0
<b>Année</b>	1307821	5	261564	43.04	0
<b>Période*Année</b>	676240	5	135248	22.26	0
<b>Erreur</b>	171859737	28282	6077		

Les résultats sont donc que ces deux facteurs sont significatifs, de plus la courbe représentant les moyennes par année selon la période semble confirmer l'hypothèse d'une amélioration de la qualité de la circulation au fil des années.

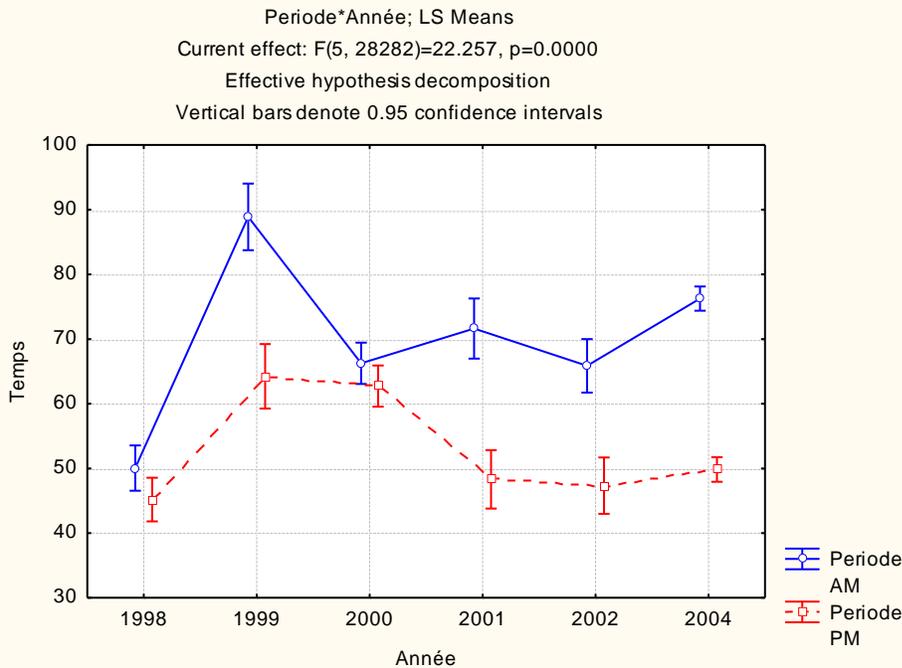


Figure 93. Moyennes selon année et période

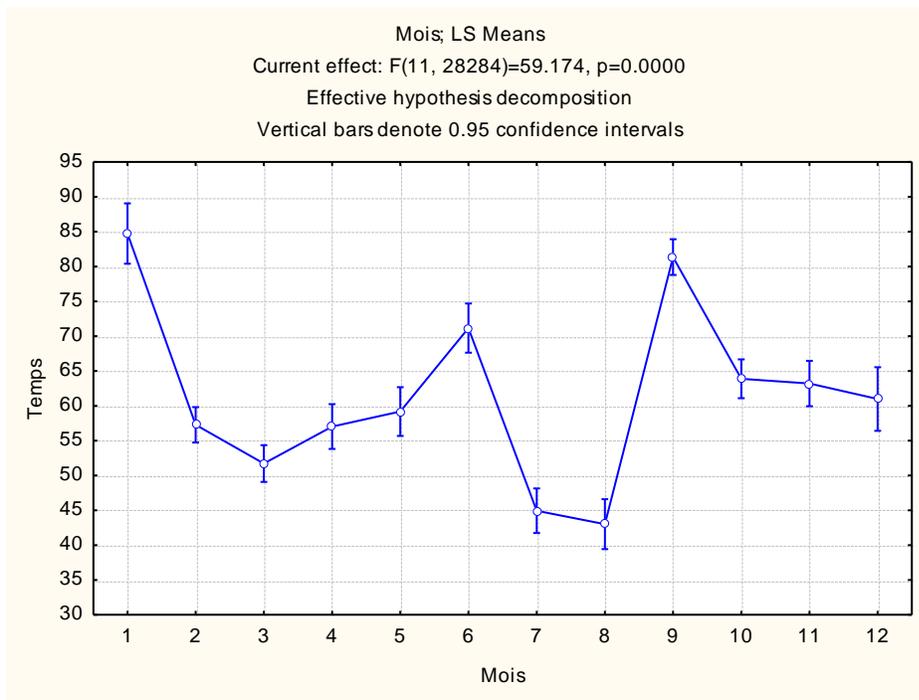
### 8.1.5 Analyse de la variance selon le mois

Désormais dans le but d'observer les différences significatives entre les mois, commentons les résultats obtenus après une analyse de la variance des temps de parcours selon le mois d'observations

Tableau 53. ANOVA selon le mois

Tableau ANOVA					
	SS	Deg de liberté	MS	F	p
<b>Constante</b>	9.38E+07	1	9.38E+07	15390.02	0
<b>Mois</b>	3.97E+06	11	3.61E+05	59.17	0
<b>Erreur</b>	1.72E+08	28284	6.09E+03		

L'ANOVA donne directement le mois comme un facteur significatif.



**Figure 94. Moyennes des temps de parcours selon le mois**

**Tableau 54. Calcul du coefficient de signification**

Coefficient de signification			
	R	R <sup>2</sup>	R <sup>2</sup> ajusté
<b>Temps</b>	0.15	0.02	0.02

Le R ajusté demeure faible, par contre nous pouvons voir une courbe de tendance intéressante. Les mois de juillet et aout possèdent de très largement inférieures moyennes de temps de parcours en raison surement des vacances estivales, alors que justement le mois de juin et de septembre, correspondraient aux départs et à la rentrée.

Tableau 55. Coefficients et intervalle de confiance par mois

Estimation des coefficients								
	Niveau	Colonne	Temps Coeff	Temps Err. Std.	Temps t	Temps p	-95.00%	95.00%
<b>Constante</b>		1	61.54	0.50	124.06	0.00	60.56	62.51
<b>Mois</b>	1	2	23.19	2.08	11.17	0.00	19.12	27.26
<b>Mois</b>	2	3	-4.26	1.28	-3.32	0.00	-6.78	-1.75
<b>Mois</b>	3	4	-9.82	1.33	-7.40	0.00	-12.41	-7.22
<b>Mois</b>	4	5	-4.51	1.58	-2.86	0.00	-7.61	-1.42
<b>Mois</b>	5	6	-2.36	1.70	-1.38	0.17	-5.69	0.98
<b>Mois</b>	6	7	9.62	1.72	5.59	0.00	6.25	13.00
<b>Mois</b>	7	8	-16.60	1.57	-10.57	0.00	-19.68	-13.52
<b>Mois</b>	8	9	-18.52	1.74	-10.65	0.00	-21.93	-15.11
<b>Mois</b>	9	10	19.81	1.29	15.32	0.00	17.28	22.35
<b>Mois</b>	10	11	2.34	1.39	1.69	0.09	-0.38	5.07
<b>Mois</b>	11	12	1.66	1.59	1.04	0.30	-1.46	4.79

Tableau 56. Test LSD sur les moyennes par mois

Test LSD sur les moyennes par mois													
	Mois	{1}	{2}	{3}	{4}	{5}	{6}	{7}	{8}	{9}	{10}	{11}	{12}
<b>1</b>	1		0	0	0	0	0.000002	0	0	0.188013	0	0	0
<b>2</b>	2	0.00		0.00	0.90	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.16
<b>3</b>	3	0.00	0.00		0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>4</b>	4	0.00	0.90	0.01		0.37	0.00	0.00	0.00	0.00	0.00	0.01	0.17
<b>5</b>	5	0.00	0.39	0.00	0.37		0.00	0.00	0.00	0.00	0.04	0.10	0.54
<b>6</b>	6	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00
<b>7</b>	7	0.00	0.00	0.00	0.00	0.00	0.00		0.43	0.00	0.00	0.00	0.00
<b>8</b>	8	0.00	0.00	0.00	0.00	0.00	0.00	0.43		0.00	0.00	0.00	0.00
<b>9</b>	9	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00
<b>10</b>	10	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00		0.76	0.29
<b>11</b>	11	0.00	0.00	0.00	0.01	0.10	0.00	0.00	0.00	0.00	0.76		0.44
<b>12</b>	12	0.00	0.16	0.00	0.17	0.54	0.00	0.00	0.00	0.00	0.29	0.44	

Les tests de différences entre les moyennes des temps de parcours par mois sont relativement faibles, hormis quelques cas. On peut citer par exemple le mois d'avril et le mois de février (90%), mai et février (38%), avril et mai (37%), décembre et mai (54%), aout et juillet (43%). Peut-on alors en dégager une forme de conclusion ? Il semble que l'on confirme l'idée faite sur les mois de juillet et aout. Par contre, il semblerait inopportun de dégager une conclusion sur les autres valeurs car ceci dépend peut-être des conditions météorologiques rencontrées durant ces périodes.

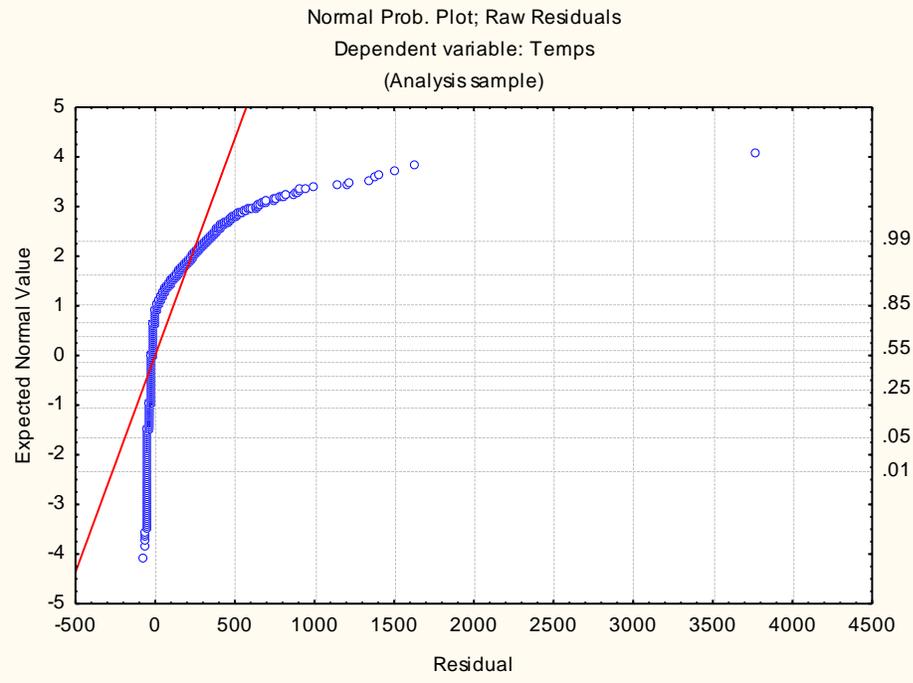


Figure 95. Résidus sur échelle normale

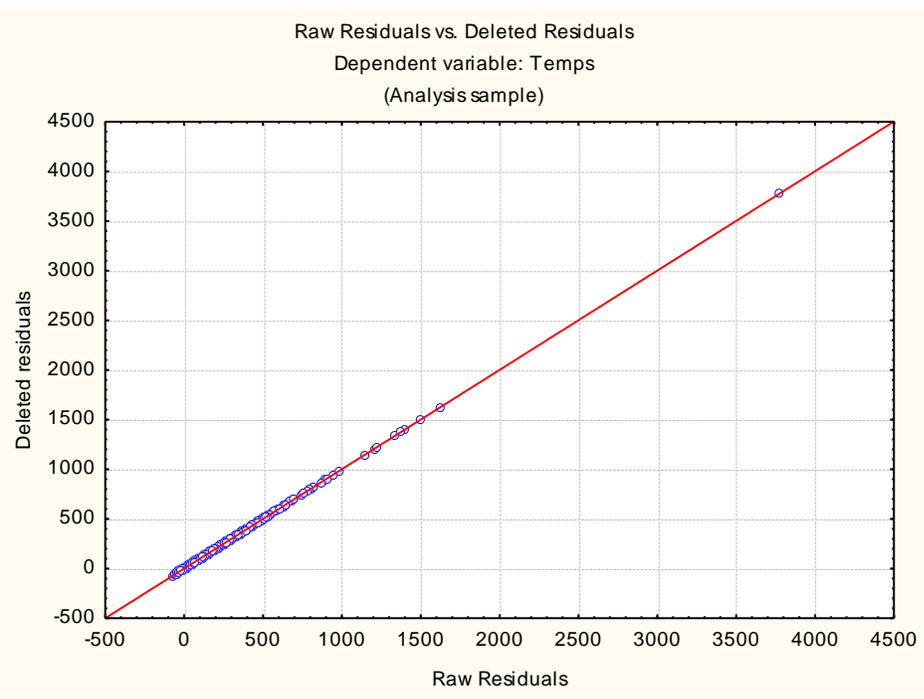


Figure 96. Résidus vs résidus exclus

Tableau 57. Coefficients de corrélation

Tableau de corrélation des relevés par mois																
	Niveau	Colonne	Effet	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7	Col. 8	Col. 9	Col. 10	Col. 11	Col. 12	Col. 13
Constante		1	Fixe													
Mois	1	2	Fixe		1.00	0.34	0.35	0.40	0.42	0.42	0.40	0.42	0.34	0.36	0.40	0.05
Mois	2	3	Fixe		0.34	1.00	0.21	0.26	0.28	0.29	0.26	0.29	0.20	0.22	0.27	-
Mois	3	4	Fixe		0.35	0.21	1.00	0.27	0.29	0.30	0.27	0.30	0.21	0.23	0.28	-
Mois	4	5	Fixe		0.40	0.26	0.27	1.00	0.34	0.35	0.32	0.35	0.26	0.29	0.32	-
Mois	5	6	Fixe		0.42	0.28	0.29	0.34	1.00	0.37	0.34	0.37	0.29	0.31	0.35	-
Mois	6	7	Fixe		0.42	0.29	0.30	0.35	0.37	1.00	0.35	0.37	0.29	0.31	0.35	-
Mois	7	8	Fixe		0.40	0.26	0.27	0.32	0.34	0.35	1.00	0.35	0.26	0.28	0.32	-
Mois	8	9	Fixe		0.42	0.29	0.30	0.35	0.37	0.37	0.35	1.00	0.29	0.31	0.35	-
Mois	9	10	Fixe		0.34	0.20	0.21	0.26	0.29	0.29	0.26	0.29	1.00	0.23	0.27	-
Mois	10	11	Fixe		0.36	0.22	0.23	0.29	0.31	0.31	0.28	0.31	0.23	1.00	0.29	-
Mois	11	12	Fixe		0.40	0.27	0.28	0.32	0.35	0.35	0.32	0.35	0.27	0.29	1.00	-
Total		13			0.05	0.02	0.04	0.01	0.01	0.03	0.05	0.05	0.08	0.01	0.01	1.00

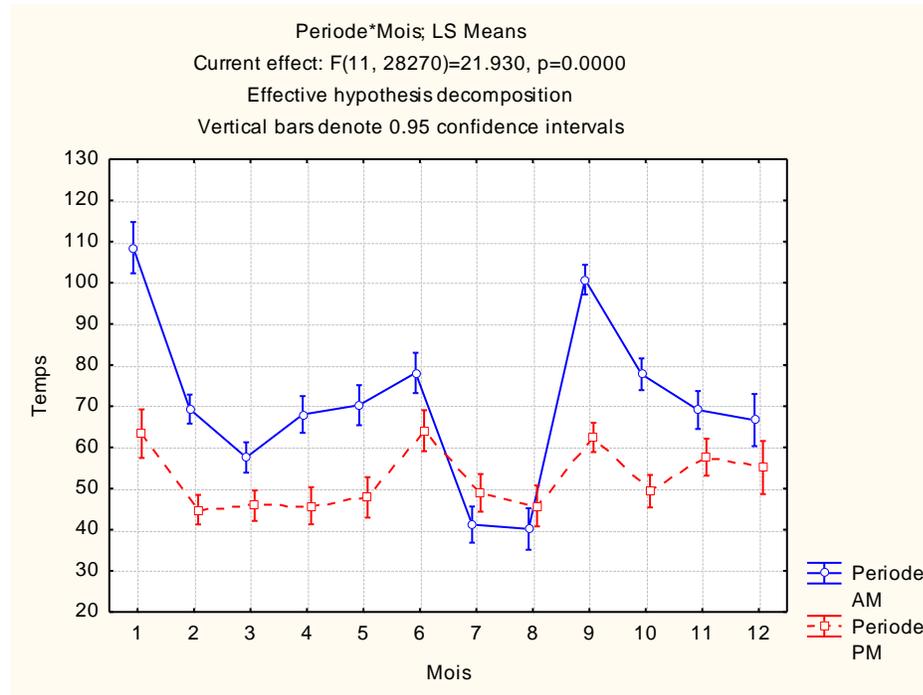


Figure 97. Moyennes des temps de parcours selon le mois et la période

L'observation des temps de parcours selon le mois et la période expose les mêmes conclusions que celles citées précédemment avec exactement les tendances identiques pour les deux périodes que la tendance générale.

### 8.1.6 Analyse de la variance selon les conditions météorologiques

Nous effectuons désormais une analyse de la variance des temps de parcours, avec comme facteur catégorique les conditions météorologiques rencontrées sur le circuit.

Tableau 58. ANOVA selon les conditions météo

Tableau ANOVA					
	SS	Degré de liberté	MS	F	p
Constante	5.74E+06	1	5.74E+06	930.4558	0
Cond	1.79E+06	5	3.58E+05	57.9769	0
Erreur	1.74E+08	28290	6168		

De nouveau, l'analyse donne ce facteur significatif. Le graphe suivant confirme l'idée entendue, d'une augmentation des temps de parcours avec la dégradation des conditions de circulation.

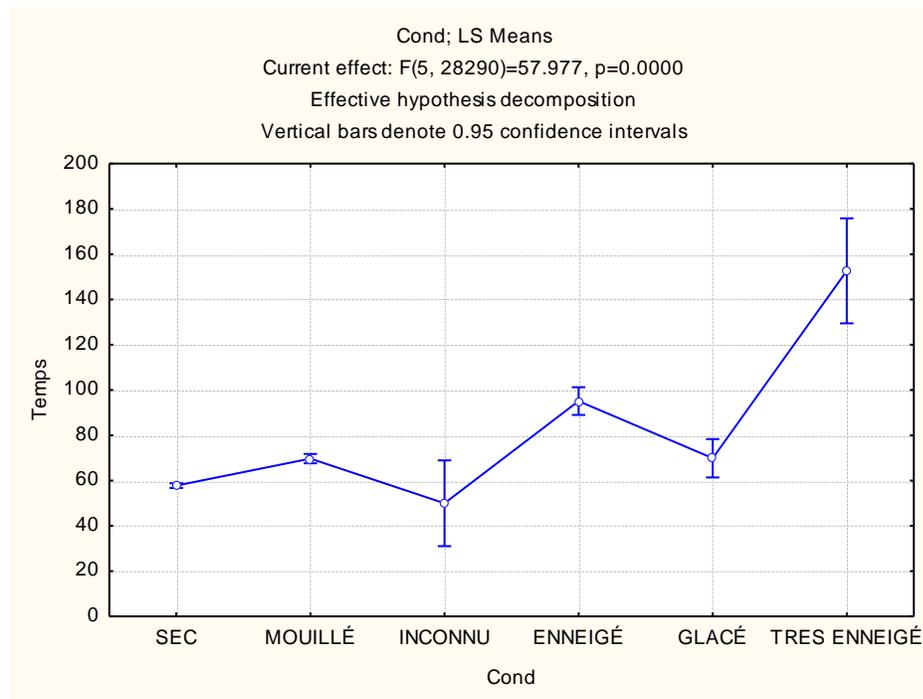


Figure 98. Moyennes des temps de parcours selon les conditions météo

Tableau 59. Coefficients de la régression et intervalles de confiance

Estimation des coefficients									
	Niveau	Colonne	Temps Coeff	Temps Std.	Err. t	Temps p	Temps p	- 95.00%	95.00%
<b>Constante</b>		1	82.52	2.71	30.50	0.00		77.22	87.82
<b>Cond</b>	SEC	2	-24.73	2.74	-9.02	0.00		-30.10	-19.36
<b>Cond</b>	MOUILLÉ	3	-12.80	2.84	-4.51	0.00		-18.36	-7.24
<b>Cond</b>	INCONNU	4	-32.56	8.34	-3.90	0.00		-48.91	-16.20
<b>Cond</b>	ENNEIGÉ	5	12.60	3.72	3.39	0.00		5.31	19.88
<b>Cond</b>	GLACÉ	6	-12.65	4.45	-2.84	0.00		-21.37	-3.93

Tableau 60. Calcul des coefficients de signification

Coefficient de signification			
	R	R <sup>2</sup>	R <sup>2</sup> ajusté
<b>Temps</b>	0.10	0.01	0.01

Tableau 61. Test LSD

Test LSD sur les moyennes par conditions météo							
	Cond	{1}	{2}	{3}	{4}	{5}	{6}
<b>1</b>	SEC		0.00	0.42	0.00	0.01	0.00
<b>2</b>	MOUILLÉ	0.00		0.04	0.00	0.97	0.00
<b>3</b>	INCONNU	0.42	0.04		0.00	0.06	0.00
<b>4</b>	ENNEIGÉ	0.00	0.00	0.00		0.00	0.00
<b>5</b>	GLACÉ	0.01	0.97	0.06	0.00		0.00
<b>6</b>	TRES ENNEIGÉ	0.00	0.00	0.00	0.00	0.00	

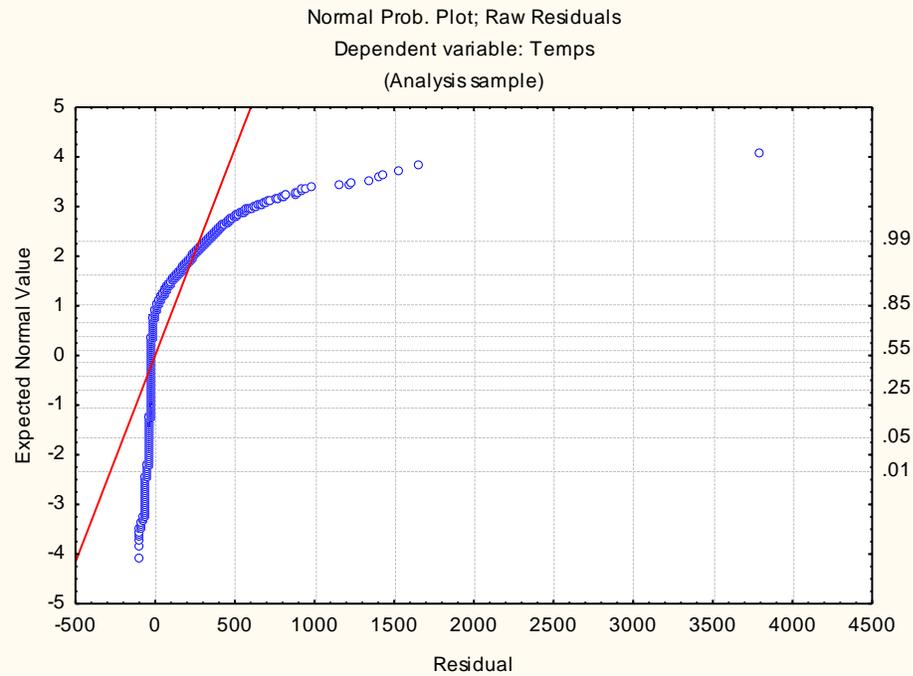


Figure 99. Résidus sur échelle normale

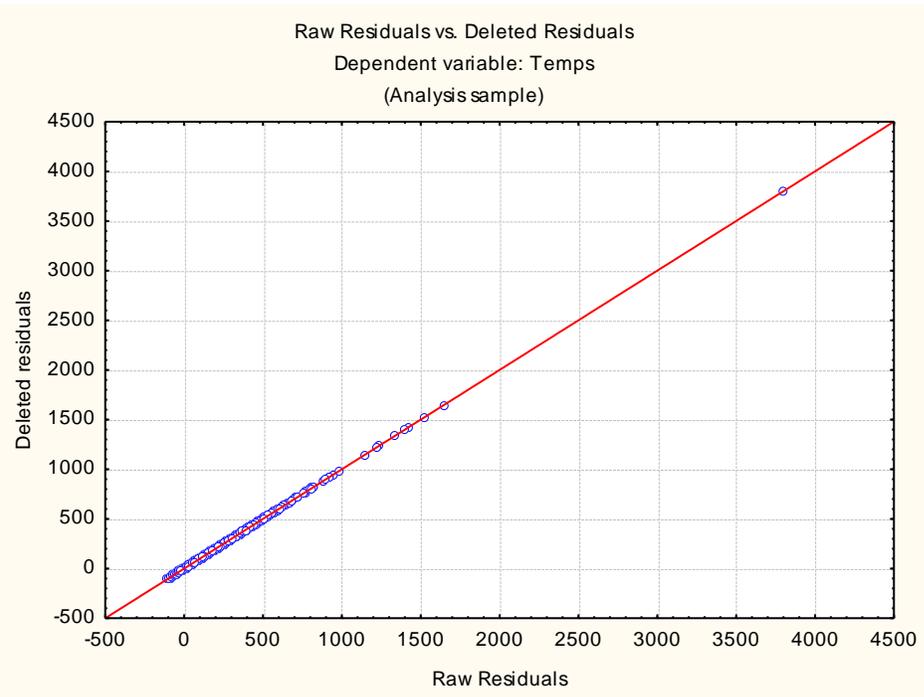


Figure 100. Résidus vs résidus exclus

**Tableau 62. Coefficients de corrélation**

Tableau de corrélations des relevés par conditions météo										
	Level	Column	Effect	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6	Col. 7
<b>Constante</b>		1	Fixed							
<b>Cond</b>	SEC	2	Fixed		1.00	-0.86	0.04	-0.22	-0.12	-0.08
<b>Cond</b>	MOUILLÉ	3	Fixed		-0.86	1.00	0.06	-0.04	-0.01	0.05
<b>Cond</b>	INCONNU	4	Fixed		0.04	0.06	1.00	0.16	0.22	-0.03
<b>Cond</b>	ENNEIGÉ	5	Fixed		-0.22	-0.04	0.16	1.00	0.08	0.05
<b>Cond</b>	GLACÉ	6	Fixed		-0.12	-0.01	0.22	0.08	1.00	0.00
<b>Total</b>		7			-0.08	0.05	-0.03	0.05	0.00	1.00

Le tableau des corrélations donnent une très forte corrélation entre les observations des temps de parcours pour temps sec et pluvieux.

### 8.1.7 Temps de parcours avec mois et année

Cette étude est effectuée dans le but de remarquer les tendances et de pouvoir évaluer la tendance observée lors de l'ANOVA des temps de parcours selon l'année.

**Tableau 63. Moyennes selon Mois par Année**

Moyennes selon mois de l'année							
	Année	Mois	Temps Mean	Temps Std. Err.	Temps -95%	Temps +95%	N
1	1998	2	49.31	4.20	41.09	57.54	342
2	1998	3	40.50	3.90	32.85	48.14	396
3	1998	4	58.31	4.99	48.53	68.08	242
4	1998	5	40.33	3.90	32.68	47.97	396
5	1998	6	54.14	4.01	46.27	62.00	374
6	1998	7	41.40	3.90	33.76	49.04	396
7	1998	8	39.41	3.90	31.76	47.05	396
8	1998	9	62.28	4.27	53.90	70.65	330
9	1998	11	47.20	2.93	41.46	52.93	704
10	1998	12	50.19	4.01	42.32	58.06	374
11	1999	1	103.18	4.59	94.18	112.17	286
12	1999	9	78.18	4.01	70.32	86.05	374
13	1999	10	66.32	3.90	58.68	73.97	396
14	1999	11	70.13	2.81	64.62	75.65	761
15	2000	1	65.96	4.01	58.10	73.83	374
16	2000	2	61.54	3.90	53.89	69.18	396
17	2000	3	50.03	3.90	42.39	57.67	396
18	2000	4	54.79	3.90	47.15	62.44	396
19	2000	5	79.07	3.90	71.43	86.72	396
20	2000	6	73.40	4.02	65.51	81.28	372

21	2000	7	54.69	3.90	47.05	62.34	396
22	2000	8	50.39	4.14	42.28	58.50	352
23	2000	9	74.14	3.90	66.50	81.79	396
24	2000	10	49.23	3.90	41.58	56.87	396
25	2000	11	86.32	4.01	78.46	94.19	374
26	2000	12	75.76	4.14	67.65	83.86	352
27	2001	7	48.34	3.90	40.69	55.98	396
28	2001	8	40.72	4.01	32.86	48.59	374
29	2001	9	87.79	3.90	80.14	95.43	396
30	2001	10	67.78	4.59	58.79	76.78	286
31	2001	11	56.10	4.01	48.23	63.96	374
32	2001	12	58.02	3.90	50.38	65.67	396
33	2002	1	67.36	4.01	59.49	75.22	374
34	2002	2	59.96	4.01	52.10	67.83	374
35	2002	3	45.47	3.90	37.83	53.12	396
36	2002	4	59.50	3.02	53.58	65.42	660
37	2002	6	71.76	3.90	64.12	79.41	396
38	2002	7	37.43	3.91	29.76	45.11	393
39	2004	1	123.22	5.31	112.83	133.62	214
40	2004	2	57.28	1.55	54.24	60.31	2509
41	2004	3	55.21	1.67	51.95	58.48	2171
42	2004	4	55.91	2.49	51.02	60.80	968
43	2004	5	58.81	2.32	54.27	63.35	1122
44	2004	6	78.45	2.88	72.80	84.09	726
45	2004	7	43.70	2.93	37.97	49.44	704
46	2004	8	42.57	2.93	36.83	48.30	704
47	2004	9	85.11	1.71	81.77	88.46	2065
48	2004	10	65.80	1.77	62.34	69.26	1933

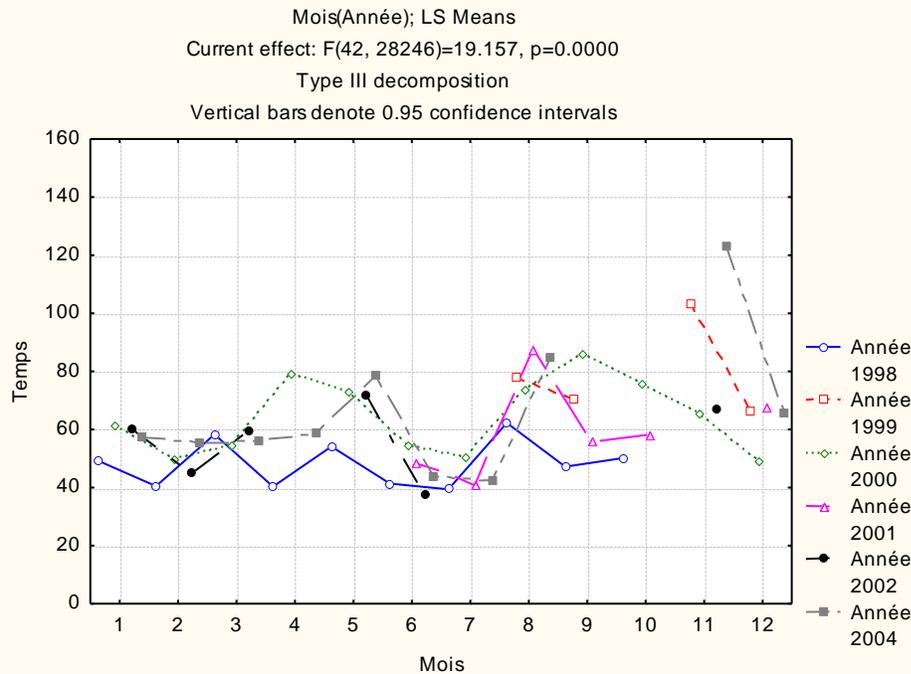


Figure 101. Moyennes selon mois par année

On constate donc que les tendances sont relativement identiques selon les années. Les différences que nous avons pu alors constater auparavant entre les années semblent implicitement liées aux mois d’observations selon l’année. En effet, en 1999, seulement les mois d’août, septembre, octobre, novembre et décembre ont été étudiés.

Tableau 64. ANOVA Nested design

Tableau ANOVA					
	Degré de libertés	SS	MS	F-Ratio	P-value
<b>Constante</b>	1	6.96E+07	6.96E+07	11548.77	0
<b>Annee</b>	5	1.46E+06	2.91E+05	48.35	0
<b>Mois(Annee)</b>	42	4.85E+06	1.15E+05	19.15	0
<b>Erreur</b>	28244	1.70E+08	6.02E+03		
<b>Total</b>	28291	1.76E+08			

On constate donc que les deux facteurs sont significatifs. Cherchons alors à différencier les mois selon la période d’observation.

Tableau 65. Moyenne et intervalle de confiance mois(année\*période)

Moyennes selon mois de l'année et période								
	Période	Année	Mois	Temps	Temps	Temps	Temps	N
1	PM	1998	2	43.06	6.15	31.01	55.11	154
2	PM	1998	3	44.05	5.42	33.42	54.68	198
3	PM	1998	4	56.31	5.75	45.04	67.58	176
4	PM	1998	5	37.71	5.42	27.09	48.34	198
5	PM	1998	6	58.87	5.75	47.60	70.15	176
6	PM	1998	7	38.45	5.42	27.82	49.07	198
7	PM	1998	8	41.16	5.42	30.53	51.78	198
8	PM	1998	9	55.94	5.42	45.32	66.57	198
9	PM	1998	11	40.03	4.07	32.06	48.00	352
10	PM	1998	12	42.64	5.42	32.01	53.27	198
11	PM	1999	1	67.18	5.75	55.90	78.45	176
12	PM	1999	9	66.85	5.44	56.20	77.51	197
13	PM	1999	10	50.57	5.42	39.95	61.20	198
14	PM	1999	11	68.96	3.99	61.14	76.77	366
15	PM	2000	1	52.90	5.42	42.27	63.53	198
16	PM	2000	2	42.75	5.42	32.12	53.38	198
17	PM	2000	3	58.35	5.42	47.72	68.97	198
18	PM	2000	4	40.32	5.42	29.70	50.95	198
19	PM	2000	5	53.63	5.42	43.00	64.25	198
20	PM	2000	6	94.12	5.78	82.78	105.46	174
21	PM	2000	7	67.08	5.42	56.45	77.71	198
22	PM	2000	8	65.14	5.75	53.86	76.41	176
23	PM	2000	9	69.41	5.42	58.78	80.04	198
24	PM	2000	10	50.37	5.44	39.71	61.02	197
25	PM	2000	11	78.01	5.14	67.92	88.09	220
26	PM	2000	12	89.13	6.15	77.08	101.18	154
27	PM	2001	7	59.06	5.42	48.44	69.69	198
28	PM	2001	8	39.55	5.75	28.27	50.82	176
29	PM	2001	9	59.45	5.42	48.82	70.07	198
30	PM	2001	10	44.00	5.75	32.73	55.28	176
31	PM	2001	11	45.28	5.42	34.66	55.91	198
32	PM	2001	12	41.09	5.42	30.47	51.72	198
33	PM	2002	1	63.11	5.75	51.84	74.38	176
34	PM	2002	2	52.29	5.75	41.01	63.56	176
35	PM	2002	3	40.85	5.42	30.22	51.48	198
36	PM	2002	4	41.38	4.51	32.53	50.22	286
37	PM	2002	6	54.95	5.42	44.32	65.58	198
38	PM	2002	7	36.28	5.46	25.57	46.99	195

39	PM	2004	1	76.62	7.38	62.16	91.07	107
40	PM	2004	2	44.38	2.15	40.16	48.60	1255
41	PM	2004	3	44.78	2.32	40.23	49.34	1078
42	PM	2004	4	46.94	3.55	39.99	53.90	462
43	PM	2004	5	49.41	3.25	43.03	55.79	550
44	PM	2004	6	57.30	3.95	49.57	65.03	374
45	PM	2004	7	45.53	4.35	37.01	54.05	308
46	PM	2004	8	42.05	3.95	34.31	49.78	374
47	PM	2004	9	62.09	2.40	57.38	66.79	1010
48	PM	2004	10	49.96	2.54	44.97	54.95	899
49	AM	1998	2	54.43	5.56	43.53	65.34	188
50	AM	1998	3	36.94	5.42	26.31	47.57	198
51	AM	1998	4	63.63	9.39	45.22	82.04	66
52	AM	1998	5	42.94	5.42	32.31	53.57	198
53	AM	1998	6	49.93	5.42	39.30	60.56	198
54	AM	1998	7	44.35	5.42	33.73	54.98	198
55	AM	1998	8	37.66	5.42	27.03	48.29	198
56	AM	1998	9	71.77	6.64	58.76	84.79	132
57	AM	1998	11	54.36	4.07	46.39	62.33	352
58	AM	1998	12	58.68	5.75	47.41	69.96	176
59	AM	1999	1	160.78	7.27	146.52	175.04	110
60	AM	1999	9	91.26	5.75	79.99	102.54	176
61	AM	1999	10	82.08	5.42	71.45	92.70	198
62	AM	1999	11	71.22	3.84	63.70	78.75	395
63	AM	2000	1	80.66	5.75	69.39	91.93	176
64	AM	2000	2	80.32	5.42	69.69	90.95	198
65	AM	2000	3	41.71	5.42	31.09	52.34	198
66	AM	2000	4	69.26	5.42	58.64	79.89	198
67	AM	2000	5	104.52	5.42	93.89	115.15	198
68	AM	2000	6	55.18	5.42	44.55	65.81	198
69	AM	2000	7	42.31	5.42	31.68	52.94	198
70	AM	2000	8	35.64	5.75	24.37	46.92	176
71	AM	2000	9	78.88	5.42	68.25	89.51	198
72	AM	2000	10	48.32	5.42	37.69	58.95	198
73	AM	2000	11	98.21	6.15	86.16	110.26	154
74	AM	2000	12	65.35	5.42	54.73	75.98	198
75	AM	2001	7	37.61	5.42	26.99	48.24	198
76	AM	2001	8	41.77	5.42	31.15	52.40	198
77	AM	2001	9	116.13	5.42	105.50	126.76	198
78	AM	2001	10	105.83	7.27	91.57	120.09	110
79	AM	2001	11	68.26	5.75	56.99	79.53	176

80	AM	2001	12	74.96	5.42	64.33	85.58	198
81	AM	2002	1	71.13	5.42	60.50	81.76	198
82	AM	2002	2	66.79	5.42	56.16	77.41	198
83	AM	2002	3	50.10	5.42	39.47	60.72	198
84	AM	2002	4	73.36	3.95	65.63	81.09	374
85	AM	2002	6	88.57	5.42	77.95	99.20	198
86	AM	2002	7	38.57	5.42	27.94	49.19	198
87	AM	2004	1	169.83	7.38	155.37	184.29	107
88	AM	2004	2	70.18	2.15	65.96	74.41	1254
89	AM	2004	3	65.50	2.31	60.98	70.03	1093
90	AM	2004	4	64.09	3.39	57.44	70.74	506
91	AM	2004	5	67.85	3.19	61.60	74.11	572
92	AM	2004	6	100.92	4.07	92.95	108.89	352
93	AM	2004	7	42.28	3.83	34.76	49.79	396
94	AM	2004	8	43.15	4.20	34.92	51.39	330
95	AM	2004	9	107.16	2.35	102.56	111.76	1055
96	AM	2004	10	79.58	2.37	74.93	84.23	1034

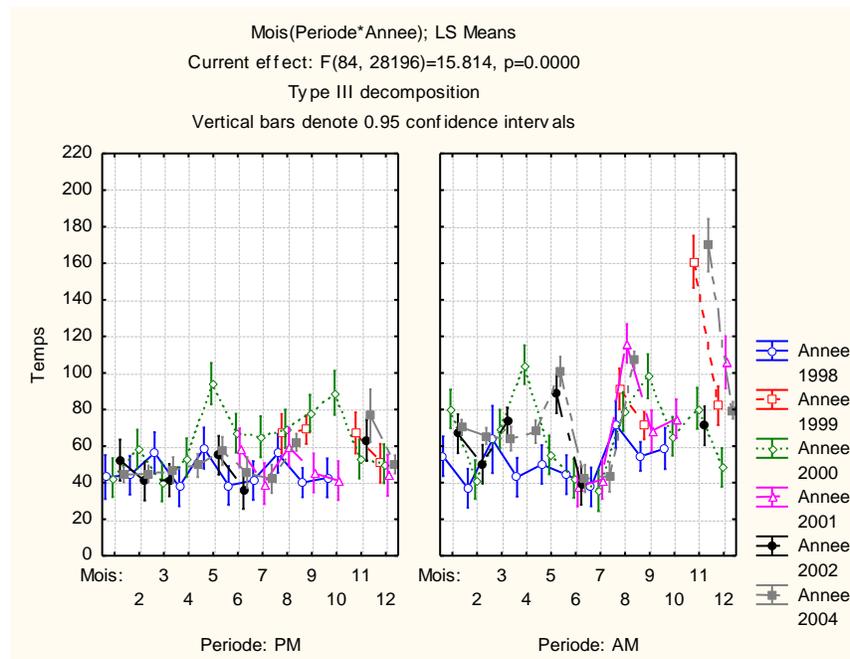


Figure 102. Moyennes et intervalles de confiance Mois(année\*période)

Tableau 66. ANOVA Mois(Année\*Période)

Tableau ANOVA Nested Design					
	Degrés de libertés	SS	MS	Rapport F	p-value
<b>Constante</b>	1	7.00E+07	7.00E+07	12019.74	0
<b>Periode</b>	1	1.64E+06	1.64E+06	282.21	0
<b>Annee(Periode)</b>	10	2.26E+06	2.26E+05	38.86	0
<b>Mois(Periode*Annee)</b>	84	7.73E+06	9.21E+04	15.81	0
<b>Erreur</b>	28196	1.64E+08	5821		
<b>Total</b>	28291	1.76E+08			

Tableau 67. Calcul du coefficient de signification

Coefficient de signification			
	R	R <sup>2</sup>	R <sup>2</sup> ajusté
<b>Temps</b>	0.26	0.07	0.07

En clair, nous pouvons souligner un meilleur R, mais aussi des tendances qui ne varient pas trop entre les années, en particulier pour la période PM, ou pour les mois de juillet aout AM. Par contre, il semble que le mois de novembre soit atypique avec une moyenne des temps de parcours plus élevée.

## 8.2 Tableaux

Tableau 68. ANOVA des temps de parcours selon Période et tronçons

Tableau ANOVA					
	Degrés de libertés	SS	MS	Rapport F	p-value
<b>Constante</b>	1	1.06E+08	1.06E+08	2.19E+04	0.00
<b>Tronçon</b>	21	2.03E+07	9.66E+05	1.99E+02	0.00
<b>Période</b>	1	2.44E+06	2.44E+06	5.02E+02	0.00
<b>Tronçon*Période</b>	21	1.66E+07	7.89E+05	1.63E+02	0.00
<b>Erreur</b>	28250	1.37E+08	4.85E+03		
<b>Total</b>	28293	1.76E+08			

**Tableau 69. ANOVA Nested Design pour la variable Mois(Année)**

Tableau ANOVA Nested Design					
	Degrés de libertés	SS	MS	Rapport F	p-value
<b>Constante</b>	1	6.96E+07	6.96E+07	1.15E+04	0.00
<b>Année</b>	5	1.46E+06	2.91E+05	4.84E+01	0.00
<b>Mois(Année)</b>	42	4.85E+06	1.15E+05	1.92E+01	0.00
<b>Erreur</b>	28244	1.70E+08	6.02E+03		
<b>Total</b>	28291	1.76E+08			

**Tableau 70. ANOVA selon conditions de la chaussée**

Tableau ANOVA					
	SS	Degrés de liberté	MS	Rapport F	p-value
<b>Constante</b>	5.74E+06	1	5.74E+06	9.30E+02	0.00
<b>Cond</b>	1.79E+06	5	3.58E+05	5.80E+01	0.00
<b>Error</b>	1.74E+08	2.83E+04	6.17E+03		