



HAL
open science

Évaluation de l'impact des systèmes de transports intelligents

Guillaume Saint Pierre

► **To cite this version:**

Guillaume Saint Pierre. Évaluation de l'impact des systèmes de transports intelligents. Statistiques [math.ST]. Université d'Evry-Val-d'Essonne, 2017. tel-01670569

HAL Id: tel-01670569

<https://hal.science/tel-01670569>

Submitted on 21 Dec 2017

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

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



CEREMA/DTerSO/DALETT/ESAD-ZELT

Habilitation à diriger des recherches

Présentée par :

Guillaume SAINT PIERRE

Évaluation de l'impact des systèmes de transports intelligents

Spécialité : Mathématiques appliquées

Soutenance le 19 octobre 2017 devant le jury constitué de :

Rapporteurs :	Oliver Carsten	Professeur, Institute for transport studies, University of Leeds
	Jean-Baptiste Lesort	Directeur de l'école nationale des travaux publics de l'État
	Jean-Michel Loubes	Professeur, Institut de Mathématiques de Toulouse
Examineurs :	Nour –Eddin El Faouzi	Directeur de recherche, IFSTTAR
	Saïd Mammari	Professeur, Université d'Evry Val d'Essonne
	Simon Cohen	Directeur de recherche émérite, Professeur à l'École des Ponts ParisTech
	Patricia Delhomme	Directeur de recherche, IFSTTAR

Remerciements

Ce mémoire fait le bilan de dix ans de recherches effectuées principalement au LCPC (Laboratoire Central des Ponts et Chaussées) et à l'IFSTTAR (Institut français des sciences et technologies des transports, de l'aménagement et des réseaux).

Ce mémoire est dédié à Didier Aubert, disparu brutalement en 2016 alors qu'il venait d'accepter d'être le directeur des travaux de cette HDR. Didier a été directeur adjoint du laboratoire LIVIC (laboratoire sur les Interactions Véhicules-Infrastructures-Conducteur) de l'IFSTTAR dans lequel j'ai effectué dix ans de travaux. Il a su m'encourager et me transmettre son expérience à maintes reprises, et ce document lui doit beaucoup.

Monsieur Oliver Carsten, Professeur à l'ITS (Institute for Transportation Studies) de l'université de Leeds (UK), est un éminent expert international de l'évaluation du comportement au volant et des systèmes d'assistance à la conduite. Je le remercie chaleureusement pour l'honneur qu'il me fait en acceptant de rapporter ce document, en français qui plus est.

Monsieur Jean-Baptiste Lesort est Directeur de l'ENTPE (école nationale des travaux publics de l'État), et me fait le privilège d'apporter son expertise et son regard transversal en acceptant la charge de rapporteur. Je le remercie de son soutien et de son intérêt pour mes travaux.

Monsieur Jean Michel Loubes est Professeur de statistique à l'IMT (institut de mathématiques de Toulouse) de l'UPS (université Paul Sabatier, Toulouse III). Je le remercie tout particulièrement d'accepter de rapporter ce mémoire et d'apporter son expertise en statistique appliquée au transport.

Monsieur Saïd Mammar, Professeur à l'université d'Evry a su m'encourager quand il le fallait. Je tiens à le remercier pour son soutien sans faille, et pour avoir accepté de participer à ce jury.

Monsieur Simon Cohen, Professeur à l'école de Ponts Paris Tech, a toujours répondu présent à mes sollicitations et constitue pour moi un repère et un exemple à suivre. Je suis fier de son soutien et de sa participation à ce jury.

Monsieur Nour-Eddin El Faouzi est Directeur de recherche à l'IFSTTAR et a su allier avec brio la statistique appliquée et la science du transport. Je le remercie pour ses conseils réguliers et l'intérêt qu'il porte à mon travail en acceptant de participer à ce jury.

Madame Patricia Delhomme est Directrice de recherche à l'IFSTTAR et est une partenaire incontournable de mes travaux impliquant la prise en compte de l'humain. Je la remercie pour ces longues années de collaboration et pour sa participation à ce jury.

Bien entendu, j'adresse les plus chaleureux remerciements à l'ensemble des personnes que j'ai côtoyées dans l'exercice du métier de chercheur. Techniciens, ingénieurs, administratifs, enseignants ou chercheurs, ils sont nombreux à avoir participé à la fertilité des idées et à avoir insufflé ce qu'il fallait de bonne humeur pour le bon accomplissement de ce long travail.

Pendant la rédaction de ce document, est né mon fils Benoît. Ce mémoire doit beaucoup à la patience et au soutien de mon épouse, ainsi qu'aux sourires de mon fils.

Présentation du document

Ce mémoire est constitué de sept parties :

- Un avant-propos synthétise mon parcours professionnel.
- Une seconde partie présente mon Curriculum Vitae détaillé. On y trouvera mon parcours professionnel, mes principales thématiques de recherche, ainsi que mon activité d'enseignement.
- La troisième partie est consacrée à mes activités d'administration et d'animation de la recherche. On y trouvera aussi mes collaborations de recherche, et l'ensemble de mes responsabilités exercées dans le cadre de projets de recherche financés.
- La quatrième partie est en lien avec l'enseignement et s'attache à décrire mes encadrements de doctorants, de post-doctorants, et de stagiaires de Master 2, ainsi que mes participations à des jurys.
- La cinquième partie propose une liste détaillée de l'ensemble de mes publications arrêtée à la date de 16 juillet 2017.
- La sixième partie est la plus conséquente de ce mémoire, et présente la synthèse de mes activités de recherche depuis 2006, et détaille mes principales contributions scientifiques.
- Une septième partie conclut ce mémoire en donnant les axes principaux de mon projet de recherche.

Table des matières

1. Avant-propos	13
2. Curriculum Vitae.....	15
2.1. Etat civil et coordonnées	15
2.2. Situation actuelle.....	15
2.3. Diplômes, titres universitaires et dates d’obtention.	16
2.4. Déroulement de carrière.....	16
2.5. Thématiques de recherche	19
2.6. Compétences.....	19
2.7. Formations	20
2.8. Enseignement	20
3. Administration et animation de la recherche.....	22
3.1. Charges administratives, animation d’équipes ou d’axes de recherche	22
3.2. Responsabilités collectives	23
3.3. Collaborations nationales et internationales.....	25
3.3.1. Institut géographique national (IGN), Sébastien Mustière (Directeur laboratoire COGIT) 25	
3.3.2. Institut du Véhicule Décarboné et Communicant et de sa Mobilité (VEDECOM), Bertrand Leroy.	26
3.3.3. Virginia Tech. Transportation Institute (VTTI), Dr. Feng Guo.	26
3.3.4. Laboratoire de Psychologie des Comportements et des mobilités (LPC-IFSTTAR), Patricia Delhomme (DR).....	26
3.3.5. Equipe perception et design sonore, Institut de recherche et coordination acoustique/musique (IRCAM), Patrick Susini (DR).....	27
3.3.6. Autres collaborations	27
3.4. Responsabilités et participations dans des projets de recherche nationaux ou internationaux.....	28
3.4.1. Résumé des implications	28
3.4.2. Détails des responsabilités	31
3.4.3. Détails des participations.....	33
4. Encadrements	36
4.1. Encadrement de doctorants.....	36
4.1.1. Résumé des encadrements de doctorants.....	36
4.1.2. Détails des encadrements de doctorants	37
4.1.3. Doctorants récompensés.....	41
4.2. Encadrements de post-doctorants.....	41
4.2.1. Résumé des encadrements de post-doctorants.....	41
4.2.2. Détails des encadrements de post-doctorants	42
4.3. Encadrement de stages (Master 2 recherche et école d’ingénieurs).....	44

4.4.	Participation à des jurys et à des instances ou comités en lien avec l'enseignement.	
		45
5.	Publications et communications	46
5.1.	Résumé des publications	46
5.2.	Présentation de 5 publications significatives	47
5.3.	Liste des publications et communications :	49
5.3.1.	Ouvrages scientifiques (OS)	49
5.3.2.	Articles dans des revues à comité de lecture répertoriées (ACL)	49
5.3.3.	Articles parus dans des revues à comité de lecture non répertoriées (ACLN)	50
5.3.4.	Communications avec actes dans congrès international (ACTI)	50
5.3.5.	Communications avec actes dans congrès national (ACTN)	52
5.3.6.	Communications invitées sans actes (INV)	53
5.3.7.	Autres publications (AP)	53
5.3.8.	Rapports de recherche y compris rapports de contrat de recherche (RAPP)	54
5.3.9.	Thèse/mémoire	55
5.3.10.	Autres productions :	55
6.	Synthèse des activités de recherche et principales contributions scientifiques	
		57
6.1.	Introduction générale	57
6.2.	Evaluation des systèmes embarqués d'assistance à la conduite : généralités	62
6.2.1.	Introduction	62
6.2.1.	Contexte	63
6.2.2.	Recueillir les données	67
6.2.3.	Définitions	70
6.2.4.	Protocoles d'évaluation des systèmes utilisant la conduite en situation naturelle	76
6.2.5.	Méthode d'analyse par question de recherche	79
6.2.6.	Modèles statistiques pour données catégorielles non indépendantes	90
6.2.7.	Evaluer l'impact sur la sécurité	96
6.2.8.	Evaluer l'impact sur l'environnement	110
6.2.9.	Publications associées à cette section	114
6.3.	Application : Systèmes régulateurs de vitesse	118
6.3.1.	Impact d'un limiteur de vitesse adaptatif sur la consommation de carburant (Projet LAVIA)	119
6.3.2.	Impact du limiteur et du régulateur de vitesse (Projet EuroFOT)	130
6.3.3.	Publications associées à cette section	137
6.4.	Application : Systèmes d'aide à l'économie d'énergie	139
6.4.1.	Introduction	139
6.4.2.	La conduite éco-responsable (écoconduite)	140
6.4.3.	Facteurs de réduction de consommation et règles d'or de l'écoconduite	142
6.4.4.	Lien entre vitesse et consommation	147
6.4.5.	Lien entre infrastructure et consommation	151
6.4.6.	Caractérisation et impact de l'écoconduite	154

6.4.7.	Conception et mise en œuvre d'un éco-index	157
6.4.8.	Impact d'un système d'aide à l'écoconduite : ecoDriver.....	161
6.4.9.	Publications associées à cette section	166
6.5.	Les Méga Données au service de la mobilité intelligente.....	168
6.5.1.	Introduction	168
6.5.2.	Modélisation des profils de vitesse.....	169
6.5.3.	Détection de l'infrastructure.....	179
6.5.4.	Conclusion.....	184
6.5.5.	Publications associées à cette section	185
6.6.	Statistique appliquée et sciences humaines : collaborations en psychologie expérimentale et acoustique.	187
6.6.1.	Introduction	187
6.6.2.	Psychophysique.....	188
6.6.3.	Psychologie du comportement et des mobilités	189
6.6.4.	Publications associées à cette section	193
7.	Projet de recherche.....	194
7.1.	Impact des systèmes coopératifs de transports intelligents.....	195
7.2.	Méga données et apprentissage statistique.....	197
7.3.	Conclusion.....	198
8.	Bibliographie.....	200

Liste des figures

Figure 1 : Evolution du nombre de participations à des projets de recherche.....	35
Figure 2 : Evolution du cumul de publications	47
Figure 3: Proposition de structuration de l'ensemble de mes recherches (2006-2016)	59
Figure 4 : Mise en évidence des axes de recherche développés dans cette section.....	63
Figure 5 : Illustration du "cycle en V" utilisé dans les FOT (Field Operational Tests).....	66
Figure 6 : Exemple d'adaptateur et de prise OBD véhicule	67
Figure 7 : Exemple d'instrumentation très complète pour l'étude de la conduite en situation naturelle pour le projet UDrive (FP7)	68
Figure 8 : Exemple d'instrumentation légère avec boîtier autonome avec capteurs intégrés et connexion à la prise OBD-II du véhicule (Projet MEDOC, Fondation MAIF).....	68
Figure 9 : Solution hybride de recueil de données du projet européen ecoDriver. En conduite normale, seul un boîtier enregistreur est connecté à la prise OBD et enregistre les données du véhicule. En période de test d'une application nomade.	69
Figure 10 : Résumé des différentes sources de données possibles lors d'une étude du comportement de conduite.....	69
Figure 11 : Illustration de la notion d'évènements et de variable situationnelle pour la conduite.	75
Figure 12: Synopsis de la méthode d'analyse d'une question de recherche par test d'hypothèse.	81
Figure 13: Illustration pour une série de trajets effectués par 6 conducteurs différents sur une période de temps identique. Les variables à valeur manquantes ne peuvent être indiquées globalement.....	82
Figure 14: Illustration de l'étape de sélection des données.....	83
Figure 15: Illustration de la forme la plus simple d'identification des situations homogènes	84
Figure 16: Illustration de l'étape d'homogénéisation pour deux facteurs.La variable "Contrôle/Traitement"est combinée avec la variable luminosité (jour/nuit) afin d'identifier quatre conditions différentes dans les données.....	84
Figure 17: Illustration de l'étape d'homogénéisation pour trois facteurs.La variable "Contrôle/Traitement"est combinée avec la variable luminosité (jour/nuit) et type de route (3 niveaux) afin d'identifier les conditions suivantes : urbain-Jour, rural-Jour, Autoroute-Jour, urbain-nuit, rural-nuit, Autoroute-nuit, dans chaque modalité de contrôle/Traitement.....	85
Figure 18: Illustration de la méthode d'agrégation par trajet.....	85
Figure 19: Illustration de l'agrégation sur une base de temps homogène.	86
Figure 20: Exemple d'identification par micro-trajet.	87
Figure 21: Illustration de la réduction de données par micro-trajets.....	87
Figure 22: Illustration d'un calcul d'indicateur sur sections inhomogènes. La taille de la valeur indiquée représente la pondération qu'il faut appliquer lors de l'analyse.	88

Figure 23 : Nombre de tués sur les routes nationales françaises (Source ONISR).	96
Figure 24 : évolution schématique de la gravité d'un accident selon la violence du choc.....	98
Figure 25 : Pyramide de Heinrich pour la sécurité routière.....	98
Figure 26: Schéma d'analyse du lien entre incidents et conditions de circulation.....	108
Figure 27 : Interpolation de la cartographie moteur d'une Laguna II permettant de calculer la consommation instantanée (en gr/sec) en fonction du couple et du régime du moteur (en tours/min).	111
Figure 28 : Exemple de courbes reliant la vitesse moyenne aux émissions de CO2 pour une variété de véhicules (source NAEI, et Carsten et al. 2008)	113
Figure 29 : Mise en évidence des axes de recherche développés dans cette section.....	118
Figure 30 : Boîtes à moustache des consommations mesurées lors des trajets limités à 50km/h du LAVIA selon l'ordre des jours de l'expérimentation. La courbe reliant les quantiles à 95% est en bleu, la régression linéaire reliant ces mêmes quantiles est en rouge, tandis que la régression polynomiale par morceaux est en vert.	123
Figure 31 : Exemple de profils de vitesse moyens sur une section limitée à 50km/h avec présence d'un feu tricolore selon le mode du LAVIA et selon la présence d'un arrêt au feu.	126
Figure 32 : Séminaire de clôture du projet euroFOT (http://www.eurofot-ip.eu/).	130
Figure 33 : Odds-ratios de l'usage du limiteur (SL) et du régulateur (CC) selon les conditions de conduite (SRS=Speed regulation system).	133
Figure 34 : Odds ratios de certains événements de la conduite lors de l'usage du limiteur (SL, gauche) et du régulateur (CC, droite).....	133
Figure 35 : Répartition des kilomètres parcourus selon le système (Speed limiter SL, cruise control CC) et les limitations de vitesse.....	134
Figure 36 : économies de carburant potentielles en cas d'équipement complet de la flotte des véhicules personnels avec le limiteur de vitesse dans l'Europe à 27.....	136
Figure 37 : Economies de carburant potentielles en cas d'équipement complet de la flotte des véhicules personnels avec le régulateur de vitesse dans l'Europe à 27.	137
Figure 38 : Mise en évidence des axes de recherche développés dans cette section.....	140
Figure 39 : Clio III du LIVIC et système de recueil de données du véhicule et des capteurs additionnels (radar et caméras).	142
Figure 40 : Vue d'ensemble des principales stratégies de réduction de la consommation de carburant.....	147
Figure 41 : Nuage des vitesses et consommations moyennes de 6465 trajets réels (projet LAVIA), et estimation par polynômes locaux de la relation entre vitesse et consommation moyenne selon le type de conduite (conduite neutre, après formation à l'écoconduite, et en utilisant une assistance embarquée d'aide à l'écoconduite, données issues du projet GERICO).	148
Figure 42 : Estimation par polynômes locaux du lien entre consommation et vitesse instantanée pour différents rapports de boîte (données GERICO), pour l'ensemble de la conduite (gauche) et pour les seules phases d'accélération (droite).	149

Figure 43 : Analyse descriptive des séquences de rapports de boîtes (colonnes de gauche) et de classes de régimes moteurs (colonnes de droite) pour 3 catégories de conducteurs sur une distance de 1500 mètres pendant la traversée d'une agglomération.....	150
Figure 44 : Profils de vitesses moyens (30 conducteurs) selon le type de conduite, mis en rapport avec la limitation de vitesse en vigueur.	151
Figure 45 : Traversée de cinq ronds-points. Visualisation du profil de consommation moyen (haut), du profil de vitesse moyen associé (milieu), et du profil de consommation instantanée selon le type de conduite (bas).....	152
Figure 46 : Consommation cumulée lors de la traversée de 5 ronds-points, comparée au modèle linéaire simple obtenu.....	153
Figure 47 : Profils de consommation cumulée selon le nombre d'arrêts au feu tricolore.	154
Figure 48 : Liens entre valeurs observées des indicateurs associés aux règles d'or et l'éco-index associé, selon le niveau du conducteur.....	160
Figure 49 : Ecoindex global selon le niveau et pour chaque seconde d'un trajet réel.....	160
Figure 50 : Interface de l'application android ecoDriver mise au point par le CTAG.....	162
Figure 51 : Mise en évidence des axes de recherche développés dans cette section.....	168
Figure 52 : Lien entre les trois espaces : [distance temps, vitesse temps] et [vitesse distance].	171
Figure 53 : Diagramme fonctionnel illustrant la définition des profils spatiaux de vitesse.....	172
Figure 54 : Application de la procédure de lissage sur 78 profils spatiaux de vitesse.....	174
Figure 55 : Aligement des profils par la méthode des landmarks dans le cas où le feu est rouge (36 courbes). Le profil moyen est représenté en noir.....	176
Figure 56 : Boîtes à moustaches ponctuelles et fonctionnelles dans le cas où le feu est rouge (36 courbes).....	178
Figure 57 : Illustration de la méthode d'extraction par fenêtres glissantes. Dans notre cas, $L=100m$, et $t=10m$	180
Figure 58 : Quatre exemples de fenêtres obtenues. Seules les deux fenêtres du bas contiennent un feu rouge.....	181
Figure 59 : Courbes ROC et bandes de confiance à 95% (méthode Threshold Averaging avec distribution binomiale) pour chaque combinaison [approche x algorithm]. Ratio de faux positifs (FPR) en abscisse et vrai positifs (TPR) en ordonnée.	183
Figure 60 : Mise en évidence des axes de recherche développés dans cette section.....	187
Figure 61 : Effet "fer à cheval" produit par une projection sur un espace à deux dimensions de sons identifiés par les « pseudo-distances » les séparant. Les sons de la catégorie « +++ » sont artificiellement proches des sons de la catégorie « + », et l'algorithme recommande l'ajout d'une dimension supplémentaire (non réellement présente) pour tenir compte de cet aspect. Ce phénomène est dû à l'absence de la propriété d'inégalité triangulaire de la pseudo-distance utilisée.	189
Figure 62 : Exemple des 4 configurations de signalisation lumineuse étudiées.	190
Figure 63 : Moyennes estimées par un modèle GEE et intervalles de confiance pour le taux de détection en fonction de la configuration lumineuse. La différence est significative entre la configuration standard et jaune (astérisque) et faiblement significative entre la configuration	

standard et celle sur le casque (astérisque entre crochets). La configuration « casque » est cependant mieux détectée que toutes les autres lorsque le motocycliste se situe à grande distance de l'observateur. 190

Figure 64 : Mise en évidence des axes de recherche que je souhaite développer dans le cadre de ma prise de poste au Cerema..... 195

Liste des tableaux

Tableau 1 : Résumé des publications par année.....	46
Tableau 2 : Proposition de catégorisation de certains systèmes d'assistance à la conduite (source : étude IDACA).....	64
Tableau 3 : Exemples de dimensionnement de flottes expérimentales de véhicules dans différents projets auxquels j'ai participé.....	70
Tableau 4 : Vue d'ensemble des différents modèles linéaires généralisés les plus courants.....	89
Tableau 5 : Seuils utilisés pour l'identification des incidents de la route dans quelques projets de référence.....	105
Tableau 6 : Moyennes de la consommation de carburants en litres/100km pour chacune des zones étudiées, et selon le mode d'utilisation du LAVIA. La valeur de la statistique F de l'anova et sa p-value sont données dans les deux dernières colonnes.....	121
Tableau 7 : Modèle proposé pour l'évolution de la consommation selon les modes du LAVIA.	124
Tableau 8 : valeurs de t_1, t_2 et t_3 en fonction de t	124
Tableau 9 : Paramètres estimés par le modèle marginal, donnant l'effet des modes du LAVIA sur la consommation de carburant pour l'ensemble des trajets.	125
Tableau 10 : Paramètres estimés par le modèle marginal, donnant l'effet du mode du LAVIA et des limitations de vitesse sur la probabilité d'un dépassement de vitesse volontaire (type 2).....	127
Tableau 11 : Bilan des données de conduite recueillies dans euroFOT pour l'analyse du limiteur et du régulateur de vitesse.....	131
Tableau 12 : Liste des hypothèses de recherche testées lors du projet euroFOT pour l'étude du système régulateur de vitesse (Speed regulation system, SRS) composé du limiteur (speed limiter, SL) et du régulateur (cruise control, CC). Le chunk désigne un intervalle de temps de 30 secondes de conduite, homogène selon la limitation de vitesse (cf. section 6.2.5.2.4).	131
Tableau 13 : évaluation des effets globaux du limiteur et du régulateur de vitesse.	135
Tableau 14 : règles d'or de l'éco-conduite.	143
Tableau 15 : Autres conseils pour économiser le carburant.....	145
Tableau 16 : Règles d'or de l'écoconduite et indicateurs associés.	154
Tableau 17 : Evolution des indicateurs de performance selon le style de conduite, pour chacune des expérimentations effectuées, et test de significativité (T-test).....	156
Tableau 18 : Percentiles des valeurs observées pour les 4 indicateurs associés aux règles d'or de l'écoconduite (conduite normale uniquement).....	159
Tableau 19 : Principaux résultats du projet ecoDriver concernant l'impact de l'application nomade Android, en conditions contrôlées et en conditions naturelles.....	163

Tableau 20 : Performances des approches images et fonctionnelle pour la méthode Random Forest. Tous les écarts sur les indicateurs sont significatifs ($p < 0.01$). ¹ : Taux de confusion avec les signes stops, ² : Temps d'apprentissage (en secondes / fenêtre)..... 183

1. Avant-propos

J'ai obtenu mon Doctorat en Mathématiques Appliquées en Septembre 2003 à l'Université Paul Sabatier (Toulouse III) après y avoir effectué la totalité de mon cursus universitaire. J'ai ensuite effectué plusieurs Post-Doctorats dans différents organismes : Université de Talence (Bordeaux I), INRIA (équipe SELECT) à l'Université d'Orsay (Paris-Sud XI), IRCAM (Institut de Recherche et Coordination Acoustique/Musique) et INRETS (Institut National de Recherche sur les Transports et leur Sécurité).

En janvier 2007, je suis devenu Chargé de Recherches à l'IFSTTAR (Institut français des sciences et technologies des transports), organisme de recherche public issu de la fusion entre l'INRETS et le Laboratoire Central de Ponts et Chaussées (LCPC). J'ai intégré le laboratoire sur les Interactions Véhicules-Infrastructures-Conducteur (LIVIC) du département Composants et Systèmes (COSYS), dont la mission est d'exercer des activités de recherche, de développement et d'expérimentation visant à faire émerger des systèmes intelligents d'aides à la conduite. Au sein de cette équipe, je me suis employé à mettre mes compétences en statistique appliquée au service de la mission qui m'était confiée : prendre en charge l'activité d'évaluation des systèmes embarqués d'assistance à la conduite.

Pendant ces dix années, j'ai travaillé sur l'évaluation multicritère des effets des systèmes embarqués d'assistance à la conduite (limiteurs et régulateurs de vitesse par exemple). Je m'intéresse aux effets sur le comportement et la sécurité routière, avec un intérêt particulier pour les économies d'énergie. Mes travaux participent tout autant à la mise au point de ces systèmes embarqués, qu'à leur évaluation en conditions réelles mettant l'humain au cœur du processus. Ils s'appuient sur l'exploitation statistique de grandes quantités de "traces" numériques générées par les véhicules modernes. Je cherche à quantifier objectivement les effets d'un changement de style de conduite (écoconduite par exemple), d'un élément de l'infrastructure (radars, passages piétons, etc.), ou d'un système embarqué d'assistance à la conduite, sur le comportement et les actions du conducteur.

Sur le plan de l'enseignement j'ai maintenu une activité constante en étant responsable pendant huit ans d'un module de statistique appliquée en Master 2 de l'UPMC (Université Pierre et Marie Curie), et en encadrant régulièrement des stages étudiants.

Au premier mars 2017, j'ai obtenu une mutation au Cerema dans l'équipe ESAD-ZELT (DTER Sud-Ouest) dont la mission consiste à mener des activités d'études et de recherche autour de l'évaluation des systèmes d'aide au déplacement.

Mon activité de recherche a été fortement influencée par l'avènement de la mobilité numérique et la possibilité d'intégrer les « traces numériques de déplacement » dans l'évaluation multicritère d'une assistance à la mobilité. J'ai développé des méthodologies d'évaluation

adaptées à ce nouveau contexte, que j'ai pu mettre en œuvre concrètement dans de nombreux projets de recherche en relation avec le monde industriel. L'émergence d'outils efficaces pour traiter les « méga-données » générées par la mobilité numérique m'a conduit à imaginer de nouvelles solutions aux problématiques du transport, actuelles ou futures. L'activité foisonnante de ce domaine m'incite à continuer dans cette voie pour apporter des réponses aux questions soulevées par l'avènement des véhicules autonomes et connectés.

Fort d'une thématique de recherche en plein développement, et de l'expérience acquise aux niveaux national et international, je souhaite continuer mes travaux en développant mes activités d'encadrement et de transmission du savoir grâce à l'obtention de l'habilitation à diriger des recherches.

2. Curriculum Vitae

2.1. Etat civil et coordonnées

Guillaume SAINT PIERRE

Né le 08 Décembre 1975 à Toulouse (41 ans), nationalité Française

Marié, 1 enfant (né en 2016)

Adresse professionnelle :

CEREMA/DTerSO/DALETT/ESAD-ZELT,
1 Avenue du Colonel Roche,
31400 Toulouse
☎ 05 62 25 97 86

Adresse personnelle :

61 Avenue de l'URSS,
31400 Toulouse
☎ 06 72 33 43 76

✉ guillaume.saintpierre@cerema.fr

🌐 <http://perso.lcpc.fr/guillaume.saint-pierre/>

🌐 https://www.researchgate.net/profile/Guillaume_Saint_Pierre

🌐 Google Scholar : <https://lc.cx/4v9f>

2.2. Situation actuelle

Corps et grade : Chargé de recherche du développement durable (1ère classe)

Ancienneté dans le grade : 5 ans (6ème échelon)

Ancienneté totale : 10 ans au 1er janvier 2017

Organisme d'affectation : CEREMA, centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement

Unité de recherche/service : DTerSO/DALETT/ESAD-ZELT.

Date d'affectation dans l'organisme : 1er mars 2017

2.3. Diplômes, titres universitaires et dates d'obtention.

2003	Qualification aux fonctions de maître de conférences CNU section 26.
1999-2003	<p>Doctorat de Mathématiques Appliquées (Probabilités et Statistiques), Université Paul Sabatier (Toulouse III).</p> <p>Dirigé par le Pr. Bernard Garel et financé par une bourse M.N.E.R.T.</p> <p>Titre : « Identification du nombre de composants d'un mélange gaussien par chaînes de Markov dans le cas multivarié, ou par maximum de vraisemblance dans le cas univarié. »</p> <p>Devant le jury composé de :</p> <p>M. Philippe BESSE Professeur à l'Université Toulouse III Examineur</p> <p>M. Gilles CELEUX Directeur de Recherche (INRIA) Examineur</p> <p>M. Bernard GAREL Professeur à l'I.N.P. (ENSEEIH) Directeur de thèse</p> <p>M. Gérard GOVAERT Professeur à l'U.T.C. Rapporteur</p> <p>M. Christian ROBERT Professeur à l'Université Paris Dauphine Rapporteur</p> <p>Mme. Christine THOMAS-AGNAN Professeur à l'Université Toulouse I Examineur</p>
1998-1999	Master 2 recherche (D.E.A.) en Mathématiques appliquées, option statistique, mention bien, Université Paul Sabatier (Toulouse III). Boursier au mérite.
1995-1998	Licence (option Mathématiques, Informatique et Applications aux Sciences) puis Maîtrise de Mathématiques fondamentales (option statistique), Université Paul Sabatier (Toulouse III).
Avant 1995	Cursus scolaire effectué pendant 12 ans en Afrique (Tunisie, Sénégal, Côte d'Ivoire), puis au Lycée Saint Sernin à Toulouse.

2.4. Déroulement de carrière

Depuis 03.3017	Chargé de recherche dans le domaine de l'évaluation des systèmes de transports intelligents appliqués à la gestion des déplacements au CEREMA
Depuis 2016	Expert contractuel auprès de la commission Européenne : Evaluation des candidatures aux bourses Marie-Curie (H2020-MSCA-IF-2016).
2008-2016	Responsable du module « Statistique appliquée et logiciel SAS » du Master 2 recherche de Statistique de l'Université Pierre et Marie Curie, UPMC-Paris 6 (40h/an).

01.2007- 02.2017	<p>Chargé de recherche au LIVIC (LCPC puis IFSTTAR).</p> <ul style="list-style-type: none"> • Expert en évaluation des systèmes de transports intelligents (ITS). Spécialité en traitement des données massives de véhicules et évaluation de l'impact des systèmes embarqués d'assistance à la conduite.
<hr/>	
02.2006- 09.2006	<p>Chargé de recherche en Statistique à l'Institut National de Recherche sur les Transports et leur Sécurité (INRETS).</p> <ul style="list-style-type: none"> • CDD de 8 mois au sein de l'équipe GRETTIA. Projets SARI/IRCAD et INTRO. • Analyse et traitement de données météorologiques et routières (SAS). • Mise au point d'un modèle de prévision du sur-risque d'accident par temps de pluie.
<hr/>	
12.2005- 01.2006	<p>Chargé de recherche en Statistique à l'Institut de Recherche et Coordination Acoustique et Musique (IRCAM).</p> <ul style="list-style-type: none"> • CDD de 2 mois au sein de l'équipe Perception et Design Sonores, financé par l'entreprise Renault. • Analyse de données d'expérience de catégorisation orientée sur des sons électroniques. • Programmation de méthodes de type « data mining ». (Classification, Régression multiple, logistique, MDS)
<hr/>	
2004-2005	<p>Post-Doctorant à l'INRIA à l'Université d'Orsay (Paris-Sud XI).</p> <ul style="list-style-type: none"> • Equipe SELECT de l'Institut National de Recherche en Informatique et Automatique (I.N.R.I.A.). • Mise au point d'un modèle de classification pour données de distances. • Analyse Exploratoire et Discriminante de données en grande dimension (ACI Data High Dim).
<hr/>	
2003-2004	<p>Post-Doctorant à l'Université de Talence (Bordeaux I).</p> <ul style="list-style-type: none"> • Bourse du Ministère de l'Economie et des Finances en partenariat avec l'entreprise Thalès Avionics, dans le cadre d'une participation au projet Européen ACLIMAT. • Mise au point d'un modèle de fiabilité pour du matériel aéronautique embarqué hautement fiable. • Rédaction de plusieurs rapports de recherche destinés aux partenaires industriels.
<hr/>	
2002-2003	<p>Attaché Temporaire d'Enseignement et de Recherche en 26^{ème} section de l'Université Paul Sabatier (Toulouse III).</p>
<hr/>	
1999-2002	<p>Moniteur de l'Université Paul Sabatier (Toulouse III)</p>

2.5. Thématiques de recherche

Je travaille sur l'évaluation multicritère des effets des systèmes embarqués d'assistance à la conduite (limiteurs et régulateurs de vitesse par exemple). Je m'intéresse aux effets sur le comportement et la sécurité routière, avec aussi un intérêt particulier pour les économies d'énergie. Mes travaux participent tout autant à la mise au point de ces systèmes embarqués, qu'à leur évaluation en conditions réelles mettant l'humain au cœur du processus. Mon travail repose sur l'exploitation statistique de grandes quantités de "traces" numériques générées par les véhicules modernes. Je cherche à quantifier objectivement les effets d'un changement de style de conduite (écoconduite par exemple), d'un élément de l'infrastructure (radars, passages piétons, etc.), ou d'un système embarqué d'assistance à la conduite, sur le comportement et les actions du conducteur. En parallèle, je m'intéresse aux algorithmes d'apprentissage statistique, capables d'extraire de l'information pertinente pour la gestion des transports à partir des traces numériques de déplacement.

Mon activité de recherche est décrite en détail à partir de la page 57.

2.6. Compétences

Langues

Français (langue maternelle), anglais (courant, parlé et écrit), espagnol (basique, parlé).

Communication

- Excellente expérience d'écriture de rapports et de communication dans le cadre de projets de recherche internationaux en lien avec l'industrie.

Liste des publications et communications page 49, rapports de contrats page 54.

- Excellente expérience de présentation devant de larges audiences, workshops ou conférences.

Liste des conférences invitées page 53.

- Longue expérience d'enseignement (12 ans).

Activités d'enseignement page 20.

Organisation

- Très bonne expérience de coordination de la recherche.

Liste des animations d'axes de recherche page 22.

- Excellente expérience de montage, coordination, gestion et participation à des projets de recherche internationaux (SP Leader, Core group member, work package leader), et nationaux (responsable d'études ministérielles, et de tâches-projets ANR).

Détails des responsabilités et participation dans des projets de recherche page 28.

- Excellente expérience d'encadrement (11 master 2, 3 post-doctorants, 4 doctorants).

Liste des encadrements page 36.

Informatique

- Excellente maîtrise des outils de bureautique (Microsoft et open source).
- Excellente maîtrise des logiciels statistiques (SAS, R, SPSS).

2.7. Formations

- Formation « écoconduite » MEDDE (ministère de l'Ecologie, du Développement durable et de l'Energie), février 2012.
- Formation « conduite en sécurité », MPSRA (Mission prévention et sécurité routières dans les armées), 2010.
- Formation « Les transports durables », IGPDE (Institut de la gestion publique et du développement économique), 29 septembre 2008.
- Formation au logiciel MapInfo, 4 jours, du 8/9/2008 au 12/9/2008.
- Formation « Effet de serre et énergie », IGPDE, 1 journée, 27 mars 2008.
- Formation au logiciel SAS, 3 jours, INSERM (Institut national de la santé et de la recherche médicale), Mai 2007.

2.8. Enseignement

J'ai commencé à enseigner lors de mon doctorat à l'université de Paul Sabatier de Toulouse et mon poste de moniteur, et j'ai poursuivi en tant que demi ATER. Cela m'a donné le goût de l'enseignement des mathématiques de l'aléatoire. J'ai eu la chance de pouvoir enseigner les probabilités et statistiques à un public d'élèves ingénieurs (l'E.N.S.I.C.A. et E.N.S.E.E.I.H.T.) et ainsi de transmettre mon goût de la science appliquée.

Une fois en poste au LIVIC à partir de 2007, j'ai cherché à retrouver le plaisir de l'enseignement en université. J'ai eu l'opportunité de prendre la responsabilité dès 2008 d'un module complet du Master 2 de statistique appliqué du LSTA (Laboratoire de Statistique théorique et appliqué) de Paris 6. Cet enseignement de la mise en œuvre concrète sur logiciel des méthodes avancées de la statistique correspond très bien à mes compétences et mon quotidien de chercheur appliqué. Cet enseignement étant régulièrement plébiscité par les étudiants, il a été renouvelé jusqu'à mon départ en 2017. J'ai ainsi eu la chance de pouvoir transmettre mon expérience de chercheur à haut niveau, tout en ayant la possibilité de transmettre l'expertise « transport » par le biais d'encadrements de stagiaires (cf. liste des encadrements de stage *page 44*).

En outre, j'ai eu la possibilité de participer à un cycle de formation organisé par l'école des ponts à destination des professionnels de la gestion des infrastructures routières.

2008-2016 Responsable du module « Statistique appliquée et logiciel SAS » du Master 2 recherche de Statistique de l'Université Pierre et Marie Curie, UPMC-Paris 6 (40h/an).

2009, 2010 Formateur à l'école Nationale des Ponts Paris Tech sur le thème « Concilier respect des vitesses, sécurité, et économies d'énergies : Les promesses de la conduite éco-responsable » (2h/an).

Public : Professionnels du transport et de la gestion des infrastructures

2002-2003 Attaché Temporaire d'Enseignement et de Recherche de l'Université Toulouse III (demi-poste, 96h/an).

- Travaux pratiques avec le logiciel Scilab en D.E.U.G. scientifique.
 - Travaux dirigés de Statistiques à l'E.N.S.I.C.A. (Ecole Nationale Supérieure d'Ingénieurs de Constructions Aéronautiques)
 - Travaux dirigés de Statistiques à l'E.N.S.E.E.I.H.T. (Ecole Nationale Supérieure d'Electronique Electrotechnique, Informatique et Hydraulique de Toulouse).
-

1999-2002 Moniteur de l'Université Toulouse III (96h/an).

- Travaux pratiques avec les logiciels Matlab, Maple et Scilab en DEUG scientifique et classes préparatoires.
- Travaux dirigés de Probabilités et Statistiques en D.E.U.G. et licence de Mathématiques.
- Travaux dirigés de Statistiques à l'E.N.S.I.C.A. et à l'E.N.S.E.E.I.H.T.

3. Administration et animation de la recherche

3.1. Charges administratives, animation d'équipes ou d'axes de recherche

En tant que chargé de recherche, mes charges administratives sont réduites malgré mes activités d'enseignement à Paris 6. Cependant, je participe très activement à la gestion de nombreux projets de recherche, dans lesquels j'exerce des responsabilités décisionnelles, ainsi que de l'encadrement de post-doctorants (cf. liste des projets *page 28*).

Au-delà de ces aspects, je coordonne des activités de recherche sur le long terme tout autant à l'échelle nationale qu'à l'échelle de mon laboratoire. De 2014 à 2017 j'ai été en charge de l'animation de recherche dédiée au traitement de l'information pour la maîtrise de la mobilité au LIVIC. Axe que j'ai contribué à développer de par mon expérience de la statistique appliquée.

Je suis maintenant en poste au Cerema (1^{er} mars 2017), membre d'une équipe recherche en préfiguration, pour laquelle j'anime un des deux axes de recherche (Réseaux de transport et système de transports intelligents).

J'ai accepté par deux fois la charge de contribuer à coordonner une action de recherche de l'Ifsttar, dont l'objectif est de mener pendant 3-5 ans un programme de recherche concerté avec le réseau scientifique et technique (maintenant Cerema). J'ai eu la charge de coordonner les recherches autour des impacts de l'écoconduite, de superviser plusieurs études, et d'organiser des journées scientifiques sur ce thème. J'ai représenté l'Ifsttar et l'axe « étude la conduite en situation naturelle » dans la collaboration internationale qui l'associe avec le Virginia Tech Transportation Institute (VTTI, USA).

En lien avec mon expertise développée au sein du LIVIC, j'ai contribué à la vie scientifique du département Cosys de l'Ifsttar en organisant des ateliers de réflexion autour des données du transport.

2017 Animateur de l'axe « Réseaux de transport et système de transports intelligents » de l'équipe-projet recherche « Systèmes de transports » du Cerema.

2015-2016 Organisation d'atelier lors du séminaire annuel du département :

- Animation de l'atelier prospectif "Big Data & Deep Learning"
- Animation de l'atelier "Big Data et transport"

2014-2016 Animateur de l'axe « Traitement de l'information pour la maîtrise de la mobilité » (TIMM) du laboratoire LIVIC.

2014-2016 Responsable de l'axe « Naturalistic driving studies » (Collaboration Ifsttar / Virginia Tech Transportation Institute, USA)

- Collaboration internationale visant à améliorer les méthodes de détection statistique des incidents.

2009-2014 Coordinateur de l'action 5 (éco-conduite : sobriété et sécurité) de l'opération de recherche SERRES (Solutions pour une Exploitation Routière Respectueuse de l'Environnement et de la Sécurité) commune entre l'IFSTTAR et le CEREMA.

- Organisation et suivi pluriannuel d'un programme de recherche commun
- Organisation du séminaire intermédiaire (mars 2013) et du séminaire de clôture (mai 2014).
- Encadrement scientifique de plusieurs études techniques :
 - Kemmel E. « Évaluation des impacts d'une formation à l'éco-conduite sur la consommation énergétique de véhicules légers : le cas du CETE de l'Ouest », 2011-2012.
 - Romon S. « Consommation des véhicules en fonction de l'infrastructure », CETE Sud-Ouest, 2012-2013.
 - Hublart A. « Rapport d'étude de l'écoindex », CETE Normandie-Centre, 2011-2012.

3.2. Responsabilités collectives

Ma prise de poste s'est effectuée initialement au LCPC (Laboratoire central des ponts et chaussées), organisme qui était déjà sous la gestion du ministère de l'écologie (ex équipement). A ce titre, mon rôle a aussi consisté à coopérer avec les directions de l'administration centrale comme la DGITM (direction générale des infrastructures, des transports et de la mer) ou la DSCR (Direction de la sécurité et de la circulation routières). C'est donc très naturellement que, dans ce cadre, j'ai répondu aux différentes sollicitations rencontrées en adéquation avec mes thèmes de recherche.

J'ai mené plusieurs études financées par ces directions (cf. liste des projets *page 28*) et ai coordonné deux opérations de recherche communes entre l'Ifsttar et le Cerema. Ces opérations de plusieurs années m'ont permis d'élargir ma vision de mes propres recherches grâce aux liens tissés avec le réseau scientifique et technique du ministère (Cete, LRPC, etc. maintenant Cerema).

Je suis maintenant reconnu dans ces directions en tant qu'expert de l'évaluation des systèmes embarqués et des véhicules autonomes et connectés. Je suis donc sollicité régulièrement pour participer à des groupes de travail temporaires ou permanents.

Je suis ainsi membre de plusieurs groupes de travail et comités, du ministère et de l'Ifsttar, et contributeur du site web institutionnel des transports intelligents. La liste détaillée de mes responsabilités collectives est donnée ci-après.

2016	Animateur du réseau métier « analyse de données » au sein de l'Ifsttar
Depuis 2015	Membre du comité technique Cerema Ifsttar « sécurité des déplacements routiers »
Sept. 2015	Membre du jury du concours d'ingénieur d'étude pour le laboratoire SPLOTT (laboratoire Systèmes Productifs, Logistique, Organisation des Transports, et Travail), AME-IFSTTAR.
2015-2017	Membre du groupe de travail ONISR (observatoire national interministériel de la sécurité routière), « évolution à moyen terme de l'observatoire des vitesses ».
Depuis 2015	Membre de la commission de normalisation BNA/CN39/GT8 « Systèmes embarqués de commande et d'information du transport ».
Depuis 2015	Membre du groupe ISO/TC22/SC39/WG08 et de la task-force « Naturalistic driving studies ».
Depuis 2015	Membre du GERI (Groupe d'échanges et de recherches ifsttar) « écoconduite ».
2014-2016	Membre du GERI ANIM@TIC (Recueil et analyse de données spatio-temporelles pour l'étude des mobilités et des comportements de conduite).
Depuis 2010	Membre du groupe « Transports intelligents » du ministère de l'écologie (MEDDE) qui se réunit deux fois par an.
Depuis 2010	Membre du comité de rédaction du site web officiel du ministère de l'écologie (MEDDE) sur les transports intelligents (http://www.transport-intelligent.net/).
2010	Membre du comité d'organisation de la « journée véhicules traceurs », Toulouse, Décembre 2010.
2009-2012	Membre du GERI "véhicules électriques" (Ifsttar).
2008-2012	Membre de la PFI TISIC (Plate-forme intégratrice Inrets « Traitement de l'information, signal, image, connaissance »)

Mon implication dans la collectivité des chercheurs se traduit aussi par mon investissement dans les activités de représentation du personnel par la voie syndicale. Je suis élu du syndicat FSU à la CAP (Commission administrative paritaire) des chercheurs du développement durable depuis 2011. A ce titre, je suis aussi membre de la commission de réforme du ministère de l'écologie pour les CR-DR, du conseil du département COSYS de l'IFSTTAR, du comité consultatif du site de Satory (CCL), et de la commission formation de l'Ifsttar.

2015-2016	<ul style="list-style-type: none">• Membre représentant du personnel au conseil du département COSYS de l'IFSTTAR.• Membre représentant du personnel au comité consultatif local (CCL) du site de Satory.• Membre représentant du personnel à la commission formation de l'Ifsttar.
Depuis 2015	<ul style="list-style-type: none">• Membre de la commission de réforme du ministère de l'écologie (MEEM).
Depuis 2011	<ul style="list-style-type: none">• Elu à la commission administrative paritaire du corps des chercheurs de l'équipement (suppléant) puis du corps des chercheurs du développement durable (titulaire).

3.3. Collaborations nationales et internationales

Depuis ma prise de poste en 2007, j'ai pu mettre à profit mon expertise en modélisation statistique lors de différentes collaborations multidisciplinaires. Ces collaborations sont le fruit d'une trajectoire professionnelle inspirée par les différentes problématiques concrètes que j'ai pu rencontrer. J'ai eu la chance de pouvoir maintenir actives certaines collaborations anciennes tout en ouvrant de nouvelles perspectives. Les collaborations ci-dessous, excepté les plus récentes, se sont toutes traduites par des communications internationales orales ou écrites.

Mes principales collaborations de recherche actives sont détaillées ci-dessous, de la plus récente à la plus ancienne.

3.3.1. Institut géographique national (IGN), Sébastien Mustière (Directeur laboratoire COGIT)

L'aspect géolocalisé des mesures numériques produites par les véhicules modernes m'a conduit à réfléchir aux liens entre la cartographie et la perception de l'environnement par les véhicules. L'idée de fond de cette collaboration récente consiste à mettre au point des méthodes d'enrichissement et de validation de cartes géographiques à partir de données recueillies sur un grand nombre de véhicules « traceurs », enrichies par de l'information géographique (satellitaire par exemple). Dans ce contexte, un accord a été conclu pour co-encadrer le doctorat d'un ingénieur IGN en poste au laboratoire COGIT (Yann Meneroux). Cette collaboration vient de débuter en janvier 2016.

3.3.2. Institut du Véhicule Décarboné et Communicant et de sa Mobilité (VEDECOM), Bertrand Leroy.

Les thèmes de recherche traditionnels de mon laboratoire se sont retrouvés en partie associés aux recherches de l'institut Vedecom. La proximité scientifique et géographique entre le LIVIC et VEDECOM m'a conduit à collaborer plus étroitement avec des experts du véhicule autonome. Mon expérience des données produites par les véhicules autonomes associée aux nouvelles problématiques posées par leur arrivée, m'a conduit à proposer un sujet de recherche, accepté et financé par Vedecom sous la forme d'un contrat de post-doc de 18 mois. Il s'agit de construire une carte enrichie à partir de mesures recueillies par des véhicules instrumentés, et d'évaluer la qualité cette carte ainsi que son utilité comme donnée d'entrée pour un véhicule autonome. Cette collaboration a débuté en décembre 2015.

3.3.3. Virginia Tech. Transportation Institute (VTTI), Dr. Feng Guo.

Dans le cadre de mes recherches sur l'impact des systèmes sur la sécurité routière, je m'intéresse aux comportements des conducteurs observés lors de la conduite naturelle (Naturalistic driving studies, NDS). Le VTTI est un acteur majeur et reconnu dans ce domaine que j'ai eu l'occasion de visiter en tant que participant à une délégation du projet euroFOT en 2009. L'IFSTTAR entretient une collaboration avec cet institut depuis plusieurs années. Dans ce cadre, j'ai animé l'axe NDS pendant 2 ans, ce qui m'a permis d'être invité au 4^{ème} symposium sur la conduite en situation naturelle en 2014.

3.3.4. Laboratoire de Psychologie des Comportements et des mobilités (LPC-IFSTTAR), Patricia Delhomme (DR).

Cette collaboration a débuté en 2011 par le co-encadrement de la thèse de B. Palat avec P. Delhomme. Le facteur humain étant un élément majeur de l'étude de la conduite, il a été tout à fait naturel pour moi de m'intéresser aux actions du conducteur en réaction à certaines contraintes ou choix de navigation. Ce premier travail a permis, grâce à un protocole expérimental adapté, de mettre en évidence les déterminants du choix d'itinéraire en présence de feux tricolores. Ce premier travail ayant été un succès, j'ai poursuivi ma collaboration avec P. Delhomme en lien avec un de mes sujets de recherche propre qui est l'écoconduite. Nous avons ainsi mené conjointement une étude sur le lien entre sécurité et écoconduite (financement DGITM). Par ailleurs, cette collaboration a mené au financement d'un nouveau projet en commun par la fondation MAIF, dont l'objectif est d'étudier comment le traitement statistique d'une masse de données de conduite permet d'améliorer la connaissance et la prévention du risque routier.

3.3.5. Equipe perception et design sonore, Institut de recherche et coordination acoustique/musique (IRCAM), Patrick Susini (DR).

J'ai eu la chance de pouvoir travailler deux mois à l'IRCAM en 2006 avant de rejoindre l'IFSTTAR en 2007. J'ai gardé de ce passage un attrait pour les questions théoriques posées par l'analyse des protocoles expérimentaux spécifiques à la psycho-acoustique, par ailleurs très proches des protocoles de psychologie expérimentale utilisés à l'IFSTTAR-LPC. Le très bon niveau théorique de cette équipe et ses attraits humains l'ont incité à conserver et développer les liens scientifiques que nous avons créés. Je collabore régulièrement avec cette équipe à laquelle j'apporte mon expertise de l'application concrète de la statistique.

3.3.6. Autres collaborations

Par le biais de mes implications répétées dans plusieurs projets européens, j'ai eu la chance de collaborer de nombreuses années avec des chercheurs étrangers. Ces collaborations se font essentiellement dans le cadre de projets, mais la multiplication de ceux-ci conduit en pratique à de longues années de travaux menés en communs. Je travaille ainsi étroitement avec l'université de Leeds (UK, Pr. Oliver Carsten, Pr. Samantha Jamson), et le TNO (Netherlands, Rino Brouwer) depuis presque 8 ans. Nous avons en effet menés ensemble, et menons encore, plusieurs projets européens : euroFOT, ecoDriver et UDrive.


Enfin, depuis 2009 j'entreprends des actions de recherche conjointes entre le CEREMA et l'IFSTTAR. Dans ce cadre, j'ai supervisé plusieurs études d'évaluation de l'impact des systèmes d'aide à la mobilité. C'est donc naturellement que j'ai été amené à collaborer plus étroitement avec le pôle de compétence national d'évaluation des systèmes d'aide au déplacement (PCI-ESAD) situé à Toulouse et dirigé maintenant par L. Khoudour. Par le biais de projets de recherche communs (SCOOP@F), nous joignons nos compétences pour mettre au point des méthodes pertinentes d'évaluation des systèmes coopératifs de transports intelligents.

En 2016, j'ai demandé à rejoindre cette équipe scientifiquement très proche de mes activités. Cette collaboration est donc devenue permanente puisque je suis depuis le 1^{er} mars en poste au Cerema de Toulouse.

3.4. Responsabilités et participations dans des projets de recherche nationaux ou internationaux

3.4.1. Résumé des implications

Dates	Nom	Financement	Rôle	Implication Hommes/Mois	Subvention totale
2017-2018	GameECAR	H2020	WP leader	6	5M€
	Partenaires :	University of Patras, University of LEEDs, BRAINSTORM, Centro Tecnológico de Automoción de Galicia (CTAG), SPARK WORKS, Ifsttar, KITE Solutions S.l.r.			
		NB : J'ai dû mettre un terme à ma participation à ce projet en mars 2017 suite à ma mutation au Cerema qui n'est pas partenaire.			
2017-2018	C-The difference	European Commission	Participant	1	1.3M€
	 Partenaires :	MAPTM, Bordeaux, Helmond, IMTECH, TNO, GEOLOC, Cerema, Blervaque			
2015-2017	MEDOC	Fondation MAIF	Co-responsable scientifique	10	400k€
	 Partenaires :	Ifsttar, Fondation MAIF, ETS			
2014-2019	SCOOP@F	Ministère écologie & Europe	Responsable de tâche	6	10M€
	 Partenaires :	Ministère de l'écologie, Communauté Urbaine de Bordeaux, Conseils généraux des Côtes d'Armor, du Finistère, d'Ille et Vilaine, de l'Isère, Conseil régional de Bretagne, City of St Brieuc, opérateurs de réseaux français, PSA, Renault, Cerema, Ifsttar, GIE RE PSA-Renault, Université de Reims, Telecom Paris Tech, Marben Products, Neavia Technologies			
2013-2015	IDACA	DSCR (Ministère intérieur)	Participant	3	30k€
	Partenaires :	Ifsttar, Ceesar			
2012-2013	Ecoconduite et sécurité en milieu urbain	DSCR (Ministère intérieur)	Responsable scientifique	6	60k€
	Partenaires :	Ifsttar (LPC, LIVIC), CETE IdF			

Dates	Nom	Financement	Rôle	Implication Hommes/Mois	Subvention totale
2012-2016	UDrive 	UE-FP7	Participant	6	8M€
	Partenaires :	Swov, Bast, Centrum Dopravniho, Ceesar, Cidaut, DLR, Ertico, FIA, IBDM, Ifsttar, KFV, LAB, Loughborough, Or Yarok, Safer, TNO, Chemnitz Univ., ITS Leeds, Volvo			
		NB : J'ai dû mettre un terme à ma participation à ce projet en mars 2017 suite à ma mutation au Cerema qui n'est pas partenaire.			
2011-2016	ecoDriver 	UE-FP7	Membre du Core Group, Responsable de sous-projet, Responsable de « work package »	25	14M€
	Partenaires :	ITS Leeds, BMW, TNO, VTI, CTAG, Ifsttar, ERTICO, CRF, HERE, DAIMLER, SIMOTION, TOM TOM TELEMATICS, CARRIERWEB			
2011-2014	Volhand 	ANR	Responsable de tâche	6	618k€
	Partenaires :	LAMIH, CHRU, Fondation Hopale, GIPSA-Lab, INRETS, JTEKT			
2009-2012	PARTAGE 	ANR	Participant	4	417k€
	Partenaires :	IRCCyN, Ifsttar Univ. Evry, Univ. Rennes 2, OKTAL, Renault			
2009-2012	ABV 	ANR	Participant	3	2M€
	Partenaires :	INRETS, Continental automotive, IBISC, GMconseil, IEF, INDUCT, INRIA, LAMIH, VIAMETRIS, MIPS, VERI			
2008-2012	euroFOT 	UE-FP7	Responsable de « work package » Responsable scientifique Ifsttar	15	14M€
	Partenaires :	Ford, BMW, Daimler, CRF, MAN, Volvo, Volvo Tech., Volkswagen, Audi, Bosch, Continental, Delphi, Harman International, Chalmers, ICCS, IKA,			

Dates	Nom	Financement	Rôle	Implication Hommes/Mois	Subvention totale
		IZVW, Politecnico di Torino, ITS Leeds, Bast, Ceesar, CTAG, Ifsttar, TNO, Allianz, ADAS management consulting, Alcor, EICT, ERTICO			
2008-2011	CARVAL	AFM (Association Française Myopathie)	Participant	3	200k€
	Partenaires :	CEREMH, Ifsttar, LISV-UVSQ			
2008-2011	Infracall	FUI, Direction Générale des Entreprises, Conseil Général des Yvelines	Participant	2	620k€
	Partenaires :	AXIMUM, LCPC (LIVIC), AXIMUM Produits Electroniques, Bouygues Telecom, PSA Peugeot-Citroën, Conseil Général des Côtes d'Armor, Groupe SANEF (SAPN)			

3.4.2. Détails des responsabilités

- Responsable du work package “Testing & Evaluation” du Projet **GameECAR**.
Gamification of EcoDriving Behaviours through Intelligent Management of dynamic car and driver information (H2020)
 - Ce projet a pour objectif de développer et tester des solutions innovantes pour encourager les conducteurs à adopter une conduite éco-responsable, par le biais de « jeux sérieux » (serious games). Il s’agira d’encourager les économies d’énergies au volant en usant de ressorts ludiques sans pour autant dégrader la sécurité de conduite. Les logiciels développés s’adapteront de manière intelligente à la personnalité de l’utilisateur et s’appuieront sur l’expérience de conduite en temps réel pour encourager aux comportements vertueux.
 - Coordinateur des expériences internationales en conditions réelles visant à évaluer l’efficacité de l’approche proposée.

- Responsable du work package “impact assessment” du projet **SCOOP@F**
Projet de déploiement pilote de systèmes de transport intelligents coopératifs (FR & UE, 2014-2019).
 - Ce projet très ambitieux, poussé par l’état français, est le premier projet de déploiement pilote de systèmes de transport intelligents coopératifs. Ces systèmes ont pour objectif d’améliorer la sécurité des usagers, faciliter leurs déplacements et garantir la sécurité des agents qui interviennent sur les routes.
 - Coordination de l’activité d’évaluation de l’impact d’une flotte de véhicules connectés sur le trafic, la sécurité, et l’environnement.

- Co-Coordinateur du projet **MEDOC** avec P. Delhomme.
« MassE de Données de Conduite pour modélisation d’un feed-back incitant à la conduite sécuritaire selon les déplacements des automobilistes » (Fondation MAIF, 400k€, 2015-2017).
 - En collaboration avec Patricia Delhomme (Laboratoire de Psychologie des Comportements et des mobilités), nous menons une recherche sur les accidents évités au volant avec une expérimentation ambitieuse impliquant 200 conducteurs.
 - Je coordonne l’analyse des données des véhicules et mets au point les méthodes de détection des évènements à risque de la conduite.

- Leader de Sous-projet et membre du Core-Group du projet européen **ecoDriver**
« *Supporting the driver in conserving energy and reducing emissions* » (IP-FP7-14M€-2011-2016).
 - Ce projet a pour but de concevoir et tester à grande échelle différents systèmes d'aide à l'écoconduite. L'Ifsttar a conçu puis testée une application pour téléphones android.
 - Responsable du sous-projet « evaluation of effectiveness », et du work package « emissions ».
 - Membre du core-group et coordination des recherches d'une dizaine de partenaires internationaux (industriels+académiques).

- Co-coordonateur étude **IDACA** avec H. Tattegrain.
« *Evaluation de l'impact des dispositifs d'assistance à la conduite automobile (réglementaires ou optionnels) : Analyse des possibilités de généralisation des méthodes d'évaluation des dispositifs* » (financement DSCR, 2013-2015).
 - Inspirés de nos travaux respectifs dans les projets européens de type FOT ou NDS, nous avons proposé à la DSCR de synthétiser dans un rapport la méthodologie d'analyse adaptée à ces nouveaux protocoles.

- Coordinateur du projet « **Ecoconduite et sécurité en milieu urbain** »
(Financement DSCR, 2012-2013).
 - L'objet de cette étude était d'appréhender l'impact de l'écoconduite sur la sécurité par une approche conjointe entre sciences humaines et statistiques utilisant une expérimentation réelle.
 - Coordination d'une recherche expérimentale conjointe entre l'Ifsttar (LIVIC-LPC) et le CETE IdF.

- Responsable de tâches, projet ANR **Volhand**
"VOLant pour personne âgée et/ou HANDicapée : Direction Assistée Electrique Personnalisée adaptée au conducteur à mobilité réduite" (ANR, 2011-2014).
 - La nouvelle loi de commande du volant, adaptée à certaines déficiences, a été testée en conditions réelles via un protocole adapté, conçu, et analysé, par l'Ifsttar.
 - Responsable du livrable « Définition du protocole expérimental des essais sur pistes ».
 - Responsable du livrable « Tests du véhicule automobile préparé ».
 - Responsable du livrable « Tests à petite échelle des essais sur pistes ».
 - Responsable du livrable « Analyse et évaluation des essais sur pistes ».

- Responsable scientifique (Principal investigator, PI) de l'Ifsttar pour le projet européen EUROFOT « euroFOT : European Field Operational Test on Active Safety System » (IP-FP7-15M€-2008-2012).
 - Le projet euroFOT est le premier à mettre en œuvre une expérimentation réelle à grande échelle pour évaluer l'impact sur la conduite de 8 systèmes embarqués différents. L'ifsttar et le Ceesar avaient la charge d'évaluer le limiteur et le régulateur de vitesse. J'ai eu la responsabilité de cette analyse.
 - Responsable du work package « *piloting* » (coordination de l'ensemble des sites pilotes européens) et du livrable associé.

3.4.3. Détails des participations

- Projet C-The Difference.
 - Il s'agit d'un projet pilote en réponse à l'appel d'offre "Beyond traffic jams: intelligent integrated transport solutions for road infrastructure" de la commission Européenne. Ce projet s'appuie sur les infrastructures et véhicules communicants déjà déployés dans les villes pionnières que sont Bordeaux et Helmund (Netherlands) et a pour ambition de mettre au point une méthodologie commune d'évaluation de l'impact de ces technologies et des bénéfices associés à leur mise en œuvre à grande échelle.
 - *Implication personnelle* : Je participe à la mise au point de la méthode d'évaluation des C-ITS en milieu urbain, tenant compte de la sécurité, l'efficacité, et l'environnement (WP4 « Impact assessment »).
- Projet européen UDrive
 « *eUropean naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment* » (FP7-2012-2016).
 - Ce projet est la première étude européenne à large échelle de la conduite en situation naturelle. Sans aucun système particulier étudié, il s'agit d'étudier le comportement de conduite en conditions réelles en limitant les biais expérimentaux. Sur la base des résultats obtenus, l'IFSTTAR est responsable des recommandations à mettre en œuvre pour améliorer la sécurité routière et réduire les comportements responsables des émissions de GES.
 - *Implication personnelle* : Tâche « Evaluation des facteurs de risque des presque-accidents et accidents ».
 - NB : J'ai dû mettre un terme à ma participation à ce projet en mars 2017 suite à ma mutation au Cerema qui n'est pas partenaire.

- **Projet ANR PARTAGE**

« Contrôle partagé entre conducteur et assistance à la conduite automobile pour une trajectoire sécurisée » (ANR, 2009-2012).

- Ce projet a contribué à rendre les voitures plus sûres et ergonomiques, en proposant des automates d'assistance au contrôle de la trajectoire, pour prévenir les sorties involontaires de voie, en virage ou en ligne droite.
- *Implication personnelle* : Tâche 4.1 « analyse des trajectoires réelles ».

- **Projet ANR ABV**

« Automatisation Basse Vitesse de la Conduite » (ANR, 2009-2012).

- Ce projet a consisté à mettre au point un véhicule capable de conduire en autonomie complète à basse vitesse (<50km/h), ainsi qu'à procéder à l'évaluation de ses performances en conditions réelles et virtuelles.
- *Implication personnelle* : Tâche 8 « impact sur le trafic et la pollution ».

- **Projet CARVAL**

« Accessibilité des personnes handicapées (myopathes tétraplégiques) à la conduite », financement AFM (Association française contre les myopathies, 2008-2011).

- *Implication personnelle* : « Conception d'un véhicule pour le handicap et l'évaluation », Tâche « Évaluation de la conduite automobile de personnes handicapées ».

- **Projet InfraCall**

« Infrastructure Communicante pour une Route sécurisée », financement FUI (Fonds unique ministériel, Direction Générale des Entreprises, Conseil Général des Yvelines, 2008-2011).

- L'objectif était de développer un dispositif capable de détecter un impact sur un équipement d'infrastructure routière et d'émettre un message d'alerte, contenant des informations contextuelles, à destination du gestionnaire.
- *Implication personnelle* : Coordination du livrable « étude d'accidentologie dans les Yvelines ». Cette étude coût/bénéfice évalue l'opportunité d'installer le système en fonction du coût du système et des vies qu'il pourrait épargner.

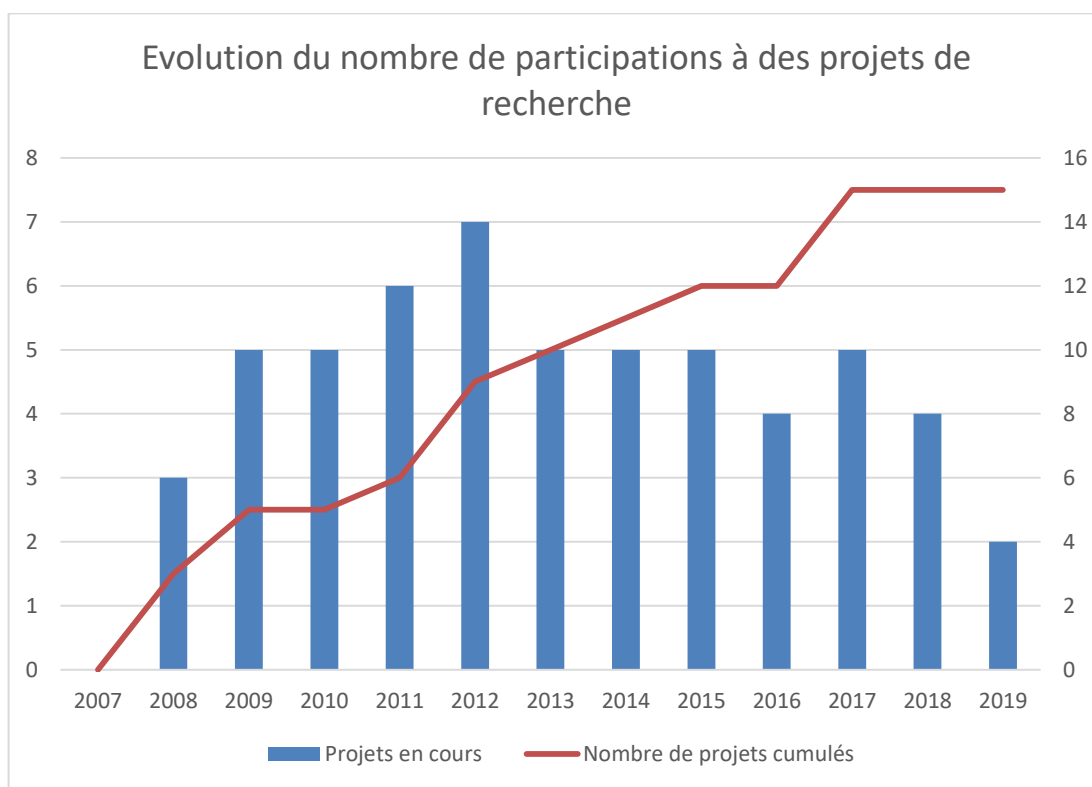


Figure 1 : Evolution du nombre de participations à des projets de recherche

4. Encadrements

4.1. Encadrement de doctorants

4.1.1. Résumé des encadrements de doctorants

<i>Dates</i>	<i>Taux d'encadrement</i>	<i>Etudiant (financement)</i>	<i>Détails</i>
2016-2019	50%	Yann Meneroux (Ingénieur IGN)	Université Paris Est, école doctorale MSTIC : « <i>Utilisation des véhicules traceurs et des données géographiques pour la construction d'une infrastructure routière numérique</i> ». Sous la direction de Sébastien Mustière, co-encadrement avec Arnaud Le-Guilcher (laboratoire COGIT-IGN) et O. Orfila (Ifsttar).
2012-2016	25%	Camila Freitas-Salguereido (Cifre Renault)	Université d'Evry, école doctorale "science et ingénierie" : « <i>Biomimétisme et véhicule décarboné : Génération de concepts innovants bio-inspirés à partir de la méthode C-K</i> ». Sous la direction de Sébastien Glaser (IDTPE IFSTTAR), co-encadrement avec P. Doublet (Renault), O. Orfila (LIVIC-IFSTTAR), S. Doncieux (ISIR-UPMC).
2010-2013	30%	Blazej Palat (Inrets)	Université de Paris-8, école doctorale N° 224 Cognition, langage, interaction : « <i>Déterminants du comportement des usagers de la route et prévention des conduites à risque</i> » Sous la direction de Patricia Delhomme (DR1 directrice du laboratoire de psychologie de la conduite à l'IFSTTAR).
2010-2013	50%	Cindie Andrieu (LCPC)	Université Paul Sabatier (Toulouse III), école Doctorale Mathématiques, Informatique, Télécommunications de Toulouse [ED MITT 475] : « <i>Modélisation statistique d'un profil de vitesse de référence pour véhicules intelligents</i> » Sous la direction du Pr. Xavier Bressaud (UPS Toulouse III).

4.1.2. Détails des encadrements de doctorants

Date :	Janvier 2016 - maintenant
Doctorant :	Yann Meneroux.
Directeur de thèse :	Sébastien Mustière (IGN).
Encadrement :	G. Saint Pierre (50%) , Arnaud Le-Guilcher (IGN, 20%), O. Orfila (Ifsttar, 20%), S. Mustière (IGN, 10%).
Sujet :	Utilisation des véhicules traceurs et des données géographiques pour la construction d'une infrastructure routière numérique.
Financement :	IGN (l'étudiant est ingénieur IGN titulaire).
Ecole doctorale :	MSTIC, Université Paris Est.
Résumé :	<p>Afin d'optimiser l'usage du réseau de transport et accroître son efficacité, il convient d'acquérir une connaissance détaillée des caractéristiques du réseau routier et de son usage. Cette connaissance est pour l'instant uniquement disponible dans les cartes numériques embarquées dans nos GPS. Ces cartes sont devenues indispensables à la conduite, mais renseignent encore assez peu sur l'infrastructure elle-même (feux rouges, passages piétons, zones de rencontre, pistes cyclables, etc ...), et contiennent parfois une information erronée (limitations de vitesse).</p> <p>Un des objectifs de ce travail est de proposer une méthode pour créer des cartes à partir de traces de véhicules, suffisamment précises pour permettre d'alimenter les « horizons électroniques » de nos futurs véhicules autonomes ou connectés. Ces horizons sont très utiles pour aider les systèmes embarqués à aider le conducteur à mieux anticiper sa conduite, mais aussi pour les futurs véhicules dont la conduite sera partiellement ou entièrement automatisée. En effet, cet "horizon électronique" permet la modélisation des contraintes imposées par l'infrastructure à la conduite, réelle ou autonome. Le premier objectif est d'obtenir une carte digitale de la route comportant notamment la position d'éléments d'infrastructure fins comme les feux rouges, les passages piétons, les cédez-le-passage, ou les dos-d'âne.</p>
Date de soutenance prévue :	2019.
Valorisation :	1 ACTI, 1 ACL en cours d'évaluation
Position actuelle :	Ingénieur IGN.

Date :	Mai 2012 – novembre 2013
Doctorant :	Camila Freitas-Salguereido.
Directeur de thèse :	S. Glaser et S. Doncieux.
Encadrement :	S. Glaser (Vedecom, 25%), S. Doncieux (ISIR-UPMC, 25%), G. Saint Pierre (25%) , O. Orfila (Ifsttar, 25%).
Sujet :	Biomimétisme et véhicule décarboné : Génération de concepts innovants bio-inspirés à partir de la méthode C-K.
Financement :	Cifre Renault.
Ecole doctorale :	Sciences et Technologies de l'Information et de la communication STIC, UEVE - Université Évry Val d'Essonne.
Résumé :	Dans le contexte actuel, le secteur des transports a pour défi d'atténuer les effets du changement climatique, d'améliorer la qualité de l'air dans les centres urbains et de réduire la dépendance au pétrole. Dans cette perspective, la recherche se consacre à réduire les émissions polluantes et à mettre au point l'exploitation d'énergies alternatives. Pour apporter une réponse rapide et significative à ces défis, des percées technologiques sont nécessaires. Les systèmes vivants sont une source d'inspiration jusqu'ici peu explorée pour innover et pour améliorer des technologies conventionnelles. Cette démarche « d'inspiration du vivant », appelée également biomimétisme, a déjà proposé des concepts novateurs comme les films antimicrobiens inspirés de la peau des requins et les algorithmes d'optimisation basés sur le comportement des colonies de fourmis. Cette thèse vise à utiliser l'approche biomimétique dans le projet du véhicule décarboné. Son objectif est d'étudier la génération de nouveaux concepts et de nouvelles solutions bio-inspirées afin de réduire les émissions de dioxyde de carbone des véhicules sur l'ensemble de leur cycle de vie. La « théorie C-K de la conception » appuiera la modélisation du processus de bio-inspiration et la génération des concepts innovants bio-inspirés.
Date de soutenance prévue :	25 octobre 2016.
Valorisation :	1 ACL, 1 ACTI.
Position actuelle :	Chargée de mission au ministère de l'économie et des finances depuis le 1 ^{er} avril 2015.

Date :	Novembre 2010 – novembre 2013
Doctorant :	Blazej Palat.
Directeur de thèse :	Patricia Delhomme.
Encadrement :	Patricia Delhomme (Ifsttar, 70%), G. Saint Pierre (30%) .
Sujet :	Etudes déclaratives et comportementales (simulateur) de la prise de décision aux feux tricolores : interactions entre facteurs socio-psychologiques et contextuels.
Financement :	IFSTTAR.
Ecole doctorale :	Cognition, Langage, Interaction, université Paris VIII – Vincennes – Saint Denis.
Résumé :	<p>Les franchissements des feux orange et rouge constituent une cause fréquente d'accidents aux intersections. Ces comportements sont plus ou moins délibérés et influencés par différents contextes. L'objectif de la thèse a été d'étudier les facteurs intervenant dans les processus de décision aux feux tricolores. Il s'est agi d'examiner comment les automobilistes arbitrent entre la sécurité, le respect de la réglementation et la réduction du temps à attendre. Cette thèse comprend une introduction suivie de deux parties. La première partie expose les théories de la motivation, de la prise de décision, de la perception du temps et de l'espace qui constituent les références pour les recherches empiriques. Elle se termine en livrant la problématique générale de la thèse. La deuxième partie comprend quatre études. La première, basée sur la théorie du comportement planifié (Ajzen, 1985), examine les prédicteurs de l'intention de franchir le feu orange. Les études suivantes, menées sur simulateur de conduite, observent l'influence (étude 2) des contextes social et personnel sur le comportement des automobilistes aux feux tricolores, (étude 3) des arrêts aux feux sur la perception de la durée d'un trajet ainsi que les conséquences de cette perception sur le comportement des automobilistes aux feux tricolores et sur le choix de trajet. La quatrième étude s'intéresse à l'utilisation de l'heuristique de numérosité dans le choix entre deux trajets selon le nombre de feux tricolores et la probabilité d'avoir à s'arrêter aux feux. Pour clore ce travail, les résultats obtenus sont discutés et des limites aux recherches réalisées sont posées avant d'aboutir à la conclusion finale de la thèse.</p>
Date de soutenance :	25 novembre 2013.
Valorisation :	1 ACL, 1 ACTN (Il existe d'autres valorisations de cette thèse dans le domaine de la psychologie auxquelles je n'ai pas participé).
Position actuelle :	Chercheur contractuel au laboratoire de psychologie des comportements et des mobilités (LPC) de l'Ifsttar.

Date :	Septembre 2010 – septembre 2013
Doctorant :	Cindie Andrieu.
Directeur de thèse :	Xavier Bressaud.
Encadrement :	G. Saint Pierre (50%), X. Bressaud (Toulouse III, 50%).
Sujet :	Modélisation statistique d'un profil de vitesse de référence pour véhicules intelligents.
Financement :	IFSTTAR
Ecole doctorale :	Mathématiques, Informatique, Télécommunications de Toulouse [ED MITT 475]
Résumé :	<p>La connaissance des vitesses pratiquées est une caractéristique essentielle du comportement des conducteurs et de leur usage du réseau routier. Cette information est rendue disponible grâce à la généralisation des véhicules connectés, mais aussi des smartphones, qui permettent d'accroître le nombre de "traceurs" susceptibles de renvoyer leur position et leur vitesse en temps réel. Dans cette thèse, nous proposons d'utiliser ces traces numériques et de développer une méthodologie, fondée sur une approche fonctionnelle, permettant d'extraire divers profils de vitesse caractéristiques. Dans une première partie, nous proposons une modélisation fonctionnelle des profils spatiaux de vitesse (i.e. vitesse vs distance parcourue) et nous étudions leurs propriétés (continuité, dérivabilité). Dans une seconde partie, nous proposons une méthodologie permettant de construire un estimateur d'un profil spatial de vitesse à partir de mesures bruitées de position et de vitesse, fondée sur les splines de lissage et la théorie des espaces de Hilbert à noyau reproduisant (RKHS). Enfin, la troisième partie est consacrée à la construction de divers profils agrégés (moyen, médian). Nous proposons notamment un alignement des profils par "landmarks" au niveau des arrêts, puis nous proposons la construction d'enveloppes de vitesse reflétant la dispersion des vitesses pratiquées.</p>
Date de soutenance :	24 Septembre 2013.
Valorisation :	1 récompense de conférence internationale, 1 prix national, 2 ACTI, 2 ACTN, 2 Arxiv.
Position actuelle :	Chercheur en statistiques chez SAFETY LINE (gestion des risques dans le domaine du transport aérien).

4.1.3. Doctorants récompensés

<i>Date</i>	<i>Nom</i>	<i>Récompense obtenue</i>
2012	Cindie Andrieu	Laureat du "Prix de la chaire Abertis France" décerné par le concessionnaire autoroutier espagnol Abertis, l'école des Ponts ParisTech et l'IFSTTAR. Ce prix récompense des travaux dans le domaine de la gestion des infrastructures de transport.
2012	Cindie Andrieu	Médaille de bronze de la compétition "Young European Arena of Research 2012" (YEAR 2012) organisée dans le cadre de la conférence "Transport Research Arena" (TRA 2012), pilier "Transport, mobility and infrastructure", Athènes, Grèce, avril 2012.

4.2. Encadrements de post-doctorants

4.2.1. Résumé des encadrements de post-doctorants

<i>Dates</i>	<i>Nom</i>	<i>Sujet / Financement / durée</i>
2016- 2017	Mohammad Ghasemi Hamed	<ul style="list-style-type: none"> • « Modèles aléatoires pour traces numériques de véhicules ». • Ingénieur de recherche Ifsttar. • Janvier 2016 - décembre 2017.
2015- 2016	Alexandre Lekina	<ul style="list-style-type: none"> • « Utilisation des véhicules traceurs pour le diagnostic d'itinéraire et la construction de cartes dynamiques ». • Financé par l'institut du Véhicule décarbonné et communicant et de sa mobilité (VEDECOM). • Décembre 2015 - juin 2017.
2015- 2016	Xinran Li	<ul style="list-style-type: none"> • « Méthodes statistiques de détection d'événements à risque pour analyser le comportement d'un panel de conducteurs ». • Projet MEDOC (financé par la Fondation MAIF). • Septembre 2015 - décembre 2016.

4.2.2. Détails des encadrements de post-doctorants

Nom :	Mohammad Ghasemi Hamed.
Encadrement :	G. Saint Pierre (80%), O. Orfila (10%), D. Gruyer (10%).
Sujet :	Modèles aléatoires pour traces numériques de véhicules.
Financement :	Ifsttar
Mission :	<p>Ce docteur en intelligence artificielle vient d'être recruté au LIVIC afin de pallier temporairement au départ récent d'un ingénieur de recherche titulaire. Son rôle est d'apporter, sous ma coordination, une expertise de haut niveau sur les méthodes récentes d'apprentissage statistique utilisées tout autant en traitement d'image que pour l'analyse de données classique. L'application de ces méthodes (réseaux neuronaux, forêts aléatoires, modèles de Markov cachés, ...) aux données générées par les véhicules intelligents, représente un challenge constitutif d'une grande part de mon dernier projet de recherche.</p> <p>M. Ghasemi mènera une revue de littérature et explorera, en collaboration avec mes autres post-doctorants, les performances des méthodes mes plus récentes sur l'apprentissage supervisé de l'infrastructure.</p>
Date :	Janvier 2016 – janvier 2017

Nom :	Alexandre Lekina.
Encadrement :	G. Saint Pierre (90%), B. Leroy (10%).
Sujet :	Utilisation des véhicules traceurs pour le diagnostic d'itinéraire et la construction de cartes dynamiques.
Financement :	Vedecom
Mission :	<p>Ce docteur en statistique appliqué vient d'être recruté par l'institut Vedecom sur ma proposition, afin de construire une méthode permettant d'extraire de l'information utile aux futurs véhicules autonomes, à partir des données générées par des véhicules instrumentés non autonomes. Il s'agira d'acquérir une connaissance détaillée des caractéristiques du réseau routier et de son usage grâce à la collecte des profils de vitesse longitudinaux sur un même trajet (données que je fournis). A. Lekina a pour mission de proposer une méthode d'extraction de l'information concernant l'infrastructure à l'aide de méthodes statistiques afin de :</p> <ul style="list-style-type: none"> • détecter a posteriori l'infrastructure (passages piétons, feux rouges, etc.); • repérer les événements « risqués » de la conduite comme les freinages brusques ou les excès de vitesse; <p>Cela permettra ensuite de produire une carte enrichie d'un itinéraire à partir des données flottantes. La finalité à long terme étant d'évaluer la qualité de la carte enrichie ainsi que son utilité comme donnée d'entrée pour un véhicule autonome.</p>
Date :	Janvier 2016 – janvier 2017

Nom :	Xinran Li
Encadrement :	G. Saint Pierre (100%).
Sujet :	Méthodes statistiques de détection d'événements à risque pour analyser le comportement d'un panel de conducteurs.
Financement :	Projet MEDOC, fondation MAIF
Mission :	<p>Le candidat recruté est en charge de l'analyse statistique des résultats d'une expérimentation à grande échelle menée en collaboration avec Patricia Delhomme, directrice de recherche au laboratoire de Psychologie des comportements et des mobilités. Ce projet consiste à enregistrer la dynamique de 200 véhicules personnels de sociétaires MAIF lors de leurs déplacements personnels quotidiens, pendant deux mois chacun. Ces données anonymisées contiennent, pour chaque seconde, des informations sur les variables déterminantes de la dynamique du véhicule : vitesse, accélération, position et cap GPS, ainsi que d'autres paramètres issus de véhicule (régime moteur, frein, etc.). L'évolution récente des algorithmes adaptés à ce type de données, permet d'identifier des événements de la conduite ayant un lien avec la sécurité routière (freinages brusques, sorties de routes, coups de volants, etc.). Fort heureusement ces événements génèrent rarement des accidents. Il est cependant avéré que la fréquence d'occurrence de ceux-ci est corrélée avec le taux d'accident réel. Un des objectifs du projet MEDOC est d'approfondir l'étude de ce lien, en améliorant la capacité de détection des algorithmes actuellement disponible. En effet, le taux de bonne détection de ces "incidents" est très faible, et une grande partie des événements candidats ne sont en réalité pas représentatifs du risque routier. Les conducteurs participants à l'étude auront à charge de remplir un carnet de bord à chaque survenue d'un événement "risqué" survenu lors de leur conduite. Ils auront par ailleurs un bouton poussoir à disposition, leur permettant d'ajouter un marqueur numérique aux données signalant cet événement. La connaissance de cette "vérité terrain" permettra d'améliorer les capacités de détection des algorithmes utilisés. Les méthodes utilisées vont des modèles mixtes, à l'analyse fonctionnelle de courbes, en passant par l'ensemble de la panoplie usuelle de l'analyse de données (tests, modèles linéaires, analyse multivariée etc.).</p>
Date :	Septembre 2015 – décembre 2016

4.3. Encadrement de stages (Master 2 recherche et école d'ingénieurs)

J'encadre très régulièrement des élèves stagiaires de Master 2 recherchant une expertise en statistique appliquée. Cela me permet de remplir une mission essentielle du chercheur : la transmission du savoir.

<i>Dates</i>	<i>Nom</i>	<i>Sujet</i>
2015	A. Varfi (Master 2 UPMC)	« Statistical evaluation of the impact of a support system for safe and efficient driving: projet ecoDriver », mai-sept 2015.
2015	G. Leno (Master 2 UPMC)	« Analyse statistique et localisation des zones accidentogènes dans les Yvelines », projet MEDOC, juillet-sept. 2015.
2014	L. Ali (Master 2)	« Effet de la pratique de l'éco-conduite sur les phases d'accélération et de décélération », étude écoconduite et sécurité, avril-Sept 2014.
2013	W. Bounliphone (Master 2)	« Méthodes d'évaluation du risque routier induit par l'utilisation des systèmes d'aide à la conduite », avril-Sept 2013.
2012	A.-M. Alexandrescu (Master 2)	« Méthodes d'estimation du risque routier », projet euroFOT, avril-Sept 2012.
2011	O. Sok (Master 2)	« Analyse statistique de trajectoires et prédiction de sortie de voie », projet PARTAGE, avril-Sept 2011.
2011	J. Zho (Master 2)	« Évaluation de la conduite automobile de personnes handicapées », projet Carval, avril-sept 2011. Co-encadrement avec Francis Dupin (IR).
2011	Projet étudiant Centrale Paris	« développer, tester et valider, une aide à l'écoconduite destinée à équiper un smartphone fonctionnant sous « android » », avril-Sept 2011.
2010	P.-E. Caute (Master 2)	« Modélisation statistique de trajectoires de véhicules en virage », avril-sept. 2010.
2010	M. L. Pham (Master 2)	« Détection d'événements de conduite et analyse d'impact du système LAVIA par modèles marginaux », avril-sept. 2010.
2009	C. Andrieu (Master 2)	« Analyse statistique d'une expérience d'étude de l'éco-conduite : vers la conception d'un « éco index » », avril-sept. 2009.

4.4. Participation à des jurys et à des instances ou comités en lien avec l'enseignement.

<i>Dates</i>	<i>Description</i>
2015	Membre du jury du concours d'ingénieur d'étude pour le laboratoire SPLOTT (laboratoire Systèmes Productifs, Logistique, Organisation des Transports, et Travail), AME-IFSTTAR.
2011	Examinateur de la thèse de Duy-Hung HA. <ul style="list-style-type: none"> • « Variations of the Parameters of the Time Headway Models according to Macroscopic Variables, Fundamental Diagram and Exogenous Effects ». • Février 2011, IFSTTAR, France.
2010	Rapporteur de la thèse de Grégoire Sébastien Larue. <ul style="list-style-type: none"> • « Predicting effects of monotony on driver's vigilance » • Avril 2010, Centre for Accident Research & Road Safety – Queensland University, Australie.
2009	Rapporteur de la thèse de Samantha Chen. <ul style="list-style-type: none"> • « mining patterns and factors contributing to crash severity on road curves » • Août 2009, Centre for Accident Research & Road Safety – Queensland University, Australie.

5. Publications et communications

5.1. Résumé des publications

	Code	Nombre total	Nombre depuis 4 ans
Article dans revue à comité de lecture répertoriée dans les BDI JCR, Scopus, ERIH, HCRES	ACL	11	9
Article dans revue à comité de lecture non répertoriée dans les BDI JCR, Scopus, ERIH, AEREL	ACLN	1	
Article dans revue sans comité de lecture	ASCL		
Direction d'ouvrage	DO		
Ouvrage scientifique ou chapitre	OS	3	1
Conférence invitée dans un congrès international ou national (INV) ¹	INV	7	7
Communication avec actes dans un congrès international	ACTI	22	11
Communication avec actes dans un congrès national	ACTN	5	3
Communication orale sans actes dans un congrès international ou national	COM		
Communication par affiche dans un congrès international ou national	AFF		
Rapport de recherche	RAPP	18	12
Rapport d'expertise	RAPP-EX		
Ouvrage de vulgarisation ou chapitre	OV		

¹ Il s'agit de conférences invitées au sens de frais d'inscription et de voyages offerts, ou de conférences (semi)plénières.

Tableau 1 : Résumé des publications par année

	≤ 2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
OS		1		1						1
ACL		1	1				2	2	4	1
ACLN	1									
ACTI	4	5			2	4	1	4	2	
ACTN	1			1		2	1			
RAPP	5			1		4		4	3	1
INV						3	1	2	1	
Total	11	7	1	3	2	13	5	12	9	

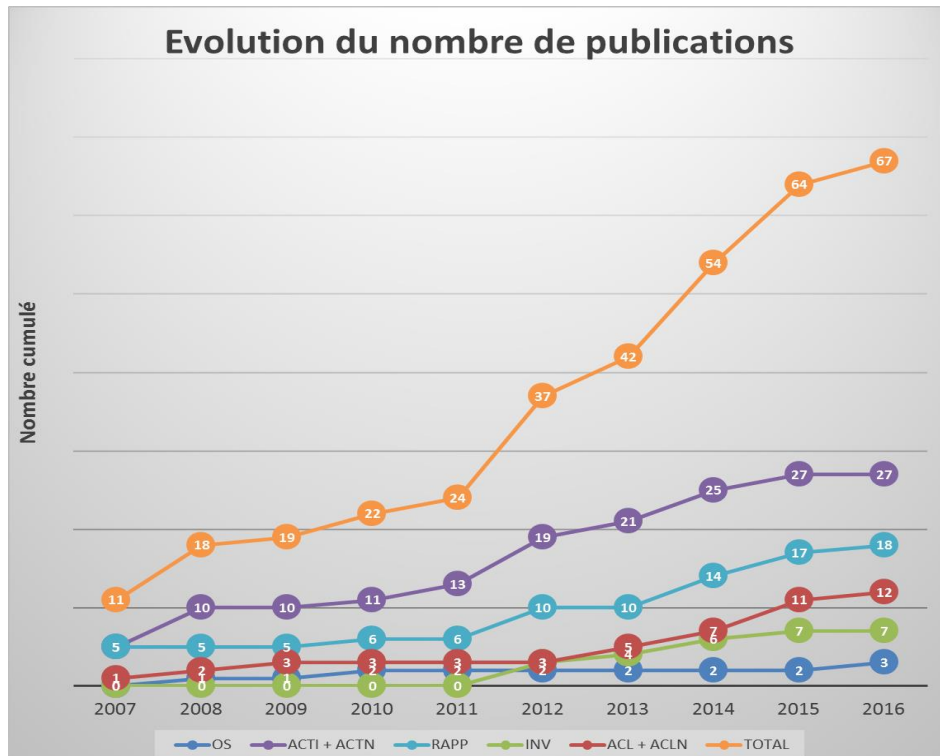


Figure 2 : Evolution du cumul de publications

5.2. Présentation de 5 publications significatives

Ouvrage scientifique (OS)

1. **Saint Pierre G.** (2010), « *Le système LAVIA et la consommation de carburant : méthodes exploratoires pour données de conduite en situation naturelle* », collection "études et recherches du laboratoire des ponts et chaussées", CR54, 120 pages.
 - *Cet ouvrage est le fruit de mes premiers travaux visant à exploiter statistiquement les données de véhicules instrumentés pour évaluer l'impact d'un système embarqué. Les mesures d'un véhicule sont géolocalisées, horodatées, et enrichies par de l'information contextuelle. Cet ouvrage présente et détaille la mise en œuvre d'une analyse statistique adaptée à ce type de données sur un grand nombre de trajets enregistrés lors du projet LAVIA (limiteur s'adaptant à la vitesse autorisée).*

Articles dans des revues à comité de lecture répertoriées (ACL)

2. Susini P., Houix O., **Saint Pierre G.** (2015), « *The Effect of Loudness on the Perceptual Representation of Sounds With Similar Timbre* », Acta Acoustica United With Acustica Vol. 101, N. 6, pp. 1174-1184(11), DOI:10.1016/j.trc.2015.04.026.
 - *Cette publication récente représente bien la collaboration scientifique que je poursuis avec l'IRCAM depuis mon passage en 2006. J'apporte mon soutien pour l'application de modèles mixtes aux protocoles de mesures répétées qui sont*

utilisés dans cette discipline. Dans ce travail, j'interprète et explique l'apparition de certains phénomènes numériques dans les analyses par MDS (Multidimensional scaling), comme l'effet « fer à cheval ».

3. Orfila O., **Saint Pierre G.**, Messias M., (2015), « *An Android Based Ecodriving Assistance System to Improve Safety and Efficiency of Internal Combustion Engine Passenger Cars* », Transportation research part C: Emerging technologies, Volume 58, Part D, Pages 772–782.
 - *Cette publication représente l'aboutissement de plusieurs années de recherche afin de mettre en œuvre concrètement l'application de méthodes statistiques d'estimation du comportement d'écoconduite. Ce travail décrit donc le fonctionnement d'une application pour téléphone android d'aide à l'écoconduite. Les effets de celle-ci sont analysés actuellement dans le cadre du projet ecoDriver que je co-coordonne.*

4. Andrieu C., **Saint Pierre G.** (2014), « *Evaluation of ecodriving performances and teaching method: comparing training and simple advices* », European Journal of Transport and Infrastructure Research, 14(3), pp. 201-213, ISSN: 1567-7141.
 - *Un de mes thèmes de recherche consiste à étudier les comportements ou systèmes permettant une réduction des émissions de polluants. Cette publication présente l'application de modèles de régression logistique à la détection de la mise en œuvre des règles d'or de l'écoconduite. Ces modèles sont ensuite utilisés pour quantifier la bonne mise en œuvre de ces consignes dans différents cas concrets.*

5. Ndong Ngu'ema E.-P., **Saint-Pierre G.** (2008), « *Model-Based Classification with Dissimilarities: a Maximum Likelihood Approach* », Formal Pattern Analysis & Applications 09/2008; 11(3):281-298. DOI:10.1007/s10044-008-0105-2.
 - *Cette publication représente les travaux de modélisation statistiques que je menais avant mon affectation à l'Ifsttar. Il s'agit ici d'un modèle de classification adapté aux données de dissimilarités (ou distances). J'utilise actuellement à nouveau ce type d'algorithmes afin de proposer des méthodes d'apprentissage de l'infrastructure routière à partir des données issues de véhicules.*

5.3. Liste des publications et communications :

5.3.1. Ouvrages scientifiques (OS)

Ouvrage

1. **Saint Pierre G.** (2010), « *Le système LAVIA et la consommation de carburant : méthodes exploratoires pour données de conduite en situation naturelle* », collection "études et recherches du laboratoire des ponts et chaussées", CR54, 120 pages.

Chapitre d'ouvrage

1. **Saint Pierre G.**, Ledoux V., Serre T., Violette E. (2017), « *La mobilité numérique au service de la sécurité routière* », publié dans "évaluation des politiques de sécurité routière, methods, outils, limites", sous la direction de Gilles Blanchard et Laurent Carnis, éditions L'Harmattan.
2. Vrignon J., Rakotonirainy A., Gruyer D., **Saint Pierre G.** (2008), « *The impact of subjective factors on driver vigilance: a driving simulator study* », in "Driver Behavior and training" vol III, chapitre 29, Ashgate publishing.

5.3.2. Articles dans des revues à comité de lecture répertoriées (ACL)

3. Freitas Salgueiredo C., Orfila O., **Saint Pierre G.**, Doublet P., Glaser S., Doncieux S., Billat V. L. (2017), « *Experimental testing and simulations of speed variations impact on fuel consumption of conventional gasoline passenger cars* », à paraître dans Transportation Research Part D: Transport and Environment.
4. Orfila O., Freitas Salgueiredo C., **Saint Pierre G.**, Sun H., Li Y., Gruyer D., Glaser S. (2016), « *Fast computing and approximate fuel consumption modeling for Internal Combustion Engine passenger cars* », Transportation Research Part D 50C (2017) pp. 14-25, DOI: 10.1016/j.trd.2016.10.016, (*JCR, Transport science, IF=1.864*).
5. Susini P., Houix O., **Saint Pierre G.** (2015), « *Loudness effect on pairwise comparisons and sorting task* », J. Acoust. Soc. Am. 138, 1918, DOI:10.1121/1.4934034 (*JCR, Psycho-acoustic, IF=1.503*).
6. Susini P., Houix O., **Saint Pierre G.** (2015), « *The Effect of Loudness on the Perceptual Representation of Sounds With Similar Timbre* », Acta Acustica United With Acustica Vol. 101, N. 6, pp. 1174-1184(11), DOI:10.1016/j.trc.2015.04.026 (*JCR, Psycho-acoustic, IF=0.783*).
7. Romon S., Bressaud X., Lassarre S., **Saint Pierre G.**, Khoudour L., (2015) « *Map-matching algorithm for large databases* », Journal of Navigation, Volume 68, Issue 05, pp 971-988, DOI:10.1017/S0373463315000156 (*JCR, Engineering, 0.949*).
8. Orfila O., **Saint Pierre G.**, Messias M., (2015), « *An Android Based Ecodriving Assistance System to Improve Safety and Efficiency of Internal Combustion Engine Passenger Cars* », Transportation research part C: Emerging technologies, Volume 58, Part D, Pages 772–782 (*JCR, Transport science, IF=2.818*).

9. Andrieu C., **Saint Pierre G.** (2014), « *Evaluation of ecodriving performances and teaching method: comparing training and simple advices* », European Journal of Transport and Infrastructure Research, 14(3), pp. 201-213, ISSN: 1567-7141 (*JCR, Transportation, IF=0.818*).
10. Palat B., Delhomme P., **Saint Pierre G.** (2014), « *Numerosity Heuristic in Route Choice Considering the Presence of Traffic Lights* », Transportation Research Part F: Traffic Psychology and Behaviour, vol 22, pp 104-112, DOI:10.1016/j.trf.2013.11.004 (*JCR, Transportation science, IF=1.473*).
11. Pinto M., Cavallo V., **Saint Pierre G.**, (2013) « *Influence of Front Light Configuration on the Visual Conspicuity of Motorcycles* », Accident Analysis & Prevention, vol 62, pp 230–237, DOI:10.1016/j.aap.2013.09.026 (*JCR, Transportation science, IF=2.07*).
12. Ponsot E., Susini P., **Saint Pierre G.**, Meunier S. (2013), « *Temporal weights of loudness for sounds with increasing and decreasing intensity profiles* », The Journal of acoustical society of America, 10/2013; 134(4):EL321-EL326. DOI:10.1121/1.4819184 (*JCR, Psycho-acoustic, IF=1.503*).
13. Popa I. S., Kharrat, A., Zeitouni K., **Saint Pierre G.** (2009), “*Base de données de capteurs à localisation mobile*”, Ingénierie des Systèmes d'information, 14, Lavoisier, pp 35-58, DOI: 10.3166/ISI.14.5.35-58 (*Scopus, Intelligence artificielle*).
14. Ndong Ngu'ema E.-P., **Saint-Pierre G.** (2008), « *Model-Based Classification with Dissimilarities: a Maximum Likelihood Approach* », Formal Pattern Analysis & Applications 09/2008; 11(3):281-298. DOI:10.1007/s10044-008-0105-2 (*JCR, Computer science, IF=0.646*).

5.3.3. Articles parus dans des revues à comité de lecture non répertoriées (ACLN)

15. d'Estampes L., Garel B., **Saint Pierre G.** (2002). « *Test séquentiel: niveau de confiance après acceptation* ». Revue de Statistique Appliquée, vol LI n°3.

5.3.4. Communications avec actes dans congrès international (ACTI)

16. Dang N-T., Cavallo V., Thomson J., Gorrini A., Piwowarczyk A., Vienne F., Bandini S., **Saint Pierre G.** (2017) « *Crossing the Street in Dyads: Social Influence on Pedestrian Behaviour in an Immersive Virtual Environment* », Road Safety and Simulation international conference, The Hague, Netherlands.
17. Méneroux Y., Le Guilcher A., Orfila O., Lusetti B., **Saint Pierre G.**, and Mustière S. (2017), « *Using Surrogate Road Network for Map-Matching: A Sensitivity Analysis of Positional Accuracy* », International Conference on GeoComputation GeoComp2017, Leeds, UK, 2017-9-4.

18. Susini P., Houix O., **Saint Pierre G.**, (2015), « *Loudness effect on pairwise comparisons and sorting* », 170th ASA meeting, Jacksonville, Etats-Unis, 2015-11-02, 138, 1918.
19. Gimonet N., Cord A., **Saint Pierre G.** (2015) « *How to predict real road state from vehicle embedded camera?* », Proceedings of the IEEE Intelligent Vehicles Symposium (IV 2015), Seoul, Korea, pp 593-598, DOI: 10.1109/IVS.2015.7225749.
20. **Saint Pierre G.**, Tattegrain H., Val C. (2014), « *Impact evaluation of speed regulation systems using naturalistic driving data: The EuroFOT example* », Transport Research Arena 2014, Paris, La Défense.
21. Freitas-Salgueiredo C., Orfila O., **Saint Pierre G.**, Glaser S., Coiret A., Vandanjon P.O. (2014), « *Overview of solutions for reducing CO2 emissions during the use phase of passenger cars* », Transport Research Arena 2014, Paris, La Défense.
22. Orfila O., **Saint Pierre G.**, Messias M. (2014), « *Development of an ecodriving assistance application for nomadic devices performing real-time and post trip coaching for road vehicles* », Transport Research Arena 2014, Paris, La Défense.
23. Romon S., **Saint Pierre G.**, Khoudour L. (2014), « *Fuel consumption and speed profiles comparisons according to infrastructure using probe vehicle data* », Poster au Transport Research Arena 2014, Paris, La Défense.
24. Andrieu C., **Saint Pierre G.**, Bressaud X. (2013), « *Estimation of space-speed profiles: A functional approach using smoothing splines* », Intelligent Vehicles Symposium 2013, June 23-26, 2013, Gold Coast, Australia, DOI: 10.1109/IVS.2013.6629594.
25. Orfila O., **Saint Pierre G.**, Andrieu C. (2012), « *Gear Shifting Behavior Model for Ecodriving Simulations Based on Experimental Data* », 15th meeting of the Euro Working Group on Transportation, EWGT 2012, 10-13 septembre, Paris, France.
26. Andrieu C., **Saint Pierre G.** (2012), « *Comparing effects of ecodriving training and simple advices on driving behavior* », 15th meeting of the Euro Working Group on Transportation, EWGT2012, 10-13 septembre, Paris, France.
27. Andrieu C., **Saint Pierre G.** (2012), « *Using statistical models to characterize eco-driving style with an aggregated indicator* », Intelligent Vehicles Symposium 2012, June 3-7, 2012, Alcalá de Henares, Spain, DOI: 10.1109/IVS.2012.6232197.
28. Andrieu C., **Saint Pierre G.** (2012), « *Speed profiles functional modeling for driving behavior and infrastructure evaluation* », Young European Arena of Research, YEAR 2012 (Finalist and bronze medal), Athènes, Greece.
29. Benmimoun M., Ljung Aust M., Faber F., **Saint Pierre G.**, Zlocki A., (2011), « *Safety analysis method for assessing the impacts of advanced driver assistance systems within the European large scale field test "euroFOT"* », European congress on intelligent transport systems, ITS2011, Lyon.
30. Rakotonirainy A., Haworth N., **Saint Pierre G.**, Delhomme P. (2011), « *Research issues in Eco-driving* », 22nd International Technical Conference on the Enhanced Safety of Vehicles, Washington DC, June.
31. Lambert A., Gruyer D., **Saint Pierre G.**, Ndjeng, A.N. (2008), « *Collision Probability Assessment for Speed Control* », 11th International IEEE Conference on Intelligent Transportation Systems, 2008. DOI: 10.1109/ITSC.2008.4732692.
32. Lambert A., Gruyer D., **Saint Pierre G.** (2008), « *A Fast Monte Carlo Algorithm for Collision Probability Estimation* », 10th International Conference on Control,

- Automation, Robotics and Vision, ICARCV, 17-20 décembre, Hanoi, Vietnam, DOI: 10.1109/ICARCV.2008.4795553.
33. Lambert A., **Saint Pierre G.** (2008), Gruyer D., « *A Monte Carlo Approach for Collision Probability Computation* », IEEE IROS08 2nd Workshop on Planning, Perception and Navigation for Intelligent Vehicles, NICE, France.
 34. **Saint Pierre G.**, Ehrlich J. (2008), « *Impact of Intelligent Speed Adaptation systems on fuel consumption and driver behaviour* », 15th World Congress on Intelligent Transport Systems.
 35. Zeitouni K., Sandu Popa I., Ehrlich J., **Saint Pierre G.**, Dupin F., Glaser S. (2008), « *On Using in-Vehicle Sensor Data for Naturalistic Driving Analysis* », 15th World Congress on Intelligent Transport Systems and ITS America Annual Meeting 2008 7 ,pp.4576.
 36. Vrignon J., Rakotonirainy A., Gruyer D., **Saint Pierre G.** (2007), « *Impact of subjective factors on driver vigilance: a driving simulator study* », in Proceedings of the Third International Driver Behaviour and Training Conference.
 37. **Saint Pierre G.**, Aron M., Bergel-Hayat R., Violette E. (2007) « *Rain reconstruction from various weather-related data sets using logistic regression: methodology and applications* », paper n°07-0916, TRB, Washington.
 38. Aron M., Bergel-Hayat R., **Saint Pierre G.**, Violette E. (2007) « *Added Risk By Rainy Weather. Some elements for the French IRCAD-SARI traffic monitoring and information system for drivers and operators* », 11th World Congress on Transportation Research, Berkeley.
 39. Garel B., **Saint Pierre G.** (2002). « *Reversible Jump Monte-Carlo Markov Chains applied to multivariate Gaussian mixtures models* ». Communication to the 26th Annual Conference of GfKl, Mannheim (Allemagne), Juillet 2002.

5.3.5. Communications avec actes dans congrès national (ACTN)

40. **Saint Pierre G.**, Andrieu C. (2013), « *Statistique et sécurité routière : L'observation de la conduite en situation naturelle* », 45ème journées de Statistique, Mai 27-31, 2013, Toulouse, France.
41. Palat B., Delhomme P., **Saint Pierre G.** (2012), « *Heuristique de numérosité dans la perception des feux tricolores et le choix de trajet* », 54ième Congrès de la Société Française de Psychologie, Montpellier, 3-5 septembre 2012.
42. Andrieu C., **Saint Pierre G.**, Bressaud X. (2012), « *Modélisation fonctionnelle et lissage sous contraintes d'un profil spatial de vitesse* », 44èmes Journées de Statistique, Bruxelles, Belgique.
43. **Saint Pierre G.**, Andrieu C. (2010). « *Caractérisation de l'éco-conduite et construction d'un indicateur dynamique pour véhicules thermiques* ». PRAC 2010.
44. Garel B., **Saint Pierre G.** (2003). « *Algorithme MCMC à sauts réversibles pour les mélanges gaussiens multivariés. Application à la détection du nombre de composants* ». Actes des 10èmes Rencontres de la Société Francophone de Classification.

5.3.6. Communications invitées sans actes (INV)

45. **Saint Pierre G.** (2016), « *La mobilité numérique au service de la sécurité routière* », exposé invité pour les 26^{ème} entretiens Jacques Cartier « Facteurs humains, technologies embarquées et numériques : Quel rôle pour les politiques de sécurité routière ? », 22-23 novembre 2016.
46. **Saint Pierre G.** (2015), « *Quantifying ecodriving performances: efficient feedback for efficient driving* », exposé invité au workshop « Human factors of green driving », TRB 94th Annual Meeting, Washington, 11 janvier 2015.
47. **Saint Pierre G.** (2014), conférencier invité lors de la table ronde « données internes, données externes », organisé par l'institut des actuaires, novembre 2014.
48. **Saint Pierre G.** (2014) «*From FOT to NDS Recent developments in UDRIVE The first large-scale European Naturalistic Driving Study* », exposé invité lors du 4ème International Symposium on Naturalistic Driving Research, Virginia Tech Transportation Institute, Blacksburg, 25-28 Août 2014.
49. **Saint Pierre G.** (2013), « *Statistique et sécurité routière : L'observation de la conduite en situation naturelle* », conférencier invité aux 45ème journées de Statistique, Toulouse, mai 2013.
50. **Saint Pierre G.** (2012), « *Design and development of mobile service for ecodriving* », conférencier invité au Workshop DECOMOBIL "Nomadic transport services for multimodal mobility", organisé par le réseau Européen HUMANIST, (<http://www.youtube.com/watch?v=-xwM56ndyNU>), décembre 2012.
51. **Saint Pierre G.** (2012), Expert invité au FOT-Net International Workshop, « *round table on Impact Assessment and Scaling up* », Vienne, Autriche, Octobre 2012.
52. **Saint Pierre G.** (2012), «*Statistical modeling of safety events in naturalistic driving data: An application to speed regulation systems* », conférencier invité à la "5th International Conference on Traffic and Transport Psychology", ICTTP 2012, Groningen, The Netherlands, August 29-31, 2012.

5.3.7. Autres publications (AP)

Pré-publications en accès libre

53. Andrieu C., **Saint Pierre G.**, Bressaud X., (2013), « *A moving fixed-interval filter/smoother for estimation of vehicle position using odometer and map-matched GPS* » (lien HAL : <http://hal.archives-ouvertes.fr/hal-00915402>, lien ArXiv : <http://arxiv.org/abs/1312.2129>).
54. Andrieu C., **Saint Pierre G.**, Bressaud X., (2013), « *A functional analysis of speed profiles: smoothing using derivative information, curve registration, and functional boxplot* » (lien HAL : <http://hal.archives-ouvertes.fr/hal-00915475>, lien ArXiv : <http://arxiv.org/abs/1312.2252>).

En cours de relecture

55. Méneroux Y., Le Guilcher A., Ghasemi Hamed M., Saint Pierre G., Mustière S., (2017), "A comparison of image and functional learning for traffic lights detection from in-vehicle GPS speed profiles", soumis à Data Mining and Knowledge Discovery.

5.3.8. Rapports de recherche y compris rapports de contrat de recherche (RAPP)

56. Ledoux, V., **Saint Pierre, G.** (2016) « Usage des véhicules sondes pour la mesure des vitesses pratiquées en France », rapport pour le groupe de travail DSCR « évolution de l'observatoire des vitesses ».
57. **Saint Pierre G.** et al. (2016), « *Eco driving in the real-world: behavioural, environmental and safety impacts* », deliverable 43.1 of the ecoDriver project, EU FP7.
58. Tattegrain, H., Letisserrand, D., **Saint Pierre, G.**, Val, C., Krishnakumar, R. (2015) « *Méthodes d'évaluation de dispositifs embarqués utilisant la conduite en situation naturelle* », livrable 2 de la convention N°: 2200683288 (DSCR-Ifsttar), Étude «IDACA» : Evaluation de l'Impact des Dispositifs d'Assistance à la Conduite Automobile.
59. Andrieu C., **Saint Pierre G.** (2015), « *Estimation fonctionnelle de profils de vitesse* », chapitre d'ouvrage collectif de l'Ifsttar intitulé "Nouveaux algorithmes de traitement de données appliqués au domaine du transport : approches multidisciplinaires."
60. Kircher K. et al., (2014), « *D41_1 Performance indicators and ecoDriver test design* », public deliverable, projet ecoDriver, FP7.
61. Woldeab, Z. et al., (2014), « *D32.1: Individual Use Cases and Test Scenarios Definition* », public deliverable, projet ecoDriver, FP7.
62. Nouveliere L. et al, (2014), « *D21.3: Powertrain Model Validation* », public deliverable, projet ecoDriver, FP7.
63. Seewald P., Orfila O., **Saint Pierre G.**, (2014), « *D23.2: Simulation and analysis document for on-line vehicle algorithms* », public deliverable, projet ecoDriver, FP7.
64. **Saint Pierre G.** (2012), « *Modélisation statistique de trajectoires réelles de véhicules en virage* », Livrable L.4.1.3 du projet PARTAGE (ANR), 59 pages.
65. Malta L., et al. (2012), « *Final results: Impacts on traffic safety* », livrable D6.4 du projet euroFOT (FP7), 215 pages.
66. Faber F., et al. (2012), « *Final results: Impacts on traffic efficiency and environment* », livrable D6.5 du projet euroFOT (FP7), 141 pages.
67. Sanchez D. et al. (2012), « *Final results: User acceptance and user-related aspects* », livrable D6.3 du projet euroFOT (FP7), 427 pages.
68. **Saint Pierre G.** et al. (2010), « *Report of the results of adaptation, in-vehicle implementations and piloting* », livrable D2.2 du projet EuroFOT.
69. Aron M., Bergel-Hayat R., Debaène J-Y., Dodet L., Le-Breton P., Lemaire E., Louah G., **Saint-Pierre G.**, Violette E. (2006), « *Sur-risque par temps de pluie* », rapport du WBS 1.6 du projet PREDIT IRCAD-SARI.
70. **Saint Pierre G.**, (2006). « *Modèles statistiques pour données qualitatives* », rapport de recherche pour l'entreprise Renault.

71. Antoch J., Deshouillers J-M., Dutuit Y., Jaruskova D., **Saint Pierre G.** (2004). « *Report on Stochastic processes and reliability, WP 1 - Calculation of relevant reliability characteristics from field operational data* ». Technical report, ACLIMAT project.
72. Antoch J., Deshouillers J-M., Dutuit Y., Jaruskova D., **Saint Pierre G.** (2004). « *Report on Stochastic processes and reliability, WP 1 - Bibliographic research and reporting* ». Technical report, ACLIMAT project.
73. Antoch J., Deshouillers J-M., Dutuit Y., Jaruskova D., **Saint Pierre G.** (2004). « *Report on Stochastic processes and reliability, WP 1 - Basic Data & Theoretical Tools for Life Test Modelling* ». Technical report, ACLIMAT project.

5.3.9. Thèse/mémoire

74. Thèse intitulée : « *Identification du nombre de composants d'un mélange gaussien par chaînes de Markov dans le cas multivarié, ou par maximum de vraisemblance dans le cas univarié* ». Soutenue le 26 septembre 2003 à l'Université Toulouse III.
75. Mémoire de DEA : Saint Pierre G. (1999). « *Estimation et tests dans les mélanges gaussiens* », sous la direction de B. Garel.

5.3.10. Autres productions :

Séminaires et conférences sans actes

Outres les conférences invitées, j'ai eu l'opportunité de donner de nombreux séminaires pour divers publics et occasions. Des exposés dans les séminaires internes du LCPC puis de l'Ifsttar, aux travaux présentés dans des conférences, j'aime aussi participer à l'interdisciplinarité en proposant quelques exposés plus pédagogiques sur la mise en œuvre de la méthodologie statistique.

Activités éditoriales

- Journal et revues : relecteur pour « *Transportation Research Part C* », « *Computational Statistics and Data Analysis* », « *sustainability* »
- Conférences : relecteur régulier pour « *Transports research arena* », « *intelligent Vehicles symposium* ».
- Expertise scientifique pour l'Agence Nationale de la Recherche (ANR).

Articles de vulgarisation & interviews

- Participant invité à la conférence-débat « *Écrire la Ville du Futur* », stand de la Région IDF au Salon du Livre 2016, Porte de Versailles.
- Interview lors de la démonstration d'écoconduite virtuelle menée par le LIVIC au Salon de l'Auto 2014, pour France Bleu.

- Interviews lors de la conférence TRA (Transportation Research Arena, Paris, avril 2014) au sujet du projet ecoDriver : Euronews et France Bleu.
<https://www.youtube.com/watch?v=6H4PV4YEEs>.
- Interview pour le magazine Turbo (M6-Mars 2013) : « *technologies embarquées : si ces nouvelles technologies sont utilisées pour améliorer la conduite, elles peuvent aussi être sources de distraction* ».
- Article « *Les smartphones s'offrent une écoconduite* » pour le magazine Trajectoires (le magazine de la recherche, de l'expertise et des métiers de l'ifsttar), novembre 2012.
- G. Saint Pierre, 2009, « *LAVIA, sécurité et économie de carburant* », article pour l'ouvrage collectif « *Paroles de chercheurs N°2* ». Editions INRETS.
- Marie Hélène Massot (LVMT), Guillaume Saint Pierre (LIVIC), Philippe Dupuy (LTE), 2009, « *Véhicules électriques... l'INRETS s'organise* » paru dans AXES (revue interne à l'INRETS).
- Interview pour les cahiers de l'OVE (Observatoire du véhicule d'entreprise) et le numéro spécial sur l'éco-conduite, 2009 : « *Le LIVIC étudie scientifiquement l'éco-conduite* ».
- Interview sur le sujet de l'éco-conduite donnée au magazine Terra Eco, 2008 (paru aussi dans LeMonde.fr).

6. Synthèse des activités de recherche et principales contributions scientifiques

6.1. Introduction générale

J'ai obtenu mon doctorat en mathématiques appliquées à l'université Paul Sabatier Toulouse III en 2003, suite à l'obtention d'une bourse ministérielle, d'un poste de moniteur, puis ensuite de demi-ATER. J'ai ainsi pu acquérir un très bonne maîtrise des outils théoriques de la statistique comme les modèles à variables latentes (mélanges gaussiens, modèles de classification), ou les chaînes de Markov dans un cadre bayésien. Cette expérience m'a forgé une exigence de rigueur et un goût pour l'application concrète de mes recherches. Ces deux aspects expliquent en grande partie mes choix de recherche et de carrière. J'ai bien entendu eu la chance d'enseigner à l'université pendant quatre années, avec une partie de mon service en écoles d'ingénieur (ENSICA, ENSEEIHT). Cela m'a donné le goût de l'enseignement, que j'ai exercé continuellement jusqu'en 2016, notamment en étant responsable d'un module de Master 2 recherche (Statistique appliquée et logiciel SAS) à l'université Pierre et Marie Curie (Paris VI).

Ma première expérience de post-doctorat s'est révélée être très appliquée et déjà dans le domaine du transport et des systèmes embarqués. Grâce à une bourse du Ministère de l'Economie et des Finances en partenariat avec l'entreprise Thalès Avionics, et dans le cadre d'une participation au projet Européen ACLIMAT, j'ai eu la charge de mettre au point un modèle de fiabilité pour du matériel aéronautique embarqué hautement fiable.

J'ai ensuite rejoint l'INRIA et l'équipe du Pr. G. Celeux à l'université d'Orsay (Paris XI). L'objectif était de construire un modèle de classification pour données de distances, cas très fréquent dans la pratique. Ce travail utilisait lui aussi les modèles de mélanges gaussiens à variables latentes.

J'ai ensuite souhaité explorer plus avant l'application de la statistique à des domaines variés, et ai accepté une mission de deux mois pour l'entreprise Renault, consistant à apporter mon soutien à une doctorante CIFRE dans l'équipe de perception et design sonores de l'IRCAM (Institut de Recherche et Coordination Acoustique et Musique). J'ai proposé un modèle statistique permettant d'expliquer la classification du niveau d'urgence de sons d'un habitacle de voiture selon leurs caractéristiques. Cette expérience très riche m'a permis de construire une collaboration régulière avec cette équipe sur laquelle je reviendrais plus loin.

Fort de cette double expérience dans les transports (Thalès, Renault) j'ai obtenu un poste de chercheur contractuel à l'INRETS (Institut national de recherche sur les transports et leur sécurité) financé par un projet du Ministère de l'Equipement. C'est dans l'équipe GRETTIA

(Génie des. Réseaux de Transport Terrestre et Informatique Avancée) que j'ai vraiment découvert la richesse et la complexité des problématiques liées au transport. Mon travail a consisté à reconstruire le niveau de précipitation le long d'un itinéraire routier à partir des informations de stations météorologiques, et d'en déduire un sur-risque d'accident par temps de pluie.

L'expérience du doctorat, combinée à celle de ces différents travaux de statistique appliquée en association avec le monde industriel, m'a convaincu du potentiel exceptionnel de ma discipline. Mon attrait pour les questions liées aux transports m'a incité à candidater avec succès au laboratoire central des ponts et chaussées (LCPC) sur un poste de chargé de recherche dans le domaine des systèmes de transports intelligents (ITS).

De janvier 2007 à février 2017, j'ai occupé un poste de chercheur au laboratoire LIVIC (Laboratoire sur les Interactions Véhicule-Infrastructure-Conducteur) grâce auquel j'ai pu mener l'essentiel des travaux présentés dans ce mémoire. Lors des deux dernières années passées dans cette équipe, j'ai pu assumer l'animation d'un des axes de recherche que j'ai contribué à développer : le traitement de l'information pour la maîtrise de la mobilité (TIMM, 2 CR, 1 IR, 3 post-docs, 1 doctorant).

Le LIVIC a pour mission d'exercer des activités de recherche, de développement et d'expérimentation visant à faire émerger des systèmes intelligents d'aide à la conduite. La récente explosion technologique de ces systèmes a conduit à l'apparition du concept de véhicule intelligent, puis de véhicule autonome, et maintenant celui de véhicule connecté. Le LIVIC étudie scientifiquement ce nouvel objet, et la manière dont ses différentes déclinaisons influencent déjà, ou vont influencer, notre façon de conduire et de nous déplacer. Pendant ces dix ans, j'étais le seul probabiliste-statisticien de l'équipe. J'ai eu pour mission de prendre en charge l'activité d'évaluation des systèmes embarqués d'assistance à la conduite.

Mon expérience dans cette équipe de recherche a été extrêmement enrichissante car il s'agit d'un domaine foisonnant faisant régulièrement l'actualité, et lié à de nombreux enjeux sociétaux. Le travail de recherche autour de ces questions implique une grande multidisciplinarité, tout autant qu'une proximité nécessaire avec le monde industriel.

Mes thématiques de recherche développées au LIVIC sont le fruit d'une adaptation entre mes connaissances en statistique appliquée et les recherche menées par l'équipe avant mon arrivée. J'ai co-encadré quatre doctorants, de nombreux stagiaires, et 3 post-doctorants dans le cadre de projets financés (cf. liste détaillée des encadrements **page 36**).

Au premier mars 2017, j'ai obtenu une mutation au Cerema dans l'équipe ESAD-ZELT (DTER Sud-Ouest) dont la mission consiste à mener des activités d'études et de recherche autour de l'évaluation des systèmes de transports intelligents appliqués à la gestion des déplacements. Faisant suite à de longues années de collaboration, cette mutation géographique me permettra

de poursuivre avantageusement les axes de recherches développés à l’Ifsttar et détaillés dans ce manuscrit.

Je candidate à l’obtention du diplôme d’habilitation à diriger des recherches de l’université d’Evry Val d’Essonne afin de pouvoir poursuivre mes activités d’encadrement de doctorants, et assumer la coordination des recherches en statistique appliquée pour l’évaluation des véhicules connectés ou autonomes. Il est important pour la suite de ma carrière d’obtenir la reconnaissance de mes pairs à travers l’obtention de l’HDR qui me permettra de continuer à développer sereinement des axes de recherche liés à l’usage des modèles aléatoires pour mieux maîtriser la mobilité de demain, par le biais d’encadrement d’équipes ou de doctorants.

J’ai su développer et entretenir de nombreuses collaborations, et ai participé à de nombreux projets de recherche depuis la fin de mon doctorat. Je souhaite présenter et détailler mon parcours de recherche à partir de la date du 1^{er} décembre 2005, premier jour de mon contrat à l’IRCAM, qui fut pour moi un changement dans la façon dont je voyais le métier de chercheur. En effet, certaines de mes collaborations, ainsi que mon goût pour l’application des statistiques à la psychologie, datent de cette expérience.

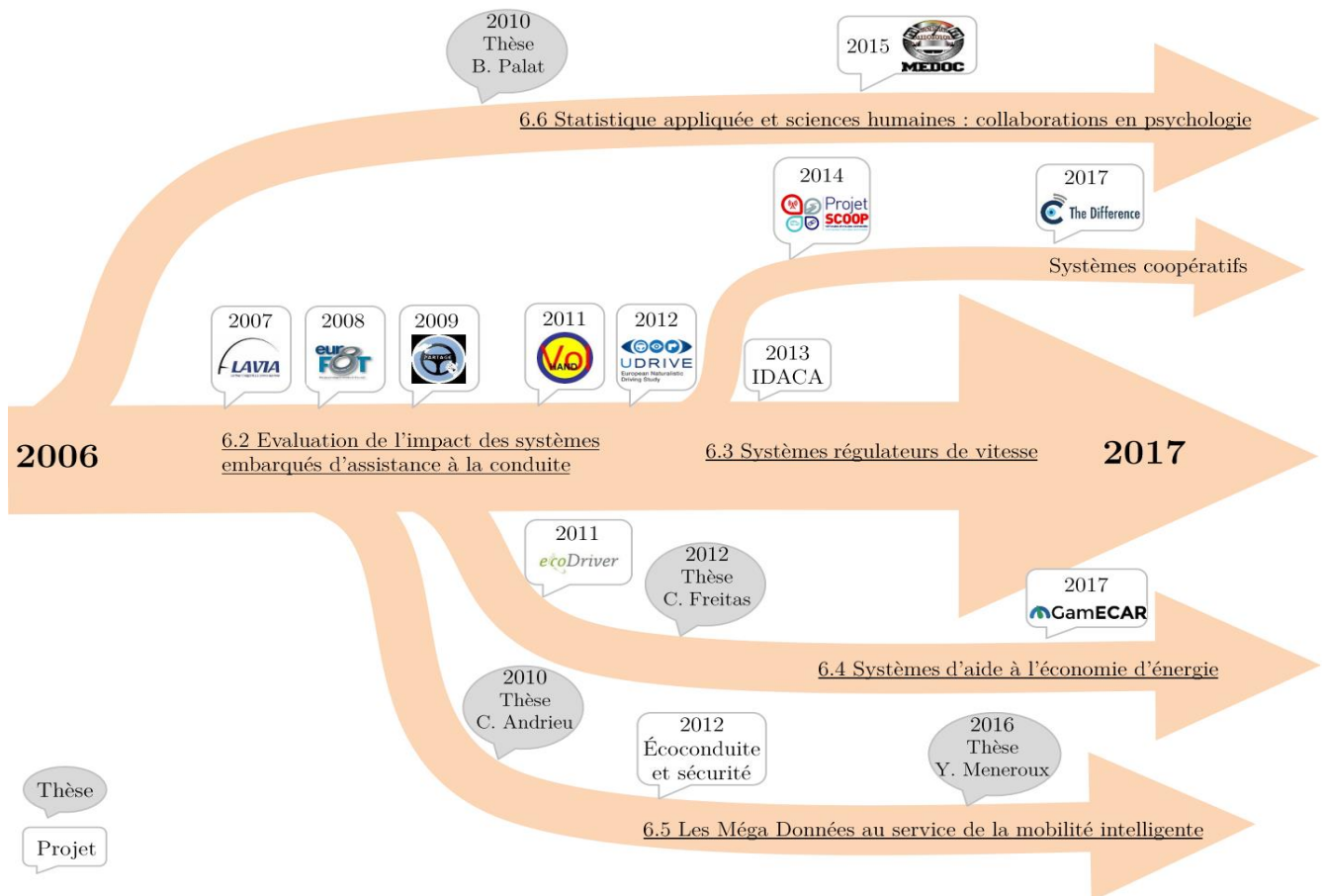


Figure 3: Proposition de structuration de l'ensemble de mes recherches (2006-2016)

La Figure 3 présente une structuration de l'ensemble de mes recherches depuis cette date. J'ai souhaité proposer un développement de mes recherches selon un axe temporel articulé autour de l'évaluation de l'impact des systèmes embarqués. Les thèmes de recherche majeurs sont indiqués dans les flèches temporelles, et sont associées à un numéro de section reprenant la numérotation des sections de ce document. Les flèches secondaires permettent de situer temporellement l'émergence d'un nouveau thème de recherche ainsi que les principaux projets de recherche et les encadrements qui lui sont associés.

L'axe principal de mes recherches concerne l'évaluation des systèmes embarqués d'assistance à la conduite au sens large, avec une expertise approfondie sur les systèmes agissant sur la vitesse (Lavia, limiteur et régulateur). J'ai ensuite élargi ce champ d'action aux systèmes d'aide aux économies d'énergies, et plus récemment aux systèmes coopératifs et connectés.

En parallèle de ces travaux, j'ai acquis une expertise sur les données générées par les véhicules modernes. Cela m'a conduit à rechercher les exploitations possibles de ces données dans le contexte de la mobilité intelligente. J'ai ainsi développé une approche statistique pour construire des indicateurs dynamiques de la conduite, que j'ai appliquée à la conduite éco-responsable. Ce travail m'a permis de comprendre les liens entre comportement de conduite et infrastructure, que j'ai exploité pour proposer des méthodes permettant d'inférer l'infrastructure routière à partir du comportement de conduite.

Depuis mon passage par l'IRCAM et la découverte des liens entre mes compétences mathématiques et le domaine des sciences humaines expérimentales, j'ai poursuivi mes recherches dans ce champ d'application, en veillant à maintenir un lien étroit avec la problématique du transport.

La suite du document s'attache à décrire plus précisément l'ensemble de ces axes de recherches. Pour introduire ma thématique de recherche principale et ainsi alléger la suite du document, j'ai choisi de proposer une première section d'ordre méthodologique qui est le fruit de mes dix années d'expérience dans ce domaine. Ainsi, la section 6.2 présente l'ensemble des étapes d'un plan d'analyse de l'impact d'un système embarqué sur la conduite, et peut s'apparenter à une formation sur ce sujet.

La section 6.3 présente l'application des méthodes proposées en 6.2 au cas des systèmes régulateurs de vitesse. La section 6.4 présente quant à elle une application aux systèmes d'aide aux économies d'énergies.

La section 6.5 constitue la description d'un axe de recherche à part entière, pour lequel j'ai développé des approches innovantes d'exploitation des données véhicules dans un autre objectif que la pure évaluation de l'impact d'un système.

La section 6.6 propose un bilan des travaux auxquels j'ai collaboré dans le domaine des sciences humaines et de la psychologie expérimentale.

Enfin, la section 7 présente mon projet de recherche que je compte mener à bien au Cerema.

6.2. Evaluation des systèmes embarqués d'assistance à la conduite : généralités

6.2.1. Introduction

Lors de mon arrivée au LIVIC en janvier 2007, j'ai dû m'adapter aux recherches menées dans l'équipe et y trouver ma place. Sur les conseils du directeur Jacques Ehrlich, je me suis investi dans le domaine de l'évaluation des systèmes d'aide à la conduite agissant sur la vitesse. C'est ainsi que j'ai eu en charge l'évaluation de l'impact sur la consommation de carburant d'un limiteur de vitesse intelligent dit LAVIA. J'ai prolongé cette expertise en représentant le LCPC (Laboratoire Central des Ponts et Chaussées) dans le projet européen EuroFOT lancé en 2008 pour étudier 8 systèmes différents, avec la responsabilité de l'étude sur le limiteur / régulateur de vitesse. A travers ces deux premiers projets, j'ai découvert le monde de la recherche appliquée s'appuyant sur des expérimentations à grande échelle et utilisant des flottes de véhicules. J'ai pu me familiariser avec les informations générées par les véhicules modernes, et comprendre les enjeux méthodologiques associés à de telles études à grande échelle.

Au même moment (2008), une proposition de méthodologie adaptée ce type de projet, appelés « FOT » (Field Operational Test), était publiée sous le nom de FESTA. Plusieurs fois remaniée (2008, 2011, 2014, 2016, cf. http://wiki.fot-net.eu/index.php/FESTA_Handbook), cette méthodologie constitue encore aujourd'hui une référence incontournable pour la pratique de l'évaluation de la conduite ou des systèmes embarqués. Similaire à la mise en œuvre de cohortes de patients dans le secteur médical, cette approche s'est généralisée pour l'étude des systèmes de transport intelligents.

Au fil de mes années d'implications dans les projets de ce type, j'ai pu contribuer à l'enrichissement de l'approche FESTA en formalisant les méthodes statistiques les plus pertinentes pour tenir compte de la spécificité des données issues de véhicules « traceurs ». Ayant eu à décrire l'ensemble de la méthodologie à maintes reprises, tout en ayant parfois la charge de l'enseigner à mes partenaires académiques ou industriels, c'est naturellement que j'en suis venu à rédiger un document plus précis que la méthodologie FESTA sur certains aspects.

Cette section représente donc ma contribution à la méthodologie d'analyse de ces systèmes, à partir de données réelles enregistrées pendant une longue durée, par des participants volontaires au volant de leurs propres véhicules. En guise d'introduction à ce manuscrit d'habilitation à diriger les recherches, j'ai donc choisi d'illustrer concrètement et dans le détail les principaux aspects de cette méthode. Cette section peut donc être vue comme la colonne vertébrale de mes recherches concernant l'évaluation de l'impact des systèmes embarqués, nécessaire à la compréhension des sections suivantes (6.3 et 6.4) plus axées sur les applications et les résultats (cf. Figure 4).

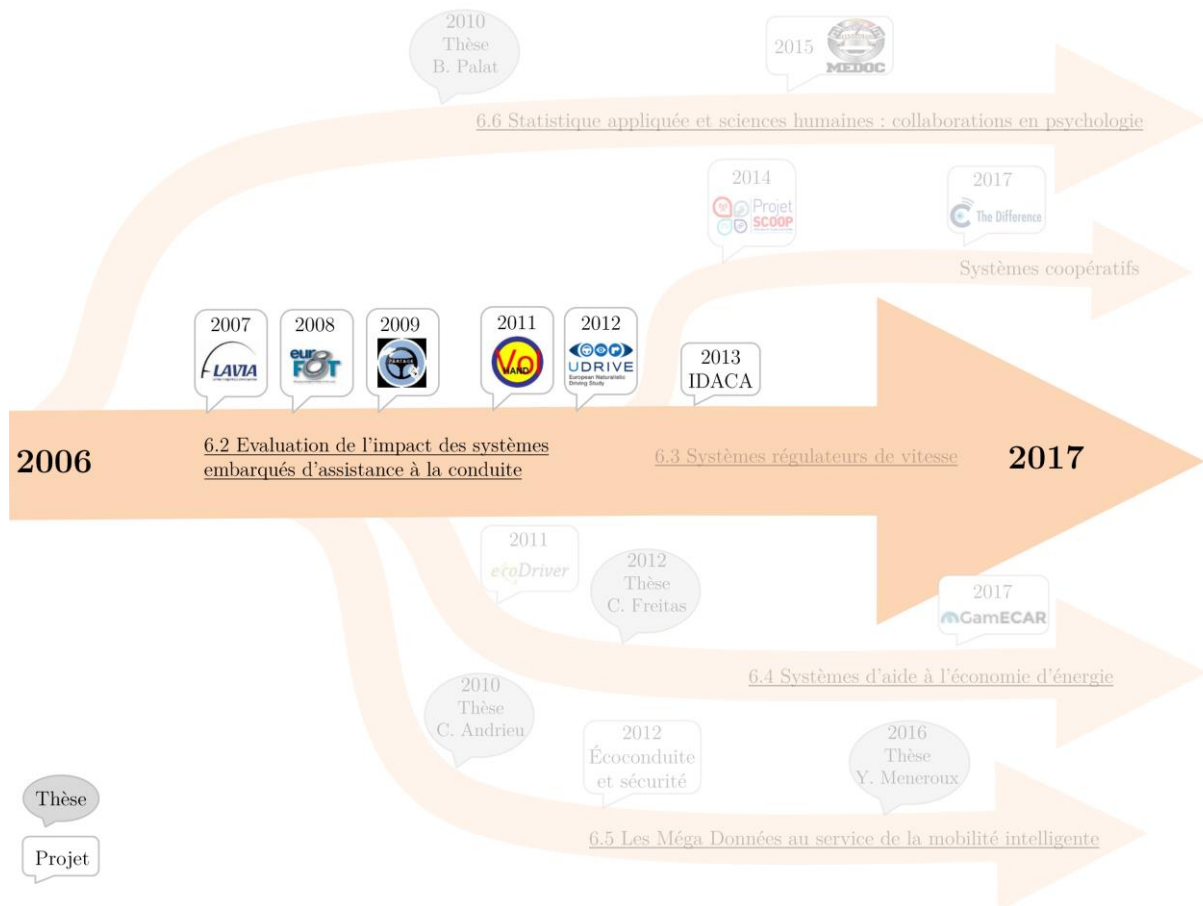


Figure 4 : Mise en évidence des axes de recherche développés dans cette section.

Cette section présente donc en premier lieu le domaine des assistances à la conduite et les origines de la méthodologie d'évaluation initiale (section 6.2.1), en poursuivant avec le contexte technique et expérimental du recueil de données (section 6.2.2). Sont ensuite présentées quelques définitions liées au vocabulaire technique (section 6.2.3) ainsi que les différents protocoles envisageables et leurs implications méthodologiques (section 6.2.4). Le procédé d'analyse par questions de recherche est présenté (section 6.2.5), suivi par une revue des modèles statistiques classiquement utilisés, ainsi que les modèles plus récents que j'ai mis en œuvre dans mes travaux (section 6.2.6). Enfin, je détaille la mise en œuvre de ces approches dans le contexte de l'évaluation de l'impact sur la sécurité (section 6.2.7) et de l'impact sur la consommation (section 6.2.8).

6.2.1. Contexte

Les systèmes embarqués d'assistance à la conduite sont nombreux et hétérogènes. Chacun d'eux vise à assister le conducteur sur un aspect bien précis de sa conduite, en usant de moyens d'actions différents, soit actifs, soit informatifs. Certains de ces systèmes sont encore peu répandus, tandis que d'autres sont devenus très familiers des conducteurs européens. Une catégorisation de ces systèmes selon leur fonctionnalité et leur type d'action est proposée au Tableau 2.

A ces systèmes développés par les acteurs de l'industrie automobile, se sont ajoutées récemment les applications d'aide à la conduite ou à la navigation fonctionnant sur des smartphones. J'ai moi-même contribué au développement d'une application d'assistance à l'écoconduite dans le cadre du projet européen ecoDriver (2011-2016).

Tableau 2 : Proposition de catégorisation de certains systèmes d'assistance à la conduite (source : étude IDACA).

Systèmes		Fonctionnalités	Type d'actions
ACC	Adaptive Cruise Control	longitudinale	Accélération/Freinage
LV	Limiteur de vitesse	longitudinale	limitation de vitesse
RV	Régulateur de vitesse	longitudinale	Accélération
ISA	Intelligent Speed Adaptation (limiteur de vitesse intelligent)	longitudinale	limitation de vitesse
DW	Distance Warning	longitudinale	Information
AEBS	Advanced Emergency Braking System (high/medium speed)	longitudinale	Freinage
AEBS City	Advanced Emergency Braking System City (low speed <30km/h)	longitudinale	Freinage
BS/BSW	Blind Spot Detection/Blind Spot Warning	latérale	Information
BSI	Blind Spot Intervention	latérale	Angle volant
CW	Collision Warning	longitudinale	Information
CA (AEB)	Collision Avoidance and Automatic Emergency Braking – not pedestrian	longitudinale	Freinage
LCA	Lane Changing Assistant	latérale	Information
LKA	Lane Keeping Assistant	latérale	Angle volant
LDW	Lane Departure Warning	latérale	Information
NV	Night Vision	perception	Information
UTA	Unstable Trajectory Alert	latérale	Information
UPA	Ultrasonic Parking Assist	manœuvre	Information
RVC	Rear View Camera +Bird View (or Allround View Camera)	manœuvre	Information
HFP	Hands Free Parking	manœuvre	Angle Volant Accélération Freinage
AHL	Automatic High Low Beam	perception	Changement position feux
FBL	Fixed Bending Lights	perception	Changement position feux

OSP	Over speed prevention	longitudinale	Information
BA	Brake Assist	longitudinale	Accélération/freinage
PBA	Predictive Assist Braking	longitudinale	Freinage
CSW	Curve Speed Warning	latérale	Information
PAB	Predictive Assist Braking	longitudinale	Freinage
EBA	Emergency Brake Assist	longitudinale	Freinage
FVCMS	Forward vehicle collision mitigation systems	longitudinale	Freinage
ABS	Antilock Braking System	longitudinale	Freinage
ESC	Electronic Stability Control	latérale	Angle volant
RD	Rollover Detection/Active rollover system/Roll stability control	latérale	Angle volant
MB	Multicollision brake	longitudinale	Freinage
Stop and Go	System Stop and Go (ou Stop & Start)	longitudinale	Accélération/Freinage
GPS	Global Positioning System	perception	Information

Qu'ils soient développés par l'industrie automobile, ou qu'ils soient sous la forme d'une application pour smartphone, la majorité de ces systèmes permettent d'influencer la conduite soit directement (contrainte sur la vitesse par exemple), soit indirectement (information fournie au conducteur). Dans ce document, nous ne traiterons pas les systèmes actifs agissant directement sur la dynamique du véhicule de façon « masquée » pour le conducteur (Esp et ABS par exemple).

L'enjeu scientifique consiste à produire une méthode suffisamment générale pour s'adapter à différents systèmes et à même de fournir des éléments objectifs pour quantifier l'impact d'un système sur le comportement de conduite. Dans ce contexte, les mesures effectuées sont le résultat des actions décidées par l'humain au volant, en fonction de sa perception de l'environnement de conduite, et des informations fournies par le système. Nous observons donc le résultat d'un phénomène de prise de décision complexe, dans un environnement lui aussi complexe. Une bonne évaluation doit donc observer le résultat de ce phénomène (les actions de conduite), les stimuli générés par le système, ainsi que les caractéristiques de l'environnement de conduite.

La complexité inhérente à la conduite automobile peut toutefois se simplifier en focalisant l'étude sur des cas d'usages spécifiques à chaque système, et en formulant des hypothèses scientifiques précises inspirées de la connaissance métier, qui peuvent être ensuite validées ou infirmées à l'aide de la théorie des tests statistiques. Ce travail passe par la définition et le calcul d'indicateurs pertinents associés à la question de recherche et au cas d'usage ciblé.

Une expérience à grande échelle destinée à évaluer en conditions réelles l'impact d'un système embarqué, passe donc par la définition d'un protocole complexe articulant séquentiellement différentes étapes d'étude, de conception, d'instrumentation, de recueil de données, et d'analyse. La gestion sur plusieurs années d'un tel projet nécessite le recours à des méthodes quasi-industrielles. Une approche existe qui est formalisée depuis longtemps par l'industrie (logicielle notamment) sous le nom de « cycle en V ». C'est sur ce type d'approche que s'appuie la méthodologie FESTA, dont le cycle en V adapté au contexte des systèmes embarqués est présenté à la Figure 5.

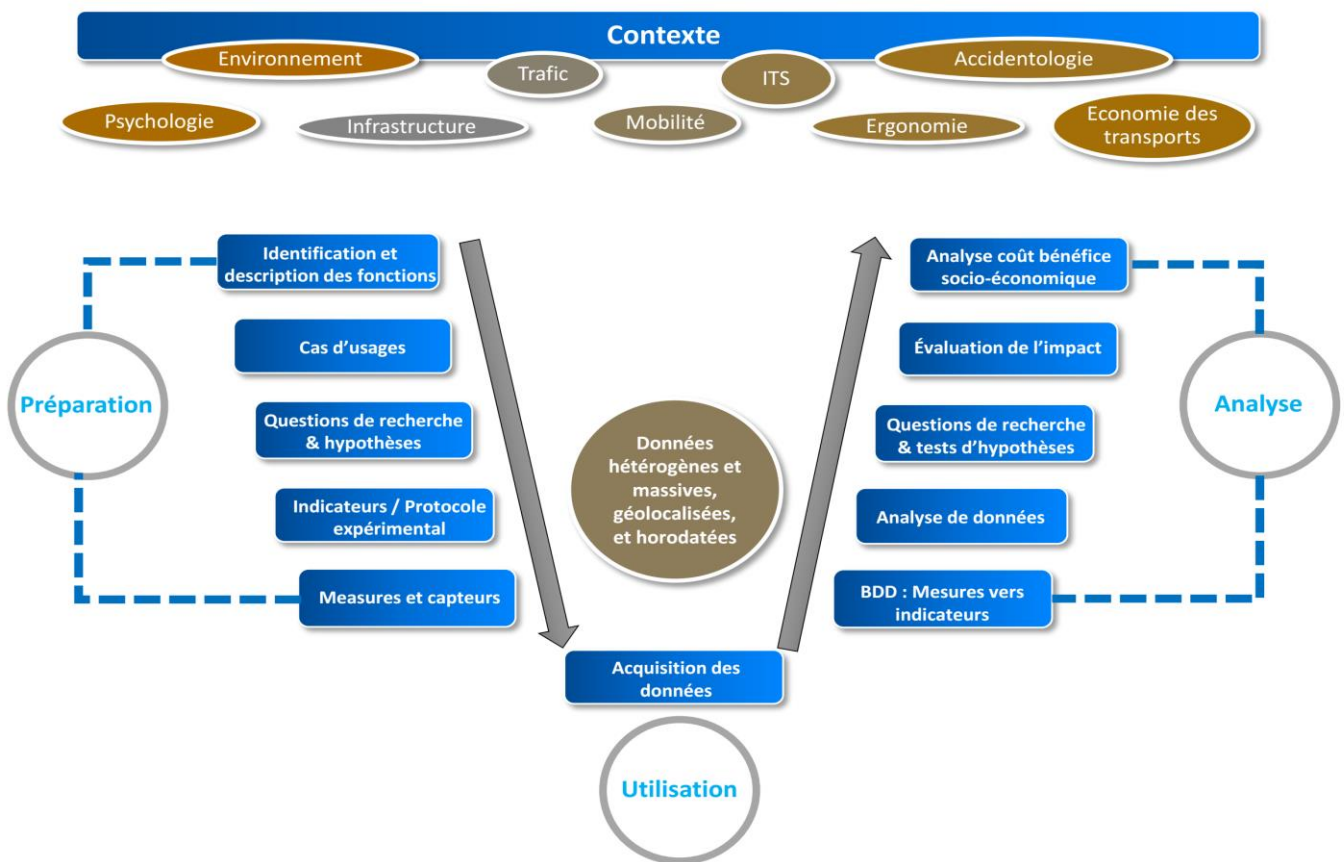


Figure 5 : Illustration du "cycle en V" utilisé dans les FOT (Field Operational Tests).

Un des défauts de ce type de méthode de gestion de projet, consiste à considérer que les différentes étapes du cycle peuvent être dissociées. En réalité, ces étapes nécessitent souvent une approche itérative. On pensera par exemple à la rédaction d'une question de recherche associée à la définition d'un indicateur, qui nécessite une information particulière, qui peut s'avérer impossible à recueillir lors de la phase expérimentale.

Le rôle du statisticien ne se limitant pas à l'analyse d'un tableau de données, il se doit d'intervenir dès le début du cycle dans la rédaction des questions de recherche, des indicateurs associés et la

rédaction du protocole expérimental. Il est évidemment très impliqué aussi dans la partie ascendante du cycle dédiée aux analyses. Une bonne part de mon activité de recherche s'est exercée dans le contexte de projets de recherche utilisant le cycle en V (cf. Tableau 3). J'ai donc souhaité, pour cette section introductive, formaliser les différentes étapes de cette démarche de façon compréhensible pour le lecteur novice. Cette section sert ainsi de référence à l'ensemble du document et sa lecture apparaît indispensable à la bonne compréhension de la suite du document.

6.2.2. Recueillir les données

La généralisation des technologies du numérique dans les véhicules modernes permet de recueillir leurs « traces numériques », ces véhicules pouvant donc être considérés comme des capteurs mobiles explorant le réseau routier en continu. Suivant les objectifs, l'instrumentation de ces véhicules peut être lourde, permettant ainsi l'enregistrement de dizaines de paramètres relatifs au véhicule, au conducteur ou à son environnement, mais elle peut au contraire être très légère grâce au développement des smartphones qui rendent facilement accessibles les informations essentielles que sont la position et la vitesse du véhicule.

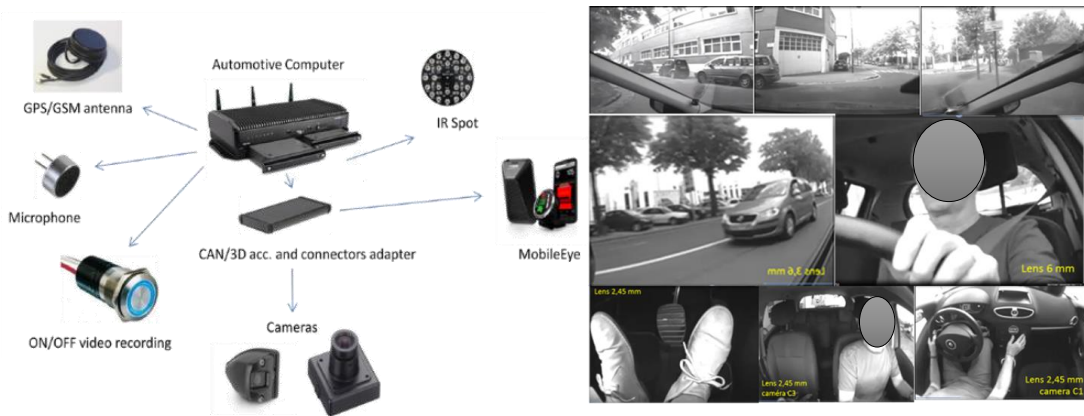
Dans le cas d'une instrumentation lourde, la plupart des paramètres relatifs au véhicule (vitesse instantanée, régime moteur, distance parcourue...) et aux actions du conducteur (position des pédales de frein et d'accélérateur, état des feux...) sont collectés à l'aide d'un enregistreur de données connecté au bus CAN du véhicule. Le bus CAN est le réseau qui permet à tous les capteurs et actionneurs du véhicule de s'échanger leurs informations. Dans un câblage traditionnel, un capteur est relié à un autre (liaison point à point) ce qui fait un câblage très complexe. Dans un réseau CAN, tous les capteurs sont reliés entre eux par un même fil (le câblage est donc très simplifié) et les informations sont échangées dans des trames avec une gestion de la priorité des trames et un contrôle d'erreurs poussé. Notons que les données transitant sur le bus CAN peuvent également être recueillies par un smartphone relié (par USB ou Bluetooth) à ce réseau à l'aide d'un adaptateur branché sur la prise OBD-II ("On-Boards Diagnostics") du véhicule.



Figure 6 : Exemple d'adaptateur et de prise OBD véhicule

D'autres capteurs peuvent également être ajoutés au véhicule. On distingue généralement deux catégories de capteurs :

- les capteurs proprioceptifs, qui permettent de mesurer des grandeurs intrinsèques du véhicule (ex : odomètre, centrale inertielle...);
- les capteurs extéroceptifs, qui permettent de situer le véhicule par rapport à son environnement (ex : GPS, radar, caméra...).



Exemple de prise de vues

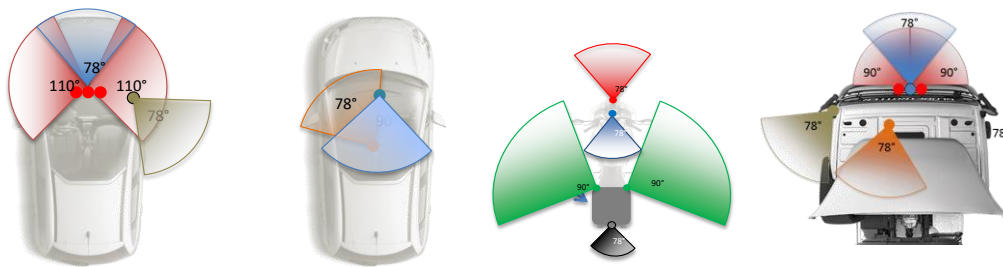


Figure 7 : Exemple d'instrumentation très complète pour l'étude de la conduite en situation naturelle pour le projet UDrive (FP7)



Figure 8 : Exemple d'instrumentation légère avec boîtier autonome avec capteurs intégrés et connexion à la prise OBD-II du véhicule (Projet MEDOC, Fondation MAIF)

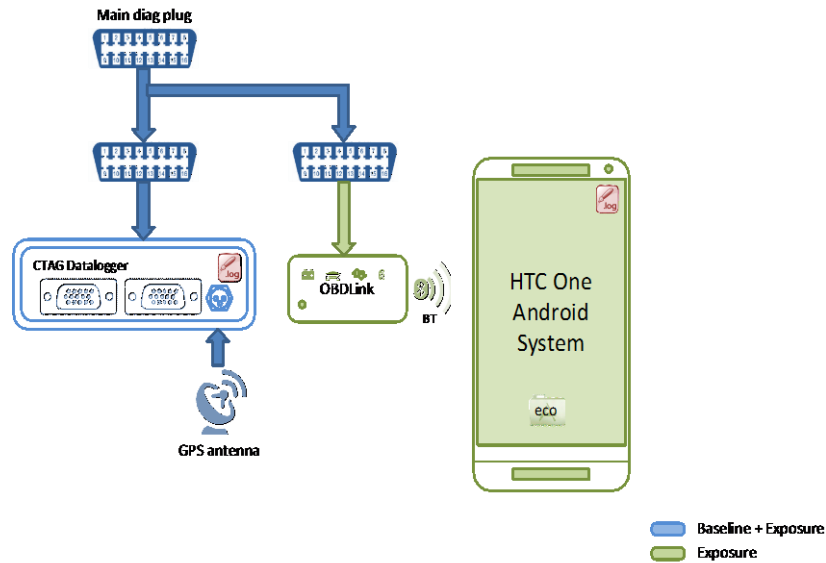


Figure 9 : Solution hybride de recueil de données du projet européen ecoDriver. En conduite normale, seul un boîtier enregistreur est connecté à la prise OBD et enregistre les données du véhicule. En période de test d’une application nomade.

Comme on peut le constater, le recueil de données lors d’une expérimentation d’étude du comportement de conduite est très hétérogène et fait appel à de nombreuses sources de données complexes (cf. **Erreur ! Source du renvoi introuvable.**). Cet aspect de l’étude du comportement en situation réelle de conduite est un véritable enjeu pour le statisticien, qui doit maîtriser les questions techniques liées à l’harmonisation et à la synchronisation des différentes sources d’information, tout autant que les pré-traitements associés aux différents capteurs

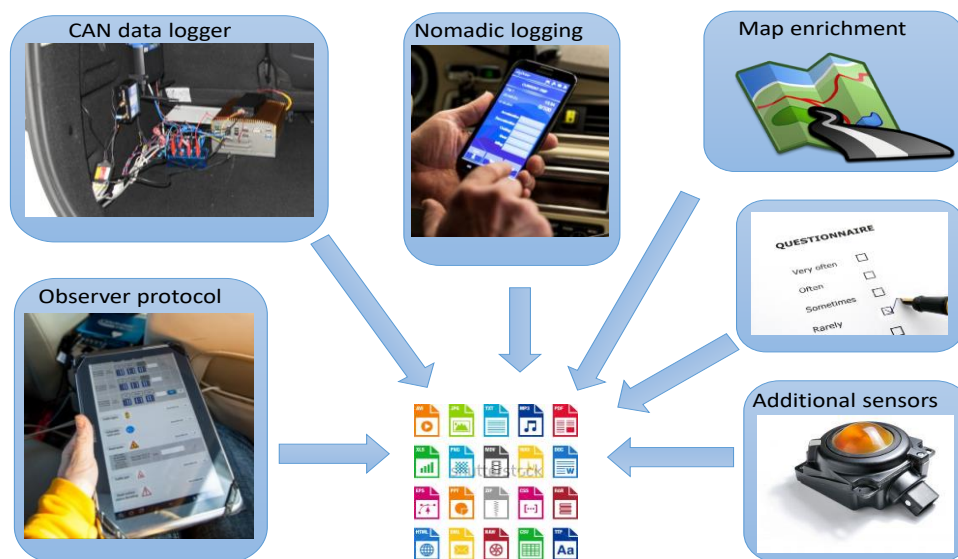


Figure 10 : Résumé des différentes sources de données possibles lors d’une étude du comportement de conduite.

utilisés. Il faut aussi citer, même si ce n'est pas l'objet du présent document, le gigantesque travail effectué par différents partenaires de ces projets pour développer les outils informatiques de traitement et d'organisation de ces données. Cela relève d'un aspect technique associé à la gestion des méga données (big data) et fait appel à une chaîne complexe de transmission des données, organisation en base de données, jusqu'aux outils de gestion et d'accès à ces données en vue de leur analyse.

Le tableau ci-dessous permet de se faire une idée de l'enjeu lié au traitement des données dans les projets visant à étudier le comportement en conditions réelles de conduite. Il présente, pour les projets auxquels j'ai participé, le nombre de véhicules impliqués (donc instrumentés) et la subvention publique accordée au projet.

Tableau 3 : Exemples de dimensionnement de flottes expérimentales de véhicules dans différents projets auxquels j'ai participé.

Années	Acronyme	Titre	Financement	Véhicules	Subvention totale
2015-2017	MEDOC	MassE de DONnées de Conduite pour la modélisation d'un feedback incitant à la conduite sécuritaire selon les déplacements des automobilistes	Fondation MAIF	200	400k€
2014-2017	SCOOP@F-1 et 2	Projet de déploiement pilote de systèmes de transport intelligents coopératifs	RTE-T	3000	6,5M€ (phase 1) 10M€ (phase 2)
2012-2017	UDrive	European naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment	UE-FP7	200	8M€
2011-2016	ecoDriver	Supporting the driver in conserving energy and reducing emissions	UE-FP7	200	14M€
2008-2012	euroFOT	European Field Operational Test on Active Safety System	UE-FP7	1000	14M€

6.2.3. Définitions

Les informations recueillies par un véhicule instrumenté et son conducteur sont très hétérogènes. Dans le cadre de mes travaux de recherche à l'Ifsttar, je me suis essentiellement intéressé aux mesures objectives enregistrées à partir de capteurs proprioceptifs ou extéroceptifs. Des exemples de mesures sont le temps et la distance parcourue, la vitesse, la température extérieure ou la consommation de carburant du véhicule. Ces mesures sont utilisées pour calculer les indicateurs de performance de la conduite ou d'un système. Il existe des mesures directes, dérivées et auto-déclarées. Ces mesures peuvent ensuite être combinées pour introduire des variables plus

pratiques pour les analyses statistiques. On parlera alors de variables situationnelles (SV) ou d'événements. Ces différents concepts sont expliqués et justifiés dans les sections suivantes.

6.2.3.1. Mesures directes

Une mesure directe est enregistrée directement à partir d'un capteur, sans autres manipulations, à l'exception des transformations linéaires.

Les mesures directes sont enregistrées directement à partir du capteur, qui est l'entité physique qui délivre une certaine valeur. Dans le cas d'un oculomètre une mesure directe peut être la direction du regard en x, y et z, et dans le cas d'un bus CAN une mesure directe pourrait être l'angle du volant ou l'angle des roues. Si la vitesse est mesurée en m/s, une transformation linéaire en km/h ne modifie pas la mesure directe en mesure dérivée.

6.2.3.2. Mesures dérivées

Une mesure dérivée est une mesure calculée à partir d'une mesure directe (cad. en appliquant des opérations mathématiques ou statistiques) ou à partir d'une combinaison de plusieurs mesures directes (ou dérivées).

Les mesures dérivées sont calculées à partir des mesures directes ou d'autres mesures dérivées ou une combinaison de ceux-ci. A titre d'exemple, si l'enregistrement du bus CAN, associé à celui d'autres capteurs, délivre à la fois la vitesse du véhicule et la distance du véhicule qui précède, on peut alors en déduire la vitesse du véhicule qui précède et le temps inter-véhiculaire (TIV). Il s'agit alors de mesures dérivées dans la mesure où les transformations nécessaires relèvent de calculs complexes, dont les définitions peuvent varier selon le contexte.

Il est important de noter que ces calculs sont effectués en temps différés sur la base de données des mesures brutes, et nécessitent d'importantes capacités de calculs.

6.2.3.3. Mesures auto-rapportées

Une mesure auto-rapportée est une mesure subjective communiquée par l'intermédiaire de questionnaires, interviews, groupes de discussion, etc.

Elle peut concerner une situation ou un comportement de conduite, mais aussi le conducteur lui-même.

Les mesures auto-rapportées ne sont en général pas obtenues de façon continue, en raison de la nature même du processus de recueil. Dans la plupart des cas, il est difficile de demander aux conducteurs de remplir trop régulièrement des questionnaires ou de remplir quotidiennement un carnet de bord. Habituellement, les conducteurs remplissent une série de questionnaires en début

d'expérimentation, répondent à quelques questions au cours de celle-ci, et complètent ces informations par des questionnaires plus complets en fin d'étude. Il est également possible de réaliser des entrevues, individuellement ou en groupes, ou de recourir à des techniques de confrontation vidéo qui consistent à faire réagir le conducteur sur un événement particulièrement intéressant de ses carnets de bord.

6.2.3.4. Mesures externes

Une mesure externe est fournie par des sources de données ne faisant pas partie de l'équipement d'enregistrement utilisé dans l'étude.

Des exemples de ces sources de données pourraient être les bases de données accident, les bases de données météorologiques, les données cartographiques, ou les bases de données de trafic routier. Dans de nombreux cas, les données obtenues par l'équipement d'enregistrement sont associées aux informations des bases de données externes via la position et le temps, et sont indépendantes du comportement de conduite. Il s'agit par exemple d'un travail d'enrichissement des données utilisant une information cartographique permettant de déduire des éléments du contexte de conduite et permettant ensuite de construire les variables décrites ci-dessous.

6.2.3.5. Variables descriptives

Les variables descriptives sont construites à partir des mesures précédentes, et permettent de déduire des variables nécessaires aux analyses statistiques du comportement de conduite. En effet, le comportement de conduite est très variable et dépend considérablement du contexte, que cela soit le comportement des véhicules environnant, ou le type d'infrastructure emprunté par le véhicule. De même, l'étude de l'impact d'un système d'aide à la conduite nécessite de s'intéresser aux situations lors desquelles il est susceptible d'agir. Cela peut être par exemple les situations d'excès de vitesse si on s'intéresse à l'usage du limiteur de vitesse. Ces situations dépendent donc non seulement des véhicules environnants, de l'infrastructure, mais aussi d'un choix délibéré du conducteur.

En conséquence, il est nécessaire de définir maintenant trois nouveaux types de variables permettant de repérer des instants de conduites (ou segments temporels) nécessaires à l'analyse.

- Variable situationnelle (SV) : Une telle variable permet de coder sous forme de catégories ou niveaux, un aspect de l'environnement routier ou une condition de conduite liée à un facteur indépendant du comportement du conducteur (mesures externes).
- Segment : un segment indique une situation de conduite particulière issue du croisement de plusieurs variables situationnelles.

- Evènement : un évènement indique une situation de conduite dépendant non seulement d'une ou plusieurs variables situationnelles, mais aussi d'un facteur lié au comportement du conducteur (identifié grâce aux mesures directes ou dérivées).

Ces variables sont définies plus en détails ci-dessous.

6.2.3.5.1. Variables situationnelles

Une variable situationnelle (SV) permet de coder sous forme de catégories ou niveaux, un aspect de l'environnement routier ou une condition de conduite liée à un facteur indépendant du comportement du conducteur (mesures externes). À tout moment, au moins un de ces niveaux doit être valide.

Ce type de variable permet de décrire l'environnement ou les conditions dans lesquelles les conducteurs évoluent, indépendamment de leur comportement propre. Chaque variable situationnelle a plusieurs niveaux, dont un au moins est valide à chaque point dans le temps. Par exemple, le type de route emprunté peut toujours être caractérisé. On utilise généralement les catégories suivantes : milieu rural, urbain, ou autoroute. De même, il est toujours possible de caractériser la météorologie pour un instant donné (ensoleillé, sec, pluie, neige, etc.). Une variable situationnelle peut aussi être déduite d'une mesure directe ou dérivée (situation de suivi de véhicule détecté à l'aide d'un radar par exemple).

Ces différents aspects sont primordiaux pour l'analyse des données car ils permettent de prendre en compte des variables dites de confusion. Lorsque l'on étudie la consommation de carburant par exemple, celle-ci dépend évidemment de la limitation de vitesse en vigueur, et il sera donc nécessaire de stratifier l'analyse selon les catégories de limitation de vitesse (ou selon le type de route, selon le choix de l'analyste). Cependant, un autre facteur de confusion peut provenir par exemple de la température. Il pourra donc être nécessaire de croiser le type de route avec la température pour une analyse plus fine. C'est l'objectif des segments définis ci-dessous.

6.2.3.5.2. Segments

Un segment indique une situation de conduite particulière issue du croisement de plusieurs variables situationnelles.

Ainsi, un segment correspond à une combinaison de niveaux spécifiques de certaines variables situationnelles. Un segment peut se raffiner à loisir :

- Un conducteur suivant un autre véhicule.
- Un conducteur suivant un autre véhicule sur l'autoroute.
- Un conducteur suivant un autre véhicule sur l'autoroute en plein jour.

- Un conducteur suivant un autre véhicule sur l'autoroute en plein jour avec un passager.
- ...

Chaque segment dans la liste ci-dessus devient donc de plus en plus spécifique, et il est de moins en moins probable de l'identifier dans les données. Mais d'autre part la variation entre les différentes instances de cette situation est susceptible d'être plus faible. Pour certaines hypothèses très spécifiques, il peut être nécessaire de rechercher des situations très particulières, tandis que pour d'autres hypothèses plus larges des situations plus générales pourraient être suffisantes.

On pourra par exemple s'intéresser aux phases de décélération avant une intersection, ou aux phases d'approche d'une infrastructure à risque.

6.2.3.5.3. Evènements

Un événement correspond à une situation de conduite dépendant non seulement d'une ou plusieurs variables situationnelles, mais aussi d'un facteur lié au comportement du conducteur (identifié grâce aux mesures directes ou dérivées).

Les événements sont des situations spécifiques de la conduite qui sont déduites des mesures directes ou dérivées, concernant le plus souvent l'état du véhicule ou sa dynamique, et donc induites par des actions du conducteur.

Les événements généralement étudiés sont par exemple les excès de vitesse (comparaison entre vitesse instantanée et variable situationnelle donnant la vitesse limite), les manœuvres de dépassement, le suivi de véhicule à distance critique, les freinages brusques, ou les coups de volants.

Il est intéressant de parler d'évènements plutôt que de mesures dérivées, pour la bonne raison que ces évènements sont généralement d'une durée réduite, et peuvent être considérés comme des événements ponctuels. Ils servent souvent à construire des indicateurs de performance en calculant des fréquences (fréquence de dépassement, d'excès de vitesse etc.). Ils peuvent aussi être caractérisés et étudiés séparément (ampleur et durée des excès de vitesse par exemple).



Figure 11 : Illustration de la notion d'évènements et de variable situationnelle pour la conduite.

6.2.3.6. Indicateurs de performance

Les définitions précédentes permettent de formaliser les différentes caractéristiques des situations de conduites enregistrées, et ainsi d'identifier des périodes temporelles d'intérêt par rapport au système étudié. L'objectif d'une étude du comportement de conduite en conditions réelles est de pouvoir comparer ces périodes d'intérêt avec des périodes de références. Toute tentative de comparaison entre périodes temporelles hétérogènes nécessite une harmonisation par rapport à une base de temps commune (normalisation). Cela passe donc par le calcul d'indicateurs caractéristiques de la période étudiée. Les indicateurs calculés sont destinés à permettre de répondre à des hypothèses posées préalablement à l'étude et destinées à être testées (approche détaillée dans les sections suivantes). Ces indicateurs sont donc en lien avec le système étudié et son fonctionnement.

Un indicateur de performance est une valeur continue ou catégorielle, relative à une période temporelle donnée, et obtenue à partir d'une ou plusieurs mesures d'une façon préalablement fixée.

- Les indicateurs de performance dépendent des hypothèses scientifiques prédéfinies que l'on souhaite tester. Chaque hypothèse est donc associée à un ou plusieurs indicateurs.
- Les indicateurs sont supposés caractériser le phénomène concerné par l'hypothèse à tester.
- Les différentes valeurs d'un indicateur sont destinées à être comparées entre elles, afin d'en étudier statistiquement les différences.

Lorsque l'on étudie un grand nombre de trajets routiers, on utilise généralement des indicateurs normalisés en temps (par heure), ou en distance (par 100km). Une période temporelle étant sélectionnée, on pourra par exemple compter la fréquence de survenue d'un évènement, ou simplement extraire des valeurs caractéristiques d'une mesure (maximum, minimum, variance, moyenne, pourcentage, etc...).

6.2.4. Protocoles d'évaluation des systèmes utilisant la conduite en situation naturelle

Les données de conduite en situation naturelle sont des enregistrements effectués de manière discrète mais continue sur un véhicule, pendant les trajets journaliers d'un conducteur. L'usage des moyens technologiques permettant cette solution s'étant démocratisé, l'enregistrement de données de ce type est devenu la norme (accès au bus CAN du véhicule notamment). Ce type de mesures permet d'accéder de façon rétrospective à des sections ou événements particuliers des données. La façon dont ces objets sont extraits des données, et les types de comparaisons que l'on en fait, sont potentiellement infinis. Chaque façon de faire définit un protocole expérimental différent, ayant des implications méthodologiques différentes. Il est donc utile de préciser les grands protocoles usuels dans les études en transport auxquels on peut se ramener pour disposer d'un cadre théorique. En effet, l'extraction d'information comparables dans un jeu de données naturalistes doit suivre un plan rigoureux afin de pouvoir extrapoler les résultats.

Rappelons qu'il existe deux grandes approches. La première approche Européenne a été de type FOT (Field Operational Test), consistant à procéder par comparaison entre une période d'enregistrement naturelle (aucun système utilisé), suivie d'une période d'utilisation d'un système précis (8 systèmes testés dans euroFOT).

- Projets Européens & Français incluant l'IFSTTAR :
 - LAVIA, FESTA (méthodologie), Scoref, EuroFOT, ecoDriver

Ensuite, les études de la conduite en situation naturelle « complètes » (NDS, Naturalistic Driving Studies) ne visent pas l'étude d'un système en particulier et tous types de comportements de conduite sont enregistrés pour analyse ultérieure. Ce type d'étude est utilisé depuis longtemps aux USA pour étudier la sécurité routière pure, et depuis peu en Europe.

- Projets Américains :
 - 100 hundred cars, SHRP2 (3000 véhicules sans aides à la conduite)
- Projets Européens & Français incluant l'IFSTTAR :
 - UDRIVE, INTERACTION, DaCoTa, S-Vrai, MEDOC

6.2.4.1. Etude transversale ("cross-sectional study")

Une approche fréquente dans les études en transport, consiste à collecter de l'information de façon ponctuelle, à différents pas de temps. Par exemple, on peut étudier une route particulière pendant plusieurs mois, pendant lesquels on comptera le nombre d'accidents observés, les conditions de circulation associées, et les caractéristiques de l'infrastructure. On pourra alors utiliser un modèle linéaire de type Régression, afin d'explorer les liens entre la variable d'intérêt (ici le nombre d'accidents) et les mesures d'exposition au risque (trafic, météo, infra, ...).

Ce type de protocole est adapté lorsque les facteurs étudiés ne changent pas beaucoup au cours du temps. C'est le cas des caractéristiques de la route elle-même. Le nombre d'accidents observés étant généralement très faible, la durée de ce type d'études est en général de plusieurs années.

Il est très difficile d'employer ce genre d'approche pour l'étude des systèmes d'assistance à la conduite.

6.2.4.2. Etude de cas-témoins ("case-control study")

Ce type d'étude cherche d'abord à identifier des cas d'intérêt pour la question étudiée. Les « cas » sont en général des événements. Il pourra s'agir soit d'une interaction avec le système d'intérêt, soit d'un événement lié à la sécurité dont l'occurrence pourrait être liée à l'usage du système (vitesse trop élevée si oubli du régulateur), soit de tout autre événement de type incident, ou presque accident.

Les événements d'intérêts, les cas, sont donc déterminés au préalable dans cette approche. Ils sont extraits des données, puis comparés à des témoins c'est à dire des situations où l'événement n'a pas eu lieu. Les facteurs d'exposition sont ensuite extraits pour évaluer leur impact sur la survenue de l'événement.

Dans notre contexte, les cas seront par exemple les conducteurs ayant vécu un événement, ou l'événement lui-même (freinages brusques par exemples), Les témoins seront les conducteurs n'ayant pas vécu cet événement, ou des séquences de conduite sans événement. Les facteurs (conditions de conduite) permettant de modéliser la probabilité de l'événement.

Le choix du groupe de contrôle (ou baseline) est déterminant pour la validité de l'étude. Le principe général est de considérer que le groupe de contrôle doit refléter les caractéristiques de la population source qui a conduit aux cas. L'implémentation pratique de ce principe dépend cependant largement du contexte. Il faut cependant noter que dans le cas d'une étude de cohorte (cf. ci-dessous), les cas sont toujours issus de la population étudiée, ce qui diminue le biais.

En effet, les études cas-témoins sont très sensibles au choix des témoins qui peuvent facilement biaiser l'étude. Un autre inconvénient est que cette approche n'autorise pas facilement une évaluation directe du risque (risque absolu d'un événement).

6.2.4.3. Etude de cohorte ou de suivi ("cohort study" or "follow-up study")

Dans une étude de cohorte, c'est l'exposition au risque qui est déterminée en premier. Le suivi longitudinal des conducteurs permet de déterminer ultérieurement la survenue ou non d'un événement (accident ou non). Une cohorte est donc un ensemble d'individus exposé en moyenne aux mêmes risques. On compare deux cohortes en observant le nombre d'événements survenus.

Par exemple dans notre contexte, on pourra comparer une cohorte de conducteurs adultes à une cohorte de jeunes conducteurs. Ces conducteurs seront suivis plusieurs années (période de l'étude) ce qui permettra d'identifier le résultat (nombre d'accidents observés).

Lorsque les facteurs étudiés sont variables au cours du temps (météo, trafic, etc.) on parle de cohorte dynamique. En réalité il s'agit du cas le plus courant des études NDS ou FOT.

Une cohorte dynamique fait référence à un ensemble d'individus exposés aux mêmes facteurs de risques, à un instant (période) donné.

Ce type d'étude est en général moins biaisé que les précédentes. En outre, la direction de l'étude, de l'exposition au risque jusqu'à la survenue de l'événement, permet de procéder directement à l'évaluation du risque.

6.2.4.4. Spécificités des études de type FOT ou NDS

Le processus de collection des données dans une étude de ce type est prospectif, et similaire au cas d'une étude de cohorte. C'est-à-dire que l'on suit pendant un temps long, suffisant pour espérer observer des événements rares, une cohorte de conducteurs exposés à différents facteurs dynamiques (conditions générales de route, activation momentanée d'un système) ou fixés (activation permanente d'un système ou non).

Dans le cas où l'on cherche à étudier la survenue de certains événements, il faut alors les extraire des données, suivis de périodes de contrôle. On est alors dans le cas d'un processus de type cas-témoin.

L'hybridation de ces deux approches conduit à parler dans ce cas d'études de cas-cohorte (case-cohort). Il s'agit d'une étude en deux étapes : La première étape consiste à enregistrer l'ensemble des données renseignant sur l'évolution des facteurs d'exposition au cours du temps. La deuxième étape est une analyse a posteriori des données, consistant à identifier les événements d'intérêts, puis choisir un groupe de contrôle. Ce type d'étude combine les mérites des études de cohortes et de cas-témoins.

6.2.4.5. Implications sur les méthodes d'analyse

Selon que l'on considère l'un ou l'autre aspect des études FOT & NDS (i.e. aspect cohorte ou cas-témoin), les méthodologies déployées sont différentes.

Les études fondées sur l'analyse de certains événements font appel à des méthodes d'évaluation du risque assez classiques. Le Tableau 5 page 105 détaille les algorithmes actuellement utilisés pour détecter les événements de la conduite les plus classiques dans les FOT & NDS. Même si il est difficile d'évaluer un risque absolu (fréquence moyenne de survenue de l'événement, dépend de la représentativité de l'échantillon), ce type d'approche permet en revanche l'évaluation d'un risque relatif, ou d'un rapport de côtes (odds-ratios).

Pour l'analyse, on fera dans ce cas appel à des modèles linéaires généralisés de type régression logistique, ou modèles mixtes si la corrélation intrinsèque des données doit être prise en compte (cf. détails page 90).

Ce type d'analyse pose plusieurs questions théoriques pour analyser correctement les données. Cela repose aussi sur la pertinence de l'événement choisi pour représenter le phénomène étudié. De nombreux aspects de cette approche sont encore de l'ordre de la recherche et ne sont pas abordés en détail dans ce document.

En revanche, l'aspect « cohorte » de l'étude rend possible une analyse complète des données (c.ad. utilisant l'ensemble des mesures, et non plus seulement une sélection de cas-témoins). C'est que l'on appelle « méthode agrégée », fondée sur un principe de réduction systématique et automatisable des données. Une fois réduites, les mesures sont sous la forme d'une table agrégée adaptée à un grand nombre d'analyses. Cette méthode est détaillée à la section suivante.

Enfin, il faut noter qu'une fois les données collectées pendant la première étape de l'étude, il est virtuellement possible de mener tout type d'analyse : séries chronologiques, analyse fonctionnelle, data mining, segmentation etc.

6.2.5. Méthode d'analyse par question de recherche

6.2.5.1. Introduction

Ce type d'analyse fondée sur une agrégation pertinente des données est appelé méthode agrégée. Cette approche permet d'étudier les changements intervenus dans la conduite entre une période de contrôle et une période de test. On cherche donc à étudier les performances de la conduite selon différentes situations. Cette approche est privilégiée par le document FESTA qui s'est fortement inspiré de l'expérience acquise par les chercheurs lors des projets récents de test à grande échelle (Field Operational Tests, projet euroFOT notamment). Les résultats obtenus ont

contribués à valider cette approche qui est maintenant reprise dans la majorité des études liées à l'impact des systèmes d'assistance.

On peut par exemple chercher à étudier comment l'usage d'un système embarqué influe sur la vitesse moyenne. On fait alors l'hypothèse que le système utilisé ne modifie pas la vitesse moyenne, et on cherche à la mettre en défaut avec les données enregistrées. Dans ce cas précis, l'indicateur de performance utilisé est la valeur moyenne de la vitesse. Plus sa variation sera positive et significative pendant la période de test, plus cette hypothèse pourra être mise en doute.

Ces indicateurs sont en général liés au phénomène que l'on veut étudier, et sont construits avec la connaissance des experts. Cependant, tous requièrent la définition d'une base de comparaison commune, que ce soit le temps (valeur rapportée à une unité de temps) ou l'espace (rapportée à la distance parcourue). Cela consiste à définir un dénominateur commun.

Les événements sont ponctuels mais ils peuvent être comparés lorsqu'une base commune est spécifiée. C'est le cas des indicateurs liés à l'occurrence de certains événements (nombre d'activation du système/100km par ex).

Si les indicateurs sont bien choisis, une évolution significative de ces derniers entre deux conditions signifiera un probable changement du comportement associé.

Des précautions méthodologiques importantes doivent être prises dans l'interprétation des résultats, qui repose en grande partie sur la capacité de l'indicateur utilisé à refléter le comportement d'intérêt. Dans le cas où plusieurs indicateurs sont étudiés, il faut veiller à la corrélation statistique si ils sont liés au même comportement.

Afin de procéder à une analyse statistique détaillée, il est recommandé de ne pas utiliser qu'un seul facteur explicatif (contrôle et traitement par exemple). Dans notre exemple précédent, la vitesse moyenne n'est pas seulement modifiée par l'usage du système, mais aussi par d'autres facteurs que l'on souhaite contrôler. Il s'agit en effet de dégager l'effet du système relativement à tous les autres effets principaux envisagés. Il faudra donc ici tenir compte par exemple de l'effet du trafic ou du type de route. Cet effet peut être codé de façon qualitative (sous forme de catégories, type de route par ex.) ou bien de façon quantitative (un autre indicateur, l'inter distance moyenne par exemple). Ces effets principaux doivent être spécifiés avant de procéder à l'extraction des indicateurs.

Pour appliquer cette méthode il faut avant tout effectuer les étapes suivantes :

1. Définir la question de recherche
2. Définir un indicateur agrégé en lien avec la question posée.
3. Définir les principaux facteurs ayant un effet potentiel sur le phénomène étudié et entrant dans le modèle d'analyse.

Exemple d'hypothèse de recherche :

La vitesse moyenne lors des périodes de vitesse stable sera plus faible lorsque le conducteur utilise un système d'aide à l'éco-conduite comparé à la condition de contrôle.

Autre formulation préférable :

La vitesse moyenne lors des périodes de vitesse stable sera identique lorsque le conducteur utilise un système d'aide à l'éco-conduite comparé à la condition de contrôle.

6.2.5.2. Etapes du plan d'analyse

Une fois les principes méthodologiques la question de recherche bien fondés (validité de l'indicateur, validité des comparaisons envisagées etc.), la personne en charge de l'analyse doit suivre différentes étapes afin de disposer d'un jeu de données adaptées à la modélisation et au test d'hypothèse.

Ces différentes étapes commencent par l'accès aux données, et finissent par les méthodes statistiques à utiliser pour leur analyse. Elles sont décrites à la Figure ci-dessous.

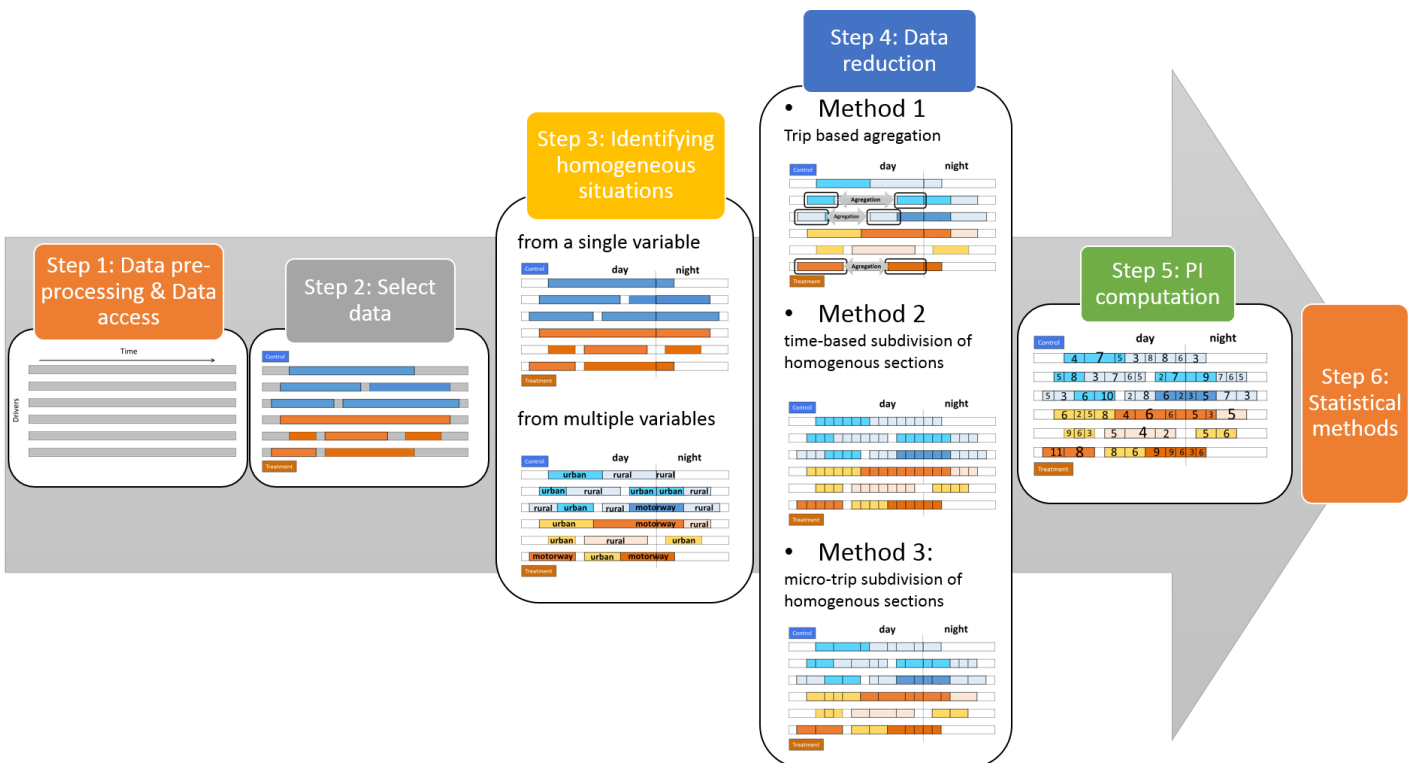


Figure 12: Synopsis de la méthode d'analyse d'une question de recherche par test d'hypothèse.

6.2.5.2.1. Etape 1 : Accès et pré-traitement des données

Une fois les données brutes collectées, une étape de prétraitement des mesures est nécessaire afin de mener les opérations suivantes :

- Contrôle de la qualité des mesures
 - Identifier les valeurs hors normes, les valeurs aberrantes, les erreurs de capteurs etc.
- Filtrage
 - Les données corrompues ne doivent pas être utilisées dans l'analyse et doivent donc être identifiées dans la base.
- Enrichissement
 - Les mesures brutes sont utilisées pour extraire les caractéristiques de la route empruntée à l'aide d'une cartographie (ex : Open Street Map)
 - Les mesures manquantes dans la carte doivent être identifiées
- Extrapolation des mesures sur une base de temps uniforme
 - Cette étape optionnelle facilite grandement le traitement des données.
 - La qualité de l'extrapolation doit être examinée avec précaution.

En fin d'étape 1, nous disposons de données extrapolées sur une base de temps commune (1Hz ou 10Hz, ou mieux si les données le permettent), avec l'ensemble des variables renseignées. Cette étape est résumée à la Figure suivante :



Figure 13: Illustration pour une série de trajets effectués par 6 conducteurs différents sur une période de temps identique. Les variables à valeur manquantes ne peuvent être indiquées globalement.

Etape 1 :

Si l'on part de l'exemple précédent d'hypothèse de recherche, cette étape consiste à enrichir les données en calculant la variable d'état « vitesse stabilisée » (cruising), qui identifie les instants de conduite où cette condition est vérifiée (définition dépendant de l'analyste).

6.2.5.2.2. Etape 2 : Sélection des données

En fonction de l'analyse souhaitée, une sélection préalable des données peut être nécessaire. Cela peut être nécessaire lors qu'un capteur est défaillant, ou tout simplement pour présélectionner une condition de conduite.

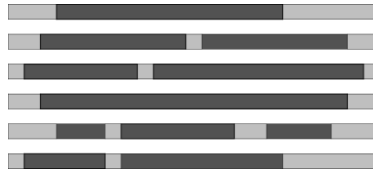


Figure 14: Illustration de l'étape de sélection des données.

Etape 2 :

Pour notre exemple, cette étape consiste à sélectionner uniquement les instants en situation de vitesse stable. Il faut s'assurer que la donnée de vitesse instantanée est de bonne qualité (non aberrante) pour pouvoir calculer l'indicateur de vitesse moyenne.

6.2.5.2.3. Etape 3 : Identification de situations homogènes

Une fois les données sélectionnées, il faut identifier dans celle-ci la valeur des principaux facteurs explicatifs du comportement. Nous pouvons avoir les cas suivants :

1. Les variables explicatives sont toutes qualitatives (i.e. uniquement des variables situationnelles, SV)
 - a. Cette approche permet notamment d'effectuer une analyse de la variance, ou par table de contingence.
2. Les variables explicatives sont toutes quantitatives (i.e. uniquement des mesures brutes ou dérivées)
 - a. Cette approche permet d'utiliser une approche par régression linéaire multiple.
 - b. Cette approche ne nécessite pas d'identifier des situations homogènes.
3. Les variables sont soit qualitatives, soit quantitatives (croisement de SV et prise en compte de mesures dérivées sous forme de covariables)
 - a. Cette approche permet d'utiliser les modèles suivants : Analyse de la covariance, Binomiale Logit Mixtes, Régression logistique, Modèles log-linéaires, Modèles à réponses multinomiales.

Rappelons que toute mesure continue peut être transformée en mesure qualitative. La majorité des facteurs influençant l'usage d'un système sont soit liés à une caractéristique de l'infrastructure

(courbure aigue, type de route, etc.), soit lié à une condition de route particulière (météo), soit lié à l'occurrence d'un événement (interaction avec l'autoradio). Ce sont donc en général des mesures qualitatives.

L'ensemble des indicateurs numériques est calculé sur une série de sections de données identifiées en croisant les principaux facteurs qualitatifs. L'ajout de mesures quantitatives au modèle (cas n°3 précédent) reste possible par l'ajout d'autres indicateurs au modèle, bien entendu non corrélés avec l'indicateur principal que l'on cherche à observer/expliciter.

Nous souhaitons donc identifier des sections de données en fonction de un ou plusieurs facteurs qualitatifs.

Une seule variable qualitative

Le cas le plus simple consiste à étudier les variations d'un indicateur en fonction d'un seul facteur qualitatif. L'objectif est par exemple de mener une analyse de variance à un facteur.

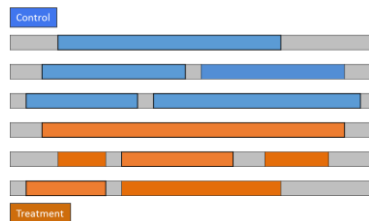


Figure 15: Illustration de la forme la plus simple d'identification des situations homogènes

Etape 3 :

Pour notre exemple, cette étape consiste à identifier les phases de contrôle et les phases de traitement. Cette valeur est binaire, et reste constante au cours d'un trajet.

Plusieurs variables qualitatives

Il est souvent utile de mener une analyse plus détaillée en croisant plusieurs facteurs indépendants. Cela implique un découpage plus fin des données.

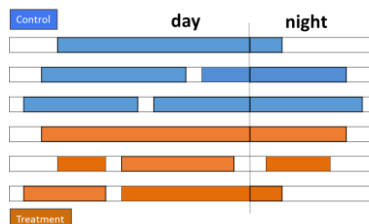


Figure 16: Illustration de l'étape d'homogénéisation pour deux facteurs. La variable "Contrôle/Traitement" est combinée avec la variable luminosité (jour/nuite) afin d'identifier quatre conditions différentes dans les données.

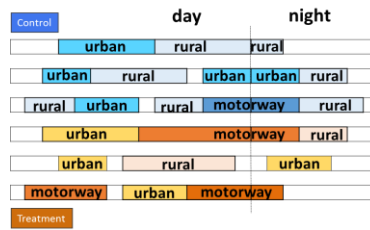


Figure 17: Illustration de l'étape d'homogénéisation pour trois facteurs. La variable "Contrôle/Traitement" est combinée avec la variable luminosité (jour/nuit) et type de route (3 niveaux) afin d'identifier les conditions suivantes : urbain-Jour, rural-Jour, Autoroute-Jour, urbain-nuit, rural-nuit, Autoroute-nuit, dans chaque modalité de contrôle/Traitement

Etape 3 :

Si l'on fait l'hypothèse que la vitesse moyenne en période de vitesse stable dépend aussi du type de route et de la luminosité, alors il faut identifier ces conditions dans les données.

6.2.5.2.4. Étape 4 : méthodes de réduction des données

Une fois identifiées les différentes combinaisons de facteurs intéressantes pour l'analyse (étape précédente), ces sections homogènes sont utilisées pour calculer les indicateurs numériques. Les choix méthodologiques effectués à cette étape peuvent avoir un impact sur l'analyse.

Différentes approches sont possibles et sont décrites ci-dessous.

Méthode 1 : Agrégation par trajet

Cette méthode est conseillée lorsque le phénomène étudié est relativement stable dans le temps, ou bien lorsque l'indicateur associé est plus pertinent lorsqu'il est calculé sur une longue période de temps (ex : consommation moyenne de carburant). L'objectif est donc ici de fusionner des sections identiques (en termes de combinaison de facteurs, et donc de situations homogènes) au sein d'un même trajet avant de calculer l'indicateur.

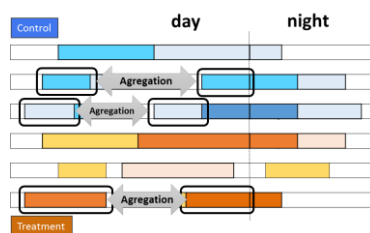


Figure 18: Illustration de la méthode d'agrégation par trajet

Étape 4 : Méthode 1

Supposons que la vitesse moyenne stabilisée dépend principalement de la variable “Contrôle/Traitement” et de la variable luminosité (jour/nuit). Chaque trajet sera donc résumé en 4 valeurs de l’indicateur pour les conditions homogènes suivantes : Contrôle-Jour, Contrôle/Nuit, Traitement-Jour, Traitement-Nuit,



→ Différentes distances et durées

→ Les données doivent être pondérées pendant l’analyse

Méthode 2 : Agrégation sur une base de temps homogène

Cette méthode est recommandée lorsque le phénomène étudié évolue fréquemment au cours d’un trajet. Par exemple, la vitesse est modifiée plusieurs fois au cours d’un trajet, étant contrainte par l’infrastructure elle-même. Si cet indicateur est calculé à l’échelle d’un trajet, il est peu probable d’observer de grandes variations.

Cette méthode offre la possibilité de spécifier un paramètre de temps (typiquement entre quelques secondes et quelques minutes) qui servira à découper des sections homogènes en fonction des facteurs, en sous-sections homogènes en fonction du temps. Les sections d’une durée inférieure au paramètre spécifié n’entrent alors pas dans l’analyse.

Cette approche est adaptée à une analyse ayant un référentiel temporel plutôt que spatial. On peut cependant noter que cette approche peut être adaptée à une analyse spatiale, en prenant comme paramètre une distance parcourue minimale.

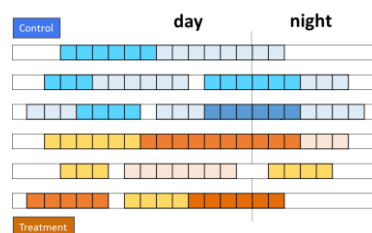



Figure 19: Illustration de l’agrégation sur une base de temps homogène.

Étape 4 : Méthode 2

Supposons que l’homogénéité a été garantie pour les variables “Contrôle/Traitement” et luminosité (jour/nuit). Chaque condition est alors subdivisée en sous sections d’une durée maximale de 30 secondes pour le calcul des indicateurs.

Chaque trajet est alors résumé par un maximum de (durée totale du trajet / 30 sec) valeurs différentes de l'indicateur de vitesse moyenne.



- Pondération inutile si analyse en temps
- Toujours nécessaire si analyse spatiale (sauf si variante spatiale de la méthode utilisée)

Méthode 3 : Agrégation selon une variable particulière

Il est possible que les comportements étudiés soient complexes, et que l'analyste souhaite subdiviser différemment les données. Par exemple, une approche par « micro-trajet » cherche à extraire des sous sections correspondant à un intervalle entre deux arrêts.

Cette approche est un compromis entre les précédentes, lorsqu'il existe suffisamment d'évidence quant à la pertinence de cette variable pour identifier des conditions comparables.

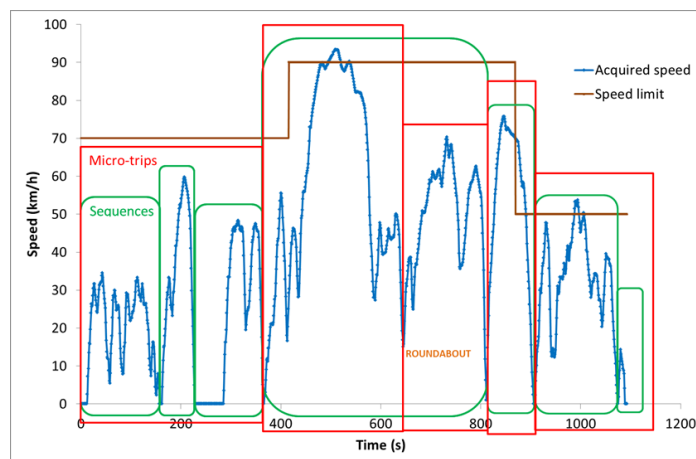


Figure 20: Exemple d'identification par micro-trajet.

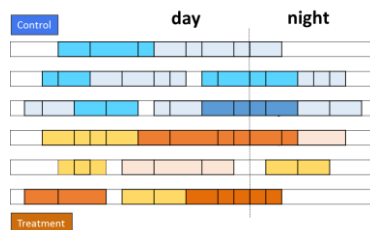



Figure 21: Illustration de la réduction de données par micro-trajets..



- Durée et distance différentes entre les sections
- Pondération nécessaire

6.2.5.2.5. Etape 5 : Calcul des indicateurs

Une fois que les sections homogènes et comparables sont identifiées (étape 3), et que la stratégie de réduction des données est définie (étape 4) il est possible de procéder au calcul des indicateurs. Chaque sous-section est donc réduite sous la forme d'une série de variables qualitatives ou quantitatives.

Ainsi, chaque facteur principal utilisé à l'étape 3 possède une valeur unique. Chaque indicateur numérique est calculé et sa valeur caractérise la sous-section. En procédant ainsi pour l'ensemble des trajets, on construit une table de données dont chaque ligne représente une sous-section remplie avec les valeurs des facteurs qualitatifs et des indicateurs numériques. Cette table est maintenant prête pour de nombreuses analyses statistiques.

Control	day						night									
	4	7	5	3	8	8	6	3								
	5	8	3	7	6	5	2	7	9	7	6	5				
	5	3	6	10	2	8	6	2	3	5	7	3				
	6	2	5	8	4	6	6	5	3	5						
	9	6	3	5	4	2			5	6						
	11	8		8	6	9	9	6	3	6						
Treatment																

Figure 22: Illustration d'un calcul d'indicateur sur sections inhomogènes. La taille de la valeur indiquée représente la pondération qu'il faut appliquer lors de l'analyse.

6.2.5.2.6. Etape 6 : Analyse statistique

La table finale obtenue à cette étape est fortement dépendante des choix effectués. Si plusieurs facteurs sont croisés à l'étape 3, une première analyse portant uniquement sur les variables catégorielles peut être menée (analyse d'un tableau de contingence).

Parmi les approches possibles pour comparer deux variables catégorielles on citera :

- Test d'association du Khi-deux de Pearson
- Mesures d'association (coefficient d'incertitude, corrélation de spearman)
- Calculs de risques relatifs, d'odd-ratios
- Si introduction d'une troisième variable : Analyses stratifiées et odd-ratios ajustés.

Ce type d'analyse ne permet cependant pas de modéliser une variable numérique en fonction d'autres facteurs. Pour aller plus loin, et afin de décrire globalement les autres approches possibles, il faut se placer dans le cadre du modèle linéaire généralisé.

Le modèle linéaire généralisé constitue un cadre général permettant d'englober une grande quantité des modèles disponibles. Cette famille de modèles permet d'étudier la liaison entre une variable dépendante ou réponse Y et un ensemble de variables explicatives ou prédicteurs X_1, \dots, X_K . Elle englobe le modèle linéaire général (régression multiple, analyse de la variance et

analyse de la covariance), le modèle log-linéaire et des techniques de modélisation telles que la régression logistique ou la régression de Poisson. Les modèles linéaires généralisés sont formés de trois composantes :

- La variable de réponse Y ; composante aléatoire à laquelle est associée une loi de probabilité.
- Les variables explicatives X_1, \dots, X_K utilisées comme prédicteurs dans le modèle, définissent sous forme d'une combinaison linéaire la composante déterministe.
- Le lien décrivant la relation fonctionnelle entre la combinaison linéaire des variables Y_1, \dots, Y_n et l'espérance mathématique de la variable de réponse Y .

Le choix du modèle linéaire généralisé dépend de la nature des données que l'on souhaite étudier. Le tableau ci-dessous résume ces différents cas.

Tableau 4 : Vue d'ensemble des différents modèles linéaires généralisés les plus courants.

Réponse (composante aléatoire)	Lien	Variables explicatives (composante déterministe)	Modèle
Normale	Identité	Quantitatives	Régression
Normale	Identité	Qualitatives	Analyse de la variance
Normale	Identité	Mixtes	Analyse de la covariance
Binomiale	Logit	Mixtes	Régression logistique
Poisson	Log	Mixtes	Modèles log-linéaires
Multinomiale	Logit généralisé	Mixtes	Modèles à réponses multinomiales

La méthode FESTA ainsi que la méthodologie décrite plus haut est essentiellement basée sur le test d'hypothèses prédéfinies. Cette approche est très classique en statistique et bien documentée. De nombreux tests peuvent être utilisés en fonction de la nature des hypothèses testées : test de Student, test de Fisher, analyse de la variance, test du χ^2 , test de Kolmogorov-Smirnov, etc.

Dans le cas des problématiques liées au transport, il est courant de chercher à tester l'influence de plusieurs facteurs sur un indicateur prédéfini. Dans le cas d'une étude de la vitesse moyenne, on cherchera par exemple à étudier les liens existant avec les facteurs suivants : disponibilité du système (binaire), limitation de vitesse en vigueur (catégorielle), âge du conducteur (continue ou catégorielle), conditions climatiques (catégorielle), heure de pointe (binaire), etc.

Les analyses menées sont donc plutôt du type analyse de la variance ou d'une analyse de la covariance. Lorsque l'indicateur utilisé est de type binaire (présence ou non d'un évènement, un excès de vitesse par exemple), l'analyse relèvera d'une approche logistique.

Cependant, un problème majeur du modèle linéaire généralisé classique, est de supposer l'indépendance des données. Cette hypothèse n'est en général pas vérifiée dans le cadre des données de conduite en situation naturelle. En effet, l'attitude générale d'un conducteur vis à vis de la conduite, qu'il soit "sportif" ou "calme", est un invariant susceptible d'introduire une corrélation dans les données dont il faut tenir compte. Cette corrélation peut être aussi plus forte au sein d'un même trajet, qu'entre deux trajets différents de la même personne.

Des modèles plus spécifiques sont donc nécessaires, parmi lesquels les modèles marginaux (Generalized estimated equations, GEE) introduits par Liang et Zeger (1986). Un autre modèle fréquemment utilisé est celui des modèles mixtes (Generalized Linear Mixed Models, GLMM). Ces modèles, fréquemment comparés dans la littérature, permettent tous deux de tenir compte de la corrélation intrinsèque aux données mais d'une façon différente. Le modèle mixte sera choisi pour ses bonnes performances lorsque l'on dispose d'un grand nombre d'observations par sujet.

La section suivante décrit plus en détail les différences méthodologiques entre l'approche classique (seul le cas de la régression logistique est abordé ici) et l'approche des modèles marginaux ou mixtes.

6.2.6. Modèles statistiques pour données catégorielles non indépendantes

6.2.6.1. Régression logistique

Lorsque l'indépendance des données est supposée, la méthode usuelle pour étudier la relation entre une variable réponse binaire et plusieurs variables explicatives est la régression logistique. On introduit les notations suivantes :

Soit T_i le nombre d'observations réalisées sur le sujet i on note $Y_i = (Y_{i1}, \dots, Y_{iT_i})$ le vecteur des observations, ces observations étant supposées indépendantes. Autrement dit, on note :

$$Y_{ij} = \begin{cases} 1 & i = 1, \dots, I; j = 1, \dots, T_i \\ 0 & \end{cases} \quad (1.1.)$$

La variable aléatoire Y_{ij} suit une loi de Bernoulli de paramètres p_{ij} où $p_{ij} = P(Y_{ij} = 1)$. Le modèle logistique ordinaire s'écrit :

$$\text{logit}(p_{ij}) = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \beta' X_{ij} \quad (1.2.)$$

$$\text{Var}(Y_{ij}) = \mu_{ij}(1 - \mu_{ij}) \text{ (fonction de variance)} \quad (1.3.)$$

Où $\mu_{ij} = p_{ij}$, X_{ij} est le vecteur des variables explicatives et β est le vecteur des paramètres. Dans ce cas, le vecteur β des paramètres est estimé par la méthode du maximum de vraisemblance. L'estimateur de β est solution des équations de vraisemblances (équations du score) obtenues en dérivant la fonction de log-vraisemblance (notée l) par rapport à β :

$$U(\beta) = \frac{\partial l}{\partial \beta} = \sum_i D'_i \Sigma_i^{-1} (Y_i - \mu_i) = 0 \quad (1.4.)$$

Où $D_i = \frac{\partial \mu_i}{\partial \beta'}$ (matrice diagonale), Σ_i est la matrice des variances-covariances (matrice diagonale $\Sigma_i = \text{diag}(\text{Var}(Y_{ij}))$), Y_i est le vecteur des Y_{ij} , et μ_i est le vecteur des μ_{ij} . Les équations ci-dessus sont appelées "independence estimating equations"(IEE) et sont résolues par des méthodes itératives : algorithme du scoring de Fisher, méthode itérative des moindres carrés pondérés (IWLS) ou algorithme de Newton modifié (Quasi-Newton). L'estimateur obtenu $\hat{\beta}$ est alors consistant (i.e. cv en probabilité vers β) et distribué asymptotiquement selon une loi normale avec pour matrice de covariance $V(\hat{\beta}) = (\sum_i D'_i \Sigma_i^{-1} D_i)^{-1}$.

Cependant, dans notre cas, les trajets effectués par un même conducteur (ou les sous-sections de trajets) sont corrélés et ne peuvent donc pas être supposés indépendants. En effet, les trajets effectués par un même conducteur ont les mêmes caractéristiques car le comportement de conduite reste similaire. Ce problème de données corrélées apparaît dans de nombreuses situations : mesures répétées dans le temps appelées données longitudinales (principalement en épidémiologie), mesures issues d'une même famille ou cluster. Dans le domaine des transports, les observations issues d'expérience de type "naturalistic driving" appartiennent clairement à cette dernière catégorie : les observations issues d'un même conducteur sont, en général, fortement corrélées. L'hypothèse d'indépendance étant violée, on ne peut pas utiliser les modèles classiques comme les modèles linéaires généralisés dont la régression logistique est un cas particulier, et qui reposent sur l'indépendance des observations.

Il existe principalement deux méthodes permettant de prendre en compte la liaison entre observations réalisées sur un même sujet :

1. les modèles marginaux ou GEE (Generalized Estimating Equations)
2. les modèles mixtes

Ces deux méthodes sont décrites dans les sections suivantes

6.2.6.2. Equations d'estimation généralisées (GEE)

Les modèles marginaux basés sur le concept d'équations d'estimations généralisées ont été introduits par Liang et Zeger (1986) pour analyser les données longitudinales (mesures répétées). Ils sont l'analogie des modèles de quasi-vraisemblance pour observations non indépendantes. En

effet, l'approche GEE ne spécifie pas de fonction de vraisemblance (basée sur l'indépendance des observations) mais utilise une quasi-vraisemblance pour laquelle il n'existe pas de modèle probabiliste approprié. Pour les modèles de quasi-vraisemblance, il est suffisant de donner les deux premiers moments (moyenne et covariance) de la variable réponse, par opposition aux modèles linéaires généralisés dans lesquels toute la distribution de la variable réponse doit être définie. Ainsi les GEE sont plutôt considérées comme une méthode d'estimation que comme une véritable méthode de modélisation.

Nous avons vu à la section précédente que sous l'hypothèse d'indépendance des observations

(i.e. $\text{Corr}(Y_{ij}, Y_{ij'}) = 0$), le modèle logistique s'écrivait sous la forme 1.2. Cependant, dans le cas de données corrélées (i.e. $\text{Corr}(Y_{ij}, Y_{ij'}) \neq 0$), on définit une matrice de variances-covariances de travail :

$$\Sigma_i = A_i^{\frac{1}{2}} R_i(\alpha) A_i^{\frac{1}{2}} \quad (1.5.)$$

Où A_i est la matrice diagonale dont les éléments diagonaux sont $\text{Var}(Y_{ij})$, et $R_i(\alpha)$ est la matrice de corrélation de travail. Cette matrice de variances-covariances de travail est généralement différente de la vraie matrice de variances-covariances S_i . Plusieurs choix sont possibles pour la structure de corrélation de travail :

- Indépendant : $\text{Corr}(Y_{ij}, Y_{ij'}) = 1$ pour tout j, j' . Cas où les observations réalisés sur un même sujet sont indépendantes.
- Echangeable (type "exchangeable")

$$\text{Corr}(Y_{ij}, Y_{ij'}) = \begin{cases} 1 & \text{si } j = j' \\ 0 & \text{si } j \neq j' \end{cases} \quad (1.6.)$$

Cas où les observations réalisés sur un même sujet sont corrélées de la même manière.

- Auto-regressive d'ordre 1 (type "AR(1)") : $\text{Corr}(Y_{ij}, Y_{ij'}) = \alpha^{|j'-j|}$.
Utilisé pour les mesures répétées dans le temps : la corrélation diminue à mesure que l'écart entre j' et j augmente.
- Non spécifiée (type "unstructured") : $\text{Corr}(Y_{ij}, Y_{ij'}) = \alpha_{jj'}$ pour $j \neq j'$.
Structure la plus générale mais beaucoup de paramètres à estimer. (choix possible uniquement si le nombre d'observations par sujet est constant).

L'estimateur de β est alors solution des équations de quasi-vraisemblance appelées équations d'estimation généralisées ou "generalized estimating equations" (GEE) :

$$\sum_i D'_i \Sigma_i^{-1} (Y_i - \mu_i) = 0 \quad (1.7.)$$

où Σ_i est défini par 1.5. Lorsque la matrice de corrélation de travail est la matrice identité, on retrouve l'équation 1.4 obtenue dans le cas d'indépendance des observations. L'estimation de β se fait à l'aide d'un processus itératif : algorithme de Fisher modifié. Le terme "modifié" signifie

que l'on utilise Σ_i à la place de la vraie matrice de variances-covariances S_i . L'intérêt de la méthode GEE repose sur la robustesse des estimations obtenues : les inférences sur

β sont correctes même si la matrice de variances-covariances Σ_i n'est pas correctement spécifiée. On utilise pour cela un estimateur "sandwich". Ainsi, sous des conditions de régularité faibles, $\hat{\beta}$ est asymptotiquement distribué selon une loi multinormale :

$$\hat{\beta} \xrightarrow{d} N(\beta, I_0^{-1} I_1 I_0^{-1}) \quad (1.8.)$$

Où $I_0 = \sum_i D'_i \Sigma_i^{-1} D_i$ et $I_1 = \sum_i D'_i \Sigma_i^{-1} S_i \Sigma_i^{-1} D_i$ évalué à $\beta = \hat{\beta}$ et $\alpha = \hat{\alpha}$ en estimant S_i par $(Y_i - \mu_i(\hat{\beta}))(Y_i - \mu_i(\hat{\beta}))'$.

Si la matrice de variances-covariances est correctement spécifiée (i.e. $\Sigma_i = S_i$), alors $I_0 = I_1$ et on retrouve le résultat classique :

$$\hat{\beta} \xrightarrow{d} N(\beta, I_0^{-1}) \quad (1.9.)$$

I_0 étant la matrice d'information de Fisher.

Pour un exposé plus détaillé des caractéristiques de ces modèles, on pourra se référer au livre de Diggle, Liang et Zeger (cf Diggle et al. [1994]). J'ai utilisé cette approche dans le cadre de mes collaborations en psychologie expérimentale (Viola Cavallo) et dans le cadre du projet euroFOT.

6.2.6.3. Modèles mixtes

Une autre approche pour modéliser des données corrélées est l'utilisation des modèles linéaires généralisés mixtes (GLMM). La méthode vue à la section précédente modélise les distributions marginales en traitant la corrélation comme un paramètre de nuisance. La méthode présentée ici introduit un effet aléatoire spécifique à chaque sujet. On modélise ainsi la variable réponse conditionnellement à ces effets aléatoires. L'inférence est donc individuelle ("subjects specific approach"), contrairement aux modèles marginaux présentés à la section précédente où l'on modélisait la moyenne de la population ("population-averaged approach").

Le modèle logistique mixte s'écrit :

$$\text{logit}[P(y_{ij} = 1|u_i)] = x'_{ij}\beta + z'_{ij}u_i \quad (1.10.)$$

Où x_{ij} et β sont définis dans le modèle 1.2, u_i est le vecteur des effets aléatoires associés au sujet i et z_{ij} est la matrice associée.

Un cas particulier important est le cas où $z_{ij} = 1$. Ce modèle est appelé modèle logistique avec effet aléatoire sur la constante ("random-intercept logistic model" ou "logistic-normal model") et s'écrit :

$$\text{logit}[P(y_{ij} = 1|u_i)] = x'_{ij}\beta + u_i \quad (1.11.)$$

où les u_i sont indépendants et $u_i \sim N(0, \sigma^2)$. Soit y le vecteur des observations et u le vecteur des effets aléatoires. La vraisemblance marginale ("marginal likelihood") s'écrit :

$$l(\beta, \sigma^2; y) = f(y; \beta, \sigma^2) = \int f(y|u; \beta) f(u; \sigma^2) du \quad (1.12.)$$

où $f(y|u; \beta)$ est la densité conditionnelle de y sachant u et $f(u; \sigma^2)$ est la densité normale de u . En particulier, la vraisemblance marginale du modèle 1.11 s'écrit :

$$l(\beta, \sigma^2; y) = \int \prod_i \prod_j \left[\frac{\exp(x'_{ij}\beta + u_i)}{1 + \exp(x'_{ij}\beta + u_i)} \right]^{y_{ij}} \left[\frac{1}{1 + \exp(x'_{ij}\beta + u_i)} \right]^{1-y_{ij}} f(u_i; \sigma^2) \quad (1.13.)$$

L'estimation des paramètres se fait par approximation puis maximisation de cette vraisemblance. Il existe pour cela différentes méthodes :

1. Quadrature de Gauss-Hermite

Cette méthode consiste à approximer l'intégrale 1.13 par des méthodes numériques puis à maximiser cette approximation de la vraisemblance. L'approximation est réalisée par la formule de la quadrature de Gauss-Hermite :

$$\int_{-\infty}^{\infty} f(u) \exp(-u^2) du \approx \sum_{k=1}^q c_k f(s_k) \quad (1.14.)$$

où les c_k sont des poids et les s_k sont des points de quadrature. L'approximation est d'autant meilleure que le nombre de points de quadrature q est grand. Cette approximation de la vraisemblance est ensuite maximisée par des algorithmes classiques comme l'algorithme de Newton-Raphson.

2. Quasi-vraisemblance pénalisée ou "Penalized Quasi-Likelihood" (PQL)

Cette méthode consiste à approximer le modèle linéaire généralisé mixte par un modèle linéaire mixte ("linearization method"). La vraisemblance marginale 1.10 résulte de l'intégration de la distribution jointe de y et u par rapport aux effets aléatoires u . En utilisant la représentation sous forme de famille exponentielle de chaque composant de cette loi jointe, l'intégrale 1.12 se ramène à une fonction exponentielle de u . Cette fonction peut alors être approximée par une série de Taylor du 2nd ordre ("Laplace approximation").

Cette approximation, appelé quasi-vraisemblance pénalisée, est maximisée en utilisant les méthodes de maximisation des modèles linéaires mixtes (REML : Restricted Maximum Likelihood).

3. Autres méthodes

Enfin, on peut citer aussi les méthodes de Monte Carlo ainsi que l'approche Bayésienne.

Pour un exposé plus détaillé des caractéristiques de ces modèles, on pourra se référer au livre de Stroup, (2012). J'utilise préférentiellement ce modèle dans mes travaux récents. En effet, son application est rendue aisée par l'utilisation du package `lme4` du logiciel R.

6.2.6.4. Conclusion sur les méthodes d'analyse

La méthodologie développée dans les sections 6.2.1 jusqu'à 6.2.6 est d'ordre général, et a donc pour ambition de pouvoir traiter la majorité des questions posées. Néanmoins, sa mise en œuvre concrète nécessite quelques raffinements selon le cadre d'application. Des questions techniques et théoriques se posent en effet dès le calcul des variables dérivées, et se poursuivent jusqu'à l'étape finale d'extrapolation des résultats. On pourra se poser par exemple, la question des émissions de particules fines pour lesquelles aucun capteur n'est disponible, et donc se poser la question de son estimation par un modèle adapté et validé. Les conditions de validité d'un tel modèle détermineront l'interprétation et la généralisation des résultats.

En tant que mathématicien, l'application de la méthodologie précédente au contexte de la sécurité et des émissions de polluants a nécessité l'acquisition de compétences spécifiques (compétences métier) éloignées de mon champ d'expertise. Les sections suivantes résument les grandes pistes scientifiques nécessaires à l'application de la méthodologie précédente dans le contexte de l'évaluation de la sécurité et de l'impact environnemental des systèmes embarqués d'assistance à la conduite.

Il convient de noter que bien que non abordé dans ce manuscrit, le domaine du géo-positionnement est transversal à l'ensemble de ces questions et constitue un préalable à l'analyse. Ainsi, la question du mapmatching est essentielle, et j'ai parfois eu à mettre au point de nouvelles approches pour résoudre un problème spécifique (cf. Romon et al. 2015 dans le cadre du projet LAVIA).

6.2.7. Evaluer l'impact sur la sécurité

Le comportement de conduite des français en termes de sécurité a été jusqu'à maintenant évalué de manière globale, grâce à la production de statistiques d'accidents ou de mesures globales des vitesses pratiquées. Le dénombrement au niveau national de l'ensemble des accidents corporels est effectué par le ministère de l'intérieur et le ministère qui a en charge les transports.

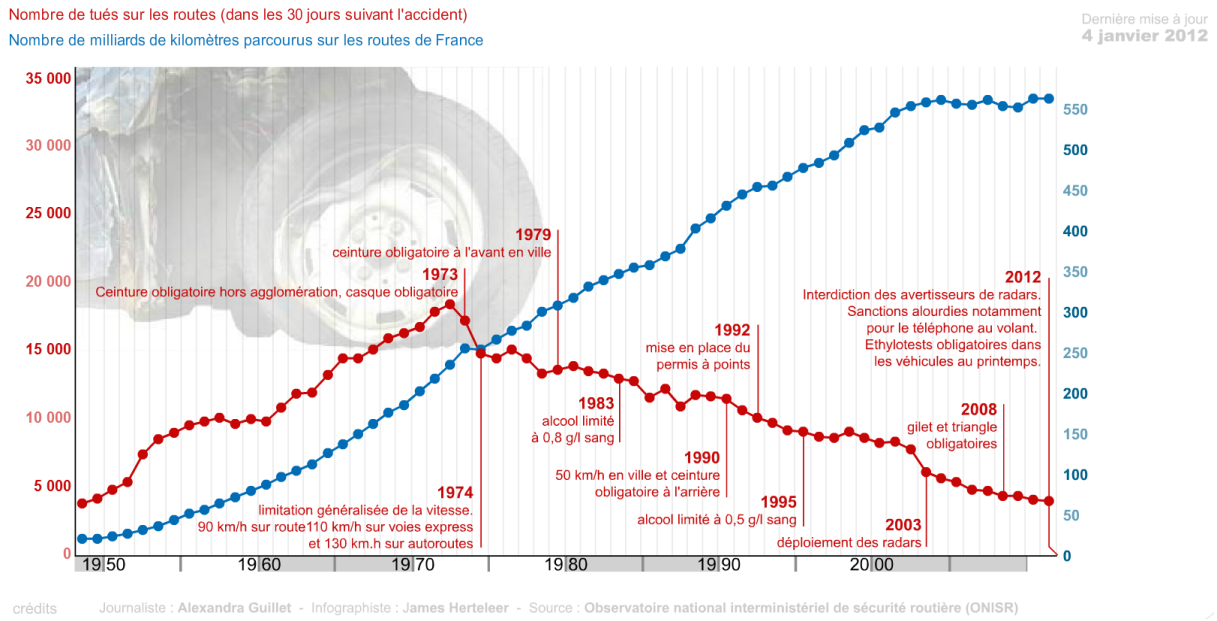


Figure 23 : Nombre de tués sur les routes nationales françaises (Source ONISR).

Des exploitations différentes de ces données sont effectuées par les différentes administrations concernées (ou par des organismes de recherche) mais la source est unique, c'est le bulletin d'analyse d'accident corporel de la circulation (BAAC). Seuls les accidents corporels sont renseignés, et une grande part de l'accidentologie reste ainsi hors de portée des statistiques. En outre, le recueil de ces informations reste complexe et le remplissage des fiches par les forces de l'ordre conduit à un taux d'information erroné proche de 30%, conduisant ainsi des difficultés d'exploitation statistique.

On peut constater à la Figure 23 la décroissance continue du nombre d'accidents corporels depuis les années 70. Etant donné la grande variabilité des situations d'accident, se pose un problème de représentativité statistique lorsque l'on cherche à comprendre finement les causes d'un accident. L'équipe MA (Mécanismes d'accidents) de l'Ifsttar procède à des études détaillées des accidents (EDA) depuis près de 25 ans. Des milliers d'accidents sont ainsi analysés par les experts qui cherchent à reconstruire pour chacun d'entre eux le scénario le plus précis possible. Cela permet le développement de recherches au plus près des problèmes réels de sécurité, que ces problèmes soient interrogés d'un point de vue disciplinaire ou interdisciplinaire. Cependant, la complexité de l'analyse détaillée d'un accident rend cette approche couteuse en temps et produit

des données non représentatives des accidents réels. En outre, lorsqu'il s'agit d'étudier l'impact d'une assistance à la conduite moderne, il devient difficile de trouver des EDA correspondant à ce type de cas. A titre d'exemple, j'ai contacté l'équipe MA en 2008 pour étudier les EDA impliquant l'usage d'un régulateur de vitesse actif au moment de l'accident, mais un seul cas de ce type avait été étudié à l'époque.

L'étude classique des accidents ou des infractions (excès de vitesse, port de la ceinture, alcool au volant, distances de sécurité) ne permet donc pas de répondre de manière satisfaisante aux enjeux modernes de la sécurité routière qui concernent l'impact des systèmes embarqués d'assistance à la conduite, ou plus récemment l'impact des véhicules connectés et/ou autonomes.

Dans les différents projets de recherche auxquels j'ai participé, un objectif majeur consistait à délivrer aux décideurs, une vision claire et argumentée de l'impact de systèmes d'assistance à la conduite embarqués sur la sécurité. Ce travail nécessite d'une part l'emploi d'une approche classique de cette évaluation, et d'autre part la recherche d'une méthodologie plus adaptée aux problématiques actuelles. Ci-après sont donnés quelques éléments méthodologiques concernant les approches classiques employées (vitesses et accidents), en entrant plus dans le détail pour l'approche novatrice à laquelle j'ai consacré une partie de mon travail de chercheur (incidentologie).

6.2.7.1. Evaluation du risque par occurrence d'accidents

Les accidents de la route sont un phénomène « aléatoire » rare, dont l'occurrence peut être décrite statistiquement par une loi de Poisson. Si λ est le nombre moyen d'accidents observés sur un réseau sur une durée donnée, la probabilité de m accident est :

$$P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!} \quad (1.15.)$$

Le risque routier est défini à partir des événements d'accidents de la route menant à un dommage corporel (blessure ou décès). Il se caractérise par les éléments suivants :

- Alea : probabilité d'occurrence de l'événement redouté
- Gravité : probabilité d'être tué/blessé en cas d'accident
- Exposition au risque : s'exprime sous forme de risque relatif ou « odd ratio »

On appelle insécurité le terme suivant :

$$\text{Insécurité} = \text{Exposition} \times \text{Alea} \times \text{Gravité} \quad (1.16.)$$

$$\text{Avec Risque} = \text{Alea} \times \text{Gravité} \quad (1.17.)$$

La gravité d'un accident dépend de la violence du choc, et peut être schématisée par la Figure 24.

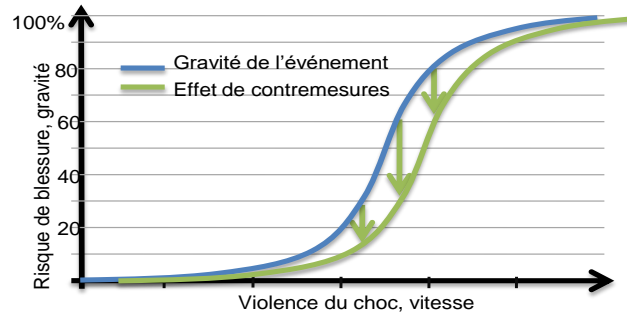


Figure 24 : évolution schématique de la gravité d'un accident selon la violence du choc.

L'évaluation du risque peut donc se faire en estimant chacune de ces composantes. Dans le contexte des véhicules autonomes, il est par exemple possible d'estimer l'aléa de manière dynamique en calculant à chaque instant une probabilité de collision, permettant de déclencher des contre-mesures d'urgence (cf. Lambert, Gruyer, Saint Pierre, 2008).

En outre, les véhicules modernes et leurs équipements embarqués peuvent agir sur le risque en réduisant soit la probabilité d'accident (sécurité active, ABS ou ESP par exemple), soit la gravité de celui-ci (sécurité passive, ceinture de sécurité par exemple).

L'occurrence d'un accident de la route est toujours un drame, mais l'on sait que cela dépend d'un grand nombre de facteurs. En général, c'est la conjonction de plusieurs d'entre eux qui déclenche l'accident. C'est-à-dire qu'un grand nombre d'accidents sont évités chaque jour pour diverses raisons. Ce phénomène peut être décrit simplement à l'aide de la pyramide de Heinrich présentée à la Figure 25 (Heinrich 1939).

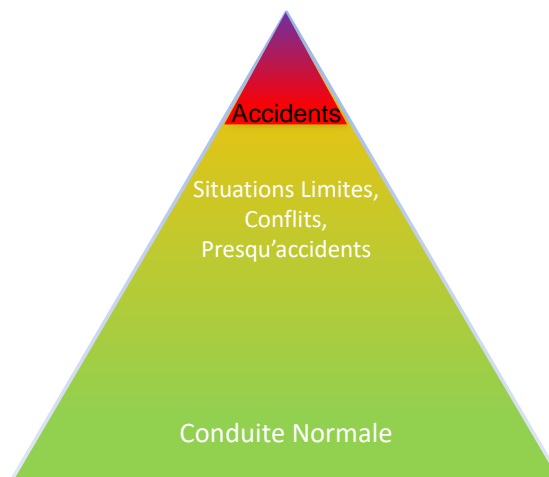


Figure 25 : Pyramide de Heinrich pour la sécurité routière.

Une situation de conduite limite (conditions dégradées liées à la météo par exemple) combinée à un événement à risque (excès de vitesse par exemple) peut être considérée comme les prémices

d'un accident de la route. Il en est de même pour un évènement de type freinage brusque, qui peut dégénérer en accident sur une chaussée humide.

Les accidents évités sont donc des évènements préfigurateurs d'un accident, et à ce titre peuvent constituer un indicateur de risque pertinent. En particulier, leur nombre est bien plus élevé que les seuls accidents corporels, et peuvent permettre des études statistiques plus pertinentes. A titre d'exemple, l'étude américaine 100-car des années 2000 qui fait référence a permis de collecter plus de 3.2 millions de km pendant plus 43000h sur 100 véhicules instrumentés, et n'a enregistré « que » 68 accidents, pour plus de 750 incidents détectés et confirmés manuellement (cf. Guo, Klauer, Hankey, and Dingus (2010). Aucun accident n'a été observé à ce jour dans les projets Européens équivalents. L'accidentologie n'est donc d'aucun recours dans un contexte expérimental, aussi ambitieux soit-il.

Nous reviendrons plus loin sur la question de la détection des incidents. Intéressons-nous maintenant aux méthodes permettant d'extrapoler le risque routier global à partir des vitesses pratiquées moyennes.

6.2.7.2. Extrapolation du risque à partir de la vitesse moyenne

Les vitesses excessives, qui comprennent les excès de vitesse par rapport à la limitation et les vitesses inappropriées compte tenu des circonstances, même au-dessous de la limitation, sont dangereuses. La vitesse est un facteur accidentogène dans environ un tiers des accidents mortels et c'est aussi un facteur déterminant sur la gravité de tous les accidents. Elle a également des effets néfastes sur l'environnement et la consommation énergétique.

L'ONISR (Observatoire national interministériel de la sécurité routière) ne se base pas uniquement sur les accidents ou les infractions, mais produit aussi une analyse descriptive du comportement de conduite des français par le biais des mesures de vitesse qui font l'objet de campagnes de mesures régulières. Depuis 2013, l'observatoire des vitesses de l'ONISR est alimenté par des mesures réalisées par le Cerema (Centre d'Etudes et d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement) sur un échantillon de 58 points représentatif du réseau routier. La méthodologie de recueil est régulièrement mise à jour, ce qui entraîne des difficultés dans le suivi des indicateurs. Le faible nombre de points représentatifs du réseau utilisés, rend complexe toute déduction sur le risque routier au niveau local ou même régional.

Cependant, des modèles globaux, fondés sur des mesures moyennes effectuées pendant une période de temps longue (quelques jours ou mois), permettent de relier la fréquence des accidents à la vitesse moyenne.

Une proportion considérable de véhicules roule à une vitesse excessive par rapport à la limitation. On pourrait en conclure que les limitations de vitesse influent peu sur les vitesses de circulation. Cependant, un très grand nombre d'études ont été menées pour évaluer l'effet des modifications

de limitation de vitesse et ont démontré que les limitations de vitesse ont bien une influence évidente sur les vitesses réelles (Finch et al. (1997)). Par exemple, les méta-analyses signalent qu'un abaissement de la limitation de 10 km/h entraîne une réduction de la vitesse de 3 à 4 km/h.

On observe également une relation évidente entre la vitesse et le taux d'accident. Cette relation peut cependant s'expliquer/se modéliser à l'aide des bases des données d'accidents et des liens avec la vitesse observée sur le réseau routier correspondant. Le modèle Puissance estime par exemple que si la vitesse moyenne sur une route donnée est réduite de 5%, le nombre de blessés baissera de 10%, et le nombre de tués, de 20%.

Il convient de remarquer qu'un changement de la limitation de vitesse, sans mesure d'accompagnement, a peu d'effet. Toute modification de la limitation de vitesse doit être accompagnée de mesures de contrôle-sanction, d'aménagement de l'infrastructure et d'une communication appropriée. L'abaissement de la limitation de vitesse, sans modification de la conception d'une route ou sans mesures permanentes de contrôle-sanction, ne permet pas un ralentissement du trafic du niveau de celui prescrit par la nouvelle limitation.

Il existe différents modèles statistiques permettant de relier les vitesses moyennes pratiquées avec une fréquence d'accident. Ceux-ci dépendent du type de route étudié, mais sont aussi liés aux données utilisées pour les calibrer, et donc au pays les ayant produits. Pour une présentation exhaustive de l'ensemble de ces modèles, nous renvoyons le lecteur à l'étude de référence produite par Taylor et al. (2000).

Les principaux modèles disponibles sont rappelés ci-dessous :

Le modèle urbain U1

Calibré sur les routes urbaines du Royaume-Uni (50-70 km/h), le modèle *U1* utilise comme mesures la vitesse moyenne et le coefficient de variation de la vitesse, et propose une relation exponentielle pour en déduire la fréquence des accidents corporels.

Le modèle urbain U2

Le modèle urbain *U2* repose sur la proportion de véhicules qui dépassent la vitesse limite et sur l'ampleur de l'excès de vitesse, qui représente la différence entre la limitation de vitesse et la vitesse moyenne de ceux qui la dépassent. Ce modèle utilise lui aussi une relation exponentielle et est calibré sur les mêmes types de routes.

Le modèle urbain de Kloeden

Kloeden et al. (2002) ont utilisé la régression logistique modifiée, et ont obtenu une fonction exponentielle polynomiale de second ordre décrivant la relation entre le risque relatif d'être impliqué dans un accident et la vitesse libre de circulation.

Le modèle rural de Kloeden

Ce modèle repose sur une approche de type cas-témoin appliquée aux chaussées rurales à voie unique de l'Australie de Sud, avec des limites de vitesses de 80 à 130km/h (Kloeden, Ponte et McLean, 2001). Une relation exponentielle associe le risque relatif d'accident à la vitesse moyenne des véhicules. Une hypothèse implicite des modèles de Kloeden est le fait que le seul choix individuel de la vitesse détermine le risque d'accident.

Le modèle puissance de Nielsson

Le Modèle Puissance a été proposé en 1982 par le chercheur suédois Göran Nilsson. Selon ce modèle, valable sur les routes entre 90 et 110 km/h, les effets des changements de vitesse sur le nombre d'accidents et la sévérité des blessures peuvent être estimés par des fonctions puissance. Le Modèle Puissance décrit la relation entre la variation de vitesse (avant et après une contre mesure) et le nombre d'accidents (ou le nombre de blessés) par six équations, selon le niveau de gravité des blessures (blessés, blessés graves, décès etc.).

La méta-analyse de Elvik et al. (2004)

En 2004, Elvik et al. ont proposé une méta-analyse qui intègre à la fois la vitesse et la sévérité d'un accident. L'objectif du rapport de Elvik et al. a été d'évaluer la validité du modèle puissance à travers une révision systématique et une méta-analyse des études appropriées. La méta-analyse a utilisé à la fois des techniques traditionnelles et des techniques de la méta-régression (modèles multivariés). Les auteurs ont suggéré que la réduction relative des accidents (ou des victimes) était liée à la réduction de la vitesse moyenne, élevée à une puissance, comme l'avait suggéré auparavant Nilsson.

Les résultats de Elvik et al. produisent un modèle général, c'est-à-dire un modèle non spécifique à aucune catégorie de routes en particulier, reliant la vitesse au risque d'accident. Ce type de modèle est très utilisé pour analyses de type coût/bénéfice, et appliqué chaque fois qu'un modèle plus approprié n'est pas été trouvé.

Application à l'étude des systèmes embarqués d'assistance à la conduite

Ces modèles très généraux sont applicables à l'étude des systèmes agissant sur la vitesse comme le régulateur ou le limiteur de vitesse. Toutefois, les expérimentations utilisant des véhicules instrumentés peinent à atteindre une représentativité statistique vu le faible effectif des conducteurs participants (quelques centaines tout au plus). L'emploi de ces approches pour évaluer l'impact des systèmes embarqués sur le nombre d'accidents et le nombre de morts, n'est donc pas entièrement satisfaisante et peut mener à des conclusions erronées (un système qui diminue la vitesse moyenne aura forcément un impact positif).

En conséquence, il est utile que la recherche explore d'autres voies pour répondre correctement à ces questions. Lors de mes propres recherches j'ai eu l'occasion de développer différentes approches permettant une évaluation du risque à partir de données collectées sur une flotte de véhicules.

6.2.7.3. Une nouvelle approche, l'incidentologie

Comme on vient de le voir, l'évaluation du risque routier procède par des analyses globales (à l'échelle nationale) afin de préserver une représentativité statistique, et afin de faciliter les choix pour les décideurs des politiques publiques. Avec la modernisation des véhicules et l'arrivée prochaine de véhicules autonomes et/ou connectés, les enjeux de sécurité routière concernent de plus en plus l'impact des systèmes embarqués d'assistance à la conduite, et l'impact de la conduite semi-autonome. Les méthodes usuelles ne permettent pas d'étudier ces aspects, et il est du ressort de la recherche d'étudier des approches nouvelles pour anticiper au mieux l'évolution de la mobilité. Fort heureusement, les évolutions technologiques récentes ont amené une nouvelle génération de véhicules personnels ouvrant de nouvelles voies d'études de la sécurité routière.

Cependant, comme vu précédemment, les accidents sont très rares, et leur étude détaillée à l'aide de véhicules instrumentés reste donc hors de portée. La recherche s'intéresse donc à l'étude des incidents (ou presque-accidents) pour mieux comprendre les causes des accidents.

La grande difficulté à laquelle est confronté le chercheur, consiste à produire des méthodes pertinentes de détection de ces incidents. Une méthode fiable de détection des incidents doit permettre de cibler au mieux les incidents potentiellement convergents vers une situation d'accident. La détection du caractère potentiellement convergent vers le risque de ces situations est cependant complexe, car il est très difficile d'appréhender le contexte de conduite avec les seules données issues du véhicule. En effet, des actions de contre-mesures peuvent être entreprises pour différentes raisons telles que prendre un risque ou à l'inverse éviter un risque. Ce sont donc des stratégies non dénuées d'ambiguïté dont on ignore l'importance. En l'état actuel des connaissances, il n'existe pas de méthode fiable de détection des événements à risque de conduite. Les résultats les plus récents nous apprennent que les méthodes automatiques les plus simples identifient correctement moins de 10% des incidents. Il existe différentes approches possibles permettant de détecter les incidents de la conduite :

6.2.7.3.1. L'identification fondée sur les réflexes de l'automobiliste

L'idée principale de cette approche s'appuie sur le fait que les automobilistes recherchent le confort de conduite, et ne seront généralement pas exposés à des manœuvres cinétiques drastiques sauf exception. Le moyen principal pour identifier ces situations dangereuses consiste à détecter les fortes accélérations/décélérations ou un changement de direction rapide. Par exemple, nous pouvons rechercher les moments où la pression sur le frein est forte, ce qui peut être interprété

comme une réaction face à un évènement. C'est ce type de déclenchement qui est utilisé pour activer les fonctions d'assistance au freinage d'urgence.

6.2.7.3.2. L'identification fondée sur les systèmes d'aide à la conduite

Une deuxième approche très naturelle utilisée pour la détection des presque-accidents consiste à utiliser l'information issue des systèmes d'aide à la conduite. Les avertissements émis lors du déclenchement d'un système embarqué (ESP, LDW, etc.) peuvent être utilisés comme identificateurs d'événements. L'inconvénient étant bien sûr que les incidents qui se déroulent en dehors de la capacité de détection du système, ou qui se produisent lorsque le système est éteint, ne sont pas pris en compte pour l'analyse des risques. Il devient alors difficile d'estimer la fréquence des conflits.

6.2.7.3.3. L'identification fondée sur les situations de conduite

Une troisième approche se focalise sur les conditions de circulation. L'idée est que pour certains véhicules et certaines conditions de trafic, les marges de sécurité sont si petites qu'il en faut très peu avant qu'une situation d'incident ne se présente. En d'autres termes, de trop petites marges risquent d'augmenter le risque de collision. Par exemple, ce risque peut être estimé par la fréquence des événements correspondants à des freinages brusques, ou par la fréquence des excès de vitesse.

6.2.7.3.4. L'identification fondée sur le passé de l'automobiliste

Une quatrième approche consiste à rechercher des événements inhabituels dans le passé des automobilistes. L'idée est que des événements inhabituels dans l'historique de conduite d'une personne sont rares, précisément parce que l'automobiliste tente d'éviter de tels événements. Ces événements représentent des situations que l'automobiliste n'est pas censé(e) effectuer, et donc au moins une certaine partie d'entre eux peuvent se rapporter à des situations potentiellement dangereuses.

Selon l'équipement des véhicules participant à l'expérimentation, l'ensemble des méthodes d'identification ne sera pas accessible. C'est le cas par exemple des alertes produites par un ESP ou un ABS qui nécessitent d'avoir accès au bus CAN des véhicules. Cet accès étant protégé par les constructeurs, il n'est en général possible d'accéder uniquement à la partie "ouverte" du bus CAN appelée OBD II. De même, le coût de l'équipement peut limiter les possibilités d'enregistrement. Ainsi, il est fréquent de ne pas disposer d'un radar avant ou arrière, qui seuls permettent d'accéder aux temps et aux distances inter-véhiculaires. Le choix des méthodes appliquées dépendra donc des possibilités du recueil de données.

Beaucoup d'efforts ont été déployés pour développer des algorithmes et des techniques qui permettent l'identification de presque-accidents. Malheureusement, à ce jour, aucune liste officielle de définition d'indicateurs d'incidents n'est disponible. La mise en place d'événements à risques et d'événements de référence nécessite une réflexion sur les mécanismes de causalité d'accidents. Le Tableau 5 permet de comparer les différentes définitions de certains projets de référence, donc les projets EuroFOT et MEDOC auxquels j'ai activement participé.

Tableau 5 : Seuils utilisés pour l'identification des incidents de la route dans quelques projets de référence.

Type de détection	100-Car study (USA)	EuroFOT (EU-FP7)	MEDOC (Fondation MAIF)	S-Vrai (DSCR)
Référence	Dingus et al. (2006) Appendix B.	Mohamed Benmimoun et al, 2011	Bagdadi, O., Várhelyi, A. (2011). Bagdadi, O. (2013).	Bagdadi, O., Várhelyi, A. (2011). Bagdadi, O. (2013).
Accélération latérale	Accélération latérale $\geq 0.7g$	$ a_{latérale} \geq a_{latérale}$ avec : <ul style="list-style-type: none"> $a_{latérale} = (7 - 2.5)m/s^2 * (v / 40 \text{ km/h}) + 2,5 \text{ m/s}^2$ [Si vitesse < 40 km/h] $a_{latérale} = 7 \text{ m/s}^2$ [Si vitesse entre 40 km/h et 50 km/h] $a_{latérale} = (4 - 7)m/s^2 * ((v - 50 \text{ km/h}) / 50 \text{ km/h}) + 7 \text{ m/s}^2$ [Si vitesse entre 60 km/h et 100 km/h] $a_{latérale} = 4 \text{ m/s}^2$ [Si vitesse > 100 km/h] 	$Accélération_{latérale} \leq -0.75$ ou $Accélération_{latérale} \geq 0.75$ (Si plusieurs détections dans un intervalle de 2 secondes, alors on n'en conserve qu'une seule)	
Accélération / décélération longitudinale	- Accélération ou décélération $\geq 0.6g$	- Accélération $\leq a_{longlow}$ avec : <ul style="list-style-type: none"> $a_{long} = -6 \text{ m/s}^2$ [Si vitesse < 50 km/h] $a_{long} = (-4 - -6) \text{ m/s}^2 * ((v - 50 \text{ km/h}) / 100 \text{ km/h}) + -6 \text{ m/s}^2$ [Si vitesse entre 50 km/h et 150 km/h] $a_{long} = -4 \text{ m/s}^2$ [Si vitesse > 150 km/h] 	- Accélération $\leq -0.65g$ - Accélération $\leq -0.3g$ si la vitesse du véhicule >64 km/h pendant au moins 20 millisecondes (Si plusieurs détections dans un intervalle de 2 secondes, alors on n'en conserve qu'une seule)	- $ Accélération > 0.61g$ avec Jerk > 20 m/s^3 - $ Accélération > 0.51g$ avec vitesse > 80 km/h - $ Accélération > 0.41g$ avec Jerk > 20 m/s^3 et vitesse > 100 km/h
Bouton d'incident	Bouton sur le tableau de bord activé par le conducteur lorsqu'il estime l'incident suffisamment pertinent		Bouton sur le tableau de bord activé par le conducteur lorsqu'il estime l'incident suffisamment pertinent	
Jerk critique			Jerk $< -1g/s$ si la vitesse du véhicule $\geq 18km/h$ pendant au moins 1 seconde (Si plusieurs détections dans un intervalle de 2 secondes, alors on n'en conserve qu'une seule)	
Yaw rate (cap ou vitesse angulaire de lacet)	Variation du cap du véhicule entre -4° et $+4^\circ$ pendant une période $\leq 3s$	$ Yaw \text{ rate} \geq \Psi_{low}$ avec : <ul style="list-style-type: none"> $\Psi_{low} = 50 \text{ }^\circ/s$ [Si vitesse < 40 km/h] $\Psi_{low} = (25 - 50) \text{ }^\circ/s * ((v - 40 \text{ km/h}) / 10 \text{ km/h}) + 50 \text{ }^\circ/s$ [Si vitesse entre 40 km/h et 50 km/h] 	Variation de yaw-rate de -8° à $+8^\circ$ ou de $+8^\circ$ à -8° en moins de 0,75 seconde et avec une vitesse du véhicule >48 km/h. (Si plusieurs détections dans un intervalle de 2 secondes, alors on n'en conserve qu'une seule)	

		<ul style="list-style-type: none"> • $\Psi_{low} = (15 - 25) \text{ }^\circ/\text{s}^*$ $((v - 50 \text{ km/h}) / 35 \text{ km/h}) + 25 \text{ }^\circ/\text{s}$ [Si vitesse entre 50 km/h et 85 km/h] • $\Psi_{low w} = 15 \text{ }^\circ/\text{s}$ [Si vitesse > 85 km/h] 		
Temps à collision (TTC)¹ avant critique	<ul style="list-style-type: none"> - Accélération ou décélération $\geq 0.5g$ avec TTC avant $\leq 4s$ - $-0.5g \leq \text{décélération} \leq -0.4g$, avec TTC avant $\leq 4s$ et distance au véhicule avant $\leq 30m$ lorsque le TTC est minimum 	<ul style="list-style-type: none"> - Temps inter-véhiculaire (TIV²) < 0.35s avec une approche lente du véhicule précédant (10 km/h < v < 20 km/h) - TIV < 0.5 s avec une approche rapide au véhicule précédant (v > 20 km/h) - TTC < 1.75 s - TTC < 1 s avec appui sur le frein - TIV < 0.35 s avec une approche rapide du véhicule précédant (v > 20 km/h) - TTC < 1 s sans freiner 		
Temps à collision (TTC) arrière critique	<ul style="list-style-type: none"> - TTC arrière $\leq 2s$ associée à une distance inter-véhiculaire arrière $\leq 15m$ - Accélération relative du véhicule suiveur $\geq 0.3g$ 			

6.2.7.3.5. Suppression des faux positifs et problème de critère de référence

Il n'y a pas d'identification de presque-accidents parfaite, toute recherche d'événements dans une base de données NDS ou FOT est susceptible de contenir un mélange de vrais positifs (événements réellement pertinentes) et de faux positifs, c'est-à-dire les événements détectés par divers algorithmes se révèlent ne pas être couplé au mécanisme d'accidentologie.

¹ On rappelle que le TTC est le temps avant collision qui est calculé à partir de la vitesse relative entre les deux véhicules.

² On rappelle que le TIV est le temps inter véhiculaire qui est calculé à partir de la vitesse absolue du véhicule.

Par conséquent, après avoir détectés tous les événements de type incident, il vient l'étape très importante de l'élimination des faux positifs dans la liste des vrais positifs. Il existe essentiellement trois façons de le faire :

- une inspection visuelle/manuelle : les chercheurs examinent chaque événement à partir de la vidéo enregistrée et décident ensuite si l'événement est vraiment pertinent ou non. Il s'ensuit également que l'évaluation varie selon les chercheurs;
- appliquer un filtre conditionnel : il s'agit de détecter des événements et d'ajouter des critères supplémentaires. Par exemple, si un conducteur a quitté la voie de circulation, on pourrait supprimer tous les cas où celui-ci n'a pas réagi dans les 3 secondes suivantes, sur l'hypothèse que si le conducteur ne réagit pas à la situation, c'est qu'elle n'est pas critique;
- une combinaison d'événements (Benmimoum et al. (2011)) : les combinaisons diverses de vitesse, de changement de lacet, de freinage, d'accélération ... permettraient d'identifier des incidents plus précis.

Par ailleurs, lors de l'analyse d'incidents, que ce soit leur fréquence relative ou un aspect particulier du comportement du conducteur durant une situation critique, il y a toujours un besoin de définir une référence de base, à savoir, une situation de comparaison. C'est l'un des aspects les plus difficiles à appliquer dans les études NDS et FOT. Dans les études NDS, il y a deux façons de sélectionner les événements de référence. La première consiste à prélever au hasard un certain nombre d'événements de référence de toutes les données qui ne sont pas critiques. L'autre consiste à adapter la sélection de la base de référence pour les participants (par exemple, si un incident se produit en ville lorsqu'il pleut, nous essayons de chercher un événement de référence qui se produit également en ville et lorsqu'il pleut). Dans les études FOT, le problème de sélection de référence est légèrement différent. L'exemple de l'Adaptive Cruise Control (ACC) illustre bien ce point. Il suffit de regarder les événements lorsque l'ACC est actif et lorsqu'il ne l'est pas.

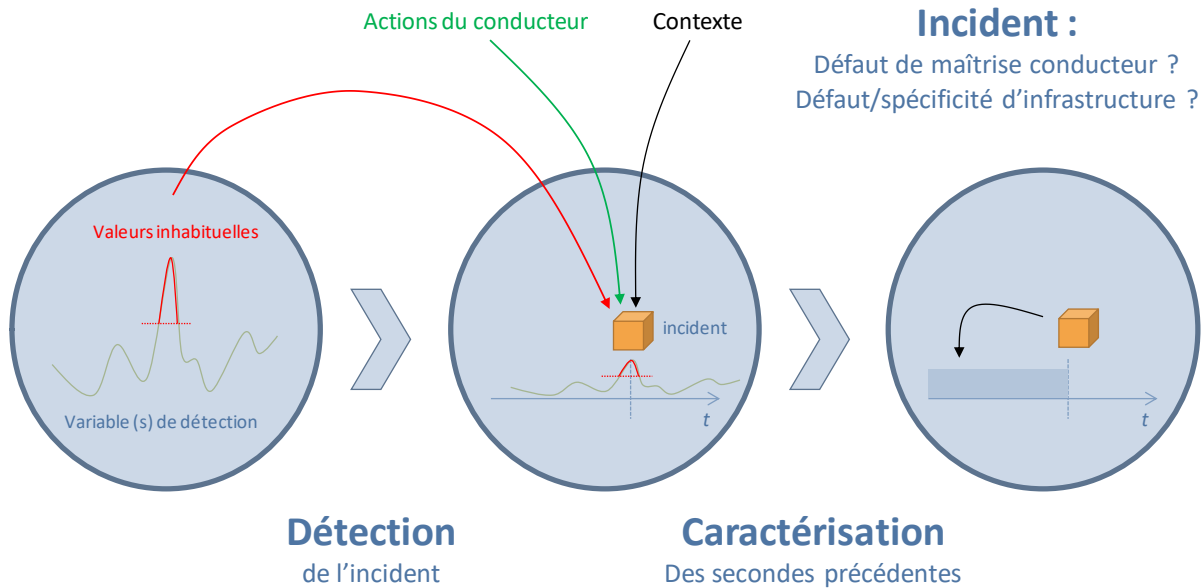


Figure 26: Schéma d'analyse du lien entre incidents et conditions de circulation

6.2.7.3.6. Lien entre incidents détectés et risque routier réel

La question de la bonne identification des incidents de la route est un vrai challenge, mais l'interprétation des résultats et la mise en évidence d'un lien avec le risque routier réel, c.a.d. celui estimé par l'occurrence des accidents de la route, l'est encore plus.

En effet, une fois l'incident détecté et caractérisé (10 ou 20 secondes de conduite précédentes annotées manuellement grâce à la vidéo dans le meilleur des cas), il devient essentiel d'en comprendre les déterminants et les facteurs critiques.

Il est néanmoins probable que l'occurrence de ceux-ci puisse caractériser le niveau de risque d'un conducteur. Bien que non encore établie scientifiquement, cette hypothèse sert de fondement à l'analyse du risque dans bon nombre d'expérimentations à grande échelle (projet euroFOT par exemple). De manière identique, il est possible d'élargir la notion d'incident à d'autres événements moins critiques, comme les excès de vitesse par exemple. La fréquence des excès de vitesse peut servir à caractériser l'usage d'un système embarqué (cas de l'évaluation du limiteur et régulateur dans le projet euroFOT, cf. section 6.3.2), mais peut aussi être corrélée au niveau de risque du conducteur.

La question de l'analyse et de l'interprétation des incidents fait l'objet de deux autres projets de recherche dans lequel je suis ou ait été impliqué (voir section 6.6.3.3 pour le projet MEDOC, le projet UDrive n'étant pas abordé en détail dans ce document).

- Dans le projet UDrive, nous espérons collecter une centaine d'incidents vérifiés et annotés à l'échelle européenne. Mon travail consistera, une fois ces informations disponibles, à analyser statistiquement les facteurs de risque déclenchant l'incident (NB : j'ai dû mettre

un terme en mars 2017 à ma participation à ce projet, le Cerema ne faisant pas partie des partenaires).

- Dans le projet MEDOC, que je coordonne avec Patricia Delhomme, nous cherchons à confirmer les incidents détectés automatiquement avec le concours des conducteurs appuyant sur le bouton d'incident. Cette « vérité » perçue par les conducteurs me permettra d'affiner les algorithmes de détection d'incident et ainsi espérer prouver un lien entre leur fréquence et l'accidentologie réelle des conducteurs.

Ces deux derniers aspects font partie de mes thèmes de recherche actuels et à venir.

6.2.8. Évaluer l'impact sur l'environnement

L'évaluation de l'impact environnemental des systèmes embarqués d'assistance à la conduite représente un enjeu considérable, tant économique que sociétal. L'actualité récente a prouvé que les industries impliquées surévaluent les économies réalisées et que la société civile ne peut plus accepter de tels arrangements.

Cependant, l'évaluation précise de l'impact d'un système embarqué sur les émissions de GES (Gaz à effet de serre) reste très compliquée à réaliser, autant pour des raisons techniques que méthodologiques.

Je me suis intéressé à cette question dès mon arrivée au LIVIC, où ma première mission a consisté à réaliser l'étude de l'impact sur la consommation d'un système de limitation adaptative de la vitesse (le LAVIA). Depuis, je n'ai eu de cesse de poursuivre dans cette voie, jusqu'à étudier plus précisément les techniques de conduite spécifiques destinées à réduire les émissions (éco-conduite).

En apparence, le problème est simple puisqu'il « suffit » d'enregistrer la consommation instantanée du/des véhicules qui est mesurée par un capteur embarqué sur la totalité des véhicules modernes. Il suffit ensuite de calculer des consommations moyennes selon les conditions de conduite en suivant la méthodologie expliquée à la section 6.2.3. Hélas, cette information est en général inaccessible pour des raisons de confidentialité, et quand bien même le serait-elle que la question des émissions de particules fines ne serait pas résolue. En effet, les émissions de GES ne sont en général pas proportionnelles à la consommation de carburant et dépendent d'un grand nombre de facteurs difficiles à estimer.

Au fil de mes recherches, j'ai abordé ces questions à plusieurs reprises, et me suis confronté à différents cas de figure. Cette section est dédiée à une présentation résumée des méthodes disponibles pour résoudre ces questions.

6.2.8.1. Enregistrement de la consommation instantanée

La consommation instantanée est toujours existante sur le bus CAN des véhicules, mais n'est quasiment jamais disponible sur sa partie « ouverte » qu'est l'OBD-II. Lorsque l'industriel fabricant du véhicule est partenaire du projet de recherche, il est parfois possible d'accéder à cette information directement.

L'emploi d'un protocole de type avant-après, permet dans ce cas de donner des mesures relatives de réduction de la consommation, sans pour autant dévoiler les mesures absolues. Cette approche a pu être employée notamment dans le projet euroFOT et ecoDriver, et dans mes propres études sur l'éco-conduite.

Il faut noter que la consommation de carburant est proportionnelle aux émissions de CO₂, mais qu'un tel lien n'existe pas pour les émissions de particules, les NO_x ou le CO dont la quantité

produite est bien plus difficile à estimer. Il faut dans ce cas utiliser des modèles d'estimation encore peu précis ou disposer d'un matériel d'enregistrement très coûteux.

6.2.8.2. Reconstitution de la consommation instantanée

Lorsque j'ai travaillé sur le projet LAVIA, les données étaient déjà recueillies, l'expérimentation terminée, et la consommation de carburant n'avait pas été enregistrée.

Il a donc fallu estimer cette dernière à partir d'autres paramètres de la conduite. Avec l'aide du constructeur Renault, j'ai pu reconstituer de manière non paramétrique la cartographie moteur d'une partie des véhicules ayant participé à l'expérimentation, et ainsi estimer la consommation instantanée (Figure 27).

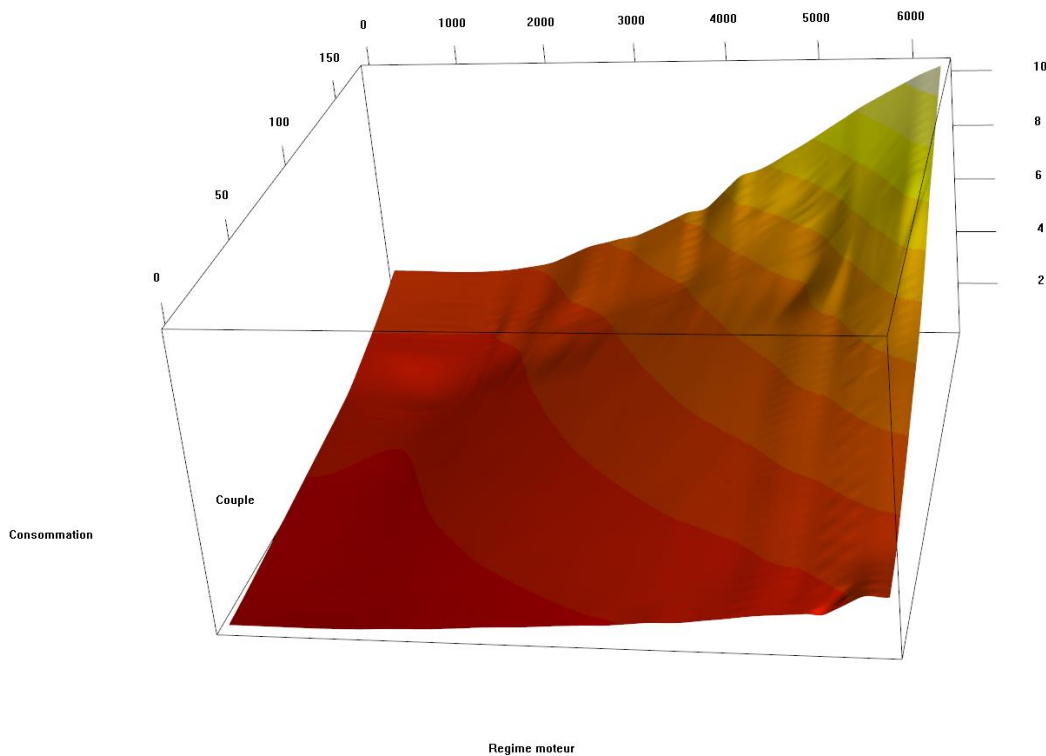


Figure 27 : Interpolation de la cartographie moteur d'une Laguna II permettant de calculer la consommation instantanée (en gr/sec) en fonction du couple et du régime du moteur (en tours/min).

6.2.8.3. Estimation à partir des émissions réelles

L'enregistrement de la position de la pédale d'accélérateur et du régime moteur permet d'identifier le point de fonctionnement du moteur. Dans certains pays, des organismes publics recensent de manière expérimentale les émissions de l'ensemble des véhicules du parc automobile

national selon différents points de fonctionnement (utilisation d'un banc à rouleau pour contrôler les paramètres). Il est alors possible de reconstituer après-coup les émissions point par point. On peut citer par exemple le UK NAEI - National Atmospheric Emissions Inventory (<http://naei.defra.gov.uk/>) qui propose un accès libre aux données relatives aux émissions de GES.

Je n'ai pas pu appliquer cette approche car ces informations font défaut en France et sont difficilement accessibles. En outre, elle repose sur l'utilisation de valeurs moyennes discrétisées, qui ne peuvent rendre compte des phénomènes d'accélération brusque fortement consommateurs d'énergie.

6.2.8.4. Estimation de la consommation à partir des vitesses (utilisation de courbes moyennes vitesse/consommation)

Lorsque les approches précédentes ne sont pas possibles, la mesure de la vitesse moyenne permet de reconstituer la consommation moyenne associée, sans toutefois tenir compte des facteurs extérieurs comme la pente de la route ou le chargement du véhicule.

De telles courbes (cf. Figure 28) sont issues de mesures en situations contrôlées (bancs à rouleaux) et ne représentent que sommairement la réalité des émissions.

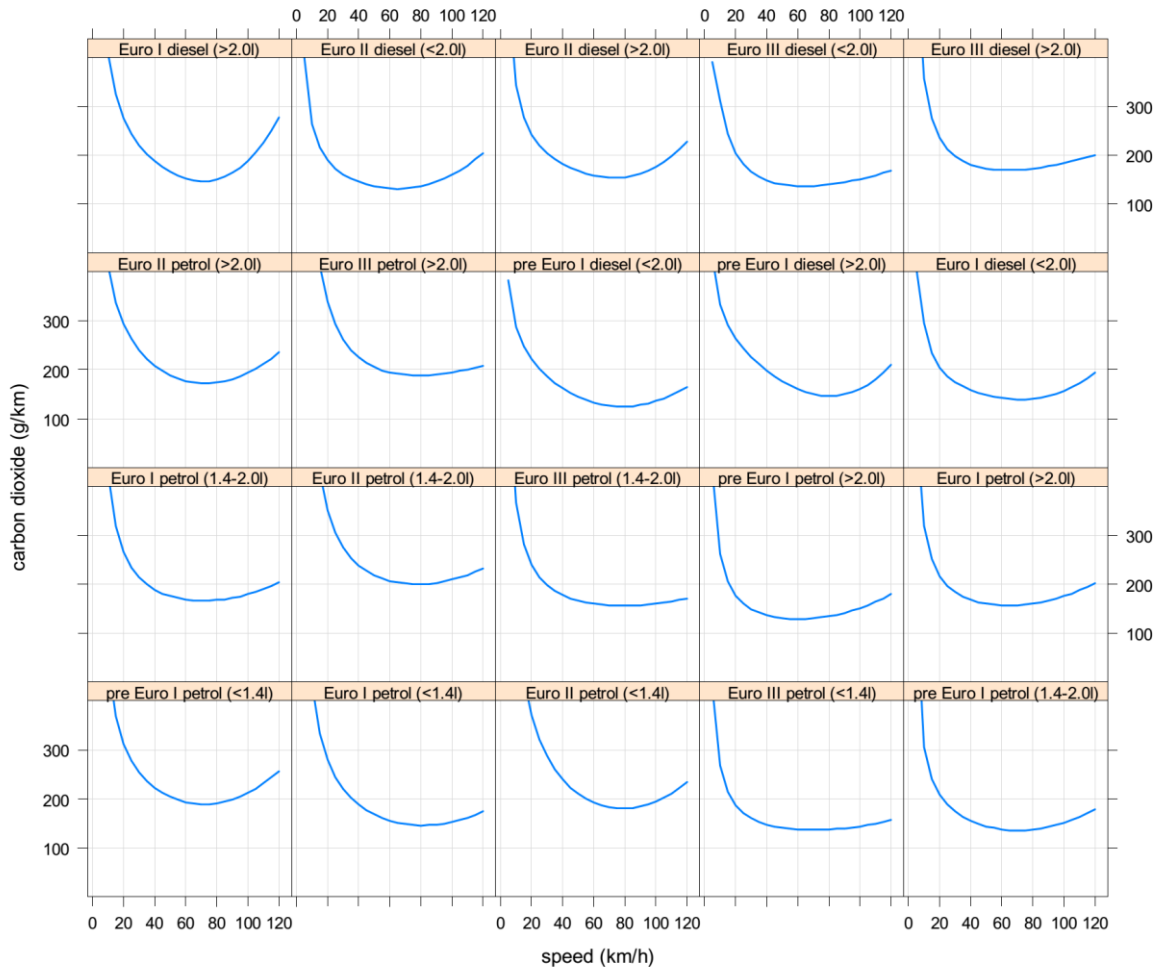


Figure 28 : Exemple de courbes reliant la vitesse moyenne aux émissions de CO₂ pour une variété de véhicules (source NAEI, et Carsten et al. 2008)

6.2.8.5. Simulation

L'emploi de la simulation permet de compléter ou de remplacer les mesures en conditions de route réelles. Afin d'estimer la consommation de carburant (et donc les émissions de CO₂), il est nécessaire de disposer des éléments suivants :

- Une représentation détaillée de l'infrastructure ciblée.
- Une représentation détaillée des mesures de trafic étudiées, ou du système embarqué visé.
- Un modèle précis du comportement du conducteur en réponse aux contraintes de l'infrastructure, du trafic, et du système étudié.
- Un modèle précis du comportement du moteur en réponse aux commandes du conducteur et des caractéristiques de l'infrastructure.
- Une représentation détaillée du trafic envisagé (mix de véhicules, évolution du trafic selon les heures, etc.).

Ces différents modèles sont ensuite intégrés dans un logiciel de simulation qui permet ensuite d'étudier par exemple l'effet de différents taux de pénétration d'un système embarqué.

La validité de telles approches, bien qu'employées systématiquement en complément des expériences réelles, pose un grand nombre de questions. En effet, la précision des différents modèles envisagés laisse souvent à désirer et les informations disponibles sont en général difficiles à obtenir.

J'ai personnellement rarement utilisé cette approche, étant un spécialiste de l'analyse des données réelles. Néanmoins, comme on vient de le voir, il n'existe pas d'approche idéale pour estimer l'impact d'un système sur l'environnement et il est recommandé d'exploiter la totalité des moyens disponibles pour confirmer les résultats. La plupart des projets auxquels j'ai participé associaient des experts de la simulation et des spécialistes de l'analyse des données réelles.

6.2.9. Publications associées à cette section

Ouvrages scientifiques (OS)

- **Saint Pierre G.**, Ledoux V., Serre T., Violette E. (2017), « *La mobilité numérique au service de la sécurité routière* », publié dans "évaluation des politiques de sécurité routière, méthodes, outils, limites", sous la direction de Gilles Blanchard et Laurent Carnis, éditions L'Harmattan.
- **Saint Pierre G.** (2010), « *Le système LAVIA et la consommation de carburant : méthodes exploratoires pour données de conduite en situation naturelle* », collection "études et recherches du laboratoire des ponts et chaussées", CR54, 120 pages.
- Vrignon J., Rakotonirainy A., Gruyer D., **Saint Pierre G.** (2008), « *The impact of subjective factors on driver vigilance: a driving simulator study* », in "Driver Behavior and training" vol III, chapitre 29, Ashgate publishing.

Articles dans des revues à comité de lecture répertoriées (ACL)

- Orfila O., Freitas Salgueiredo C., **Saint Pierre G.**, Sun H., Li Y., Gruyer D., Glaser S. (2016), « *Fast computing and approximate fuel consumption modeling for Internal Combustion Engine passenger cars* », Transportation Research Part D 50C (2017) pp. 14-25, DOI: 10.1016/j.trd.2016.10.016, (*JCR, Transport science, IF=1.864*).
- Romon S., Bressaud X., Lassarre S., **Saint Pierre G.**, Khoudour L., (2015) « *Map-matching algorithm for large databases* », Journal of Navigation, Volume 68, Issue 05, pp 971-988, DOI:10.1017/S0373463315000156 (*JCR, Engineering, 0.949*).
- Popa I. S., Kharrat, A., Zeitouni K., **Saint Pierre G.** (2009), « *Base de données de capteurs à localisation mobile* », Ingénierie des Systèmes d'information, 14, Lavoisier, pp 35-58, DOI: 10.3166/ISI.14.5.35-58 (*Scopus, Intelligence artificielle*).

Communications avec actes dans congrès international (ACTI)

- **Saint Pierre G.**, Tattegrain H., Val C. (2014), « *Impact evaluation of speed regulation systems using naturalistic driving data: The EuroFOT example* », Transport Research Arena 2014, Paris, La Défense.
- Romon S., **Saint Pierre G.**, Khoudour L. (2014), « *Fuel consumption and speed profiles comparisons according to infrastructure using probe vehicle data* », Poster au Transport Research Arena 2014, Paris, La Défense.
- Andrieu C., **Saint Pierre G.** (2012), « *Comparing effects of ecodriving training and simple advices on driving behavior* », 15th meeting of the Euro Working Group on Transportation, EWGT2012, 10-13 septembre, Paris, France.
- Andrieu C., **Saint Pierre G.** (2012), « *Using statistical models to characterize eco-driving style with an aggregated indicator* », Intelligent Vehicles Symposium 2012, June 3-7, 2012, Alcalá de Henares, Spain, DOI: 10.1109/IVS.2012.6232197.
- Benmimoun M., Ljung Aust M., Faber F., **Saint Pierre G.**, Zlocki A., (2011), « *Safety analysis method for assessing the impacts of advanced driver assistance systems within the European large scale field test “euroFOT”* », European congress on intelligent transport systems, ITS2011, Lyon.
- Lambert A., Gruyer D., **Saint Pierre G.**, Ndjeng, A.N. (2008), « *Collision Probability Assessment for Speed Control* », 11th International IEEE Conference on Intelligent Transportation Systems, 2008. DOI: 10.1109/ITSC.2008.4732692.
- Lambert A., Gruyer D., **Saint Pierre G.** (2008), « *A Fast Monte Carlo Algorithm for Collision Probability Estimation* », 10th International Conference on Control, Automation, Robotics and Vision, ICARCV, 17-20 décembre, Hanoi, Vietnam, DOI: 10.1109/ICARCV.2008.4795553.
- Lambert A., **Saint Pierre G.** (2008), Gruyer D., « *A Monte Carlo Approach for Collision Probability Computation* », IEEE IROS08 2nd Workshop on Planning, Perception and Navigation for Intelligent Vehicles, NICE, France.
- **Saint Pierre G.**, Ehrlich J. (2008), « *Impact of Intelligent Speed Adaptation systems on fuel consumption and driver behaviour* », 15th World Congress on Intelligent Transport Systems.
- Zeitouni K., Sandu Popa I., Ehrlich J., **Saint Pierre G.**, Dupin F., Glaser S. (2008), « *On Using in-Vehicle Sensor Data for Naturalistic Driving Analysis* », 15th World Congress on Intelligent Transport Systems and ITS America Annual Meeting 2008 7 ,pp.4576.
- Vrignon J., Rakotonirainy A., Gruyer D., **Saint Pierre G.** (2007), « *Impact of subjective factors on driver vigilance: a driving simulator study* », in Proceedings of the Third International Driver Behaviour and Training Conference.
- **Saint Pierre G.**, Aron M., Bergel-Hayat R., Violette E. (2007) « *Rain reconstruction from various weather-related data sets using logistic regression: methodology and applications* », paper n°07-0916, TRB, Washington.
- Aron M., Bergel-Hayat R, **Saint Pierre G.**, Violette E. (2007) « *Added Risk By Rainy Weather. Some elements for the French IRCAD-SARI traffic monitoring and information system for drivers and operators* », 11thWorld Congress on Transportation Research, Berkeley.

Communications avec actes dans congrès national (ACTN)

- **Saint Pierre G.**, Andrieu C. (2013), « *Statistique et sécurité routière : L'observation de la conduite en situation naturelle* », 45ème journées de Statistique, Mai 27-31, 2013, Toulouse, France.
- **Saint Pierre G.**, Andrieu C. (2010). « *Caractérisation de l'éco-conduite et construction d'un indicateur dynamique pour véhicules thermiques* ». PRAC 2010.

Communications invitées sans actes (INV)

- **Saint Pierre G.** (2014) « *From FOT to NDS Recent developments in UDRIVE The first large-scale European Naturalistic Driving Study* », exposé invité lors du 4ème International Symposium on Naturalistic Driving Research, Virginia Tech Transportation Institute, Blacksburg, 25-28 Août 2014.
- **Saint Pierre G.** (2013), « *Statistique et sécurité routière : L'observation de la conduite en situation naturelle* », conférencier invité aux 45ème journées de Statistique, Toulouse, mai 2013.
- **Saint Pierre G.** (2012), Expert invité au FOT-Net International Workshop, « *round table on Impact Assessment and Scaling up* », Vienne, Autriche, Octobre 2012.
- **Saint Pierre G.** (2012), « *Statistical modeling of safety events in naturalistic driving data: An application to speed regulation systems* », conférencier invité à la "5th International Conference on Traffic and Transport Psychology", ICTTP 2012, Groningen, The Netherlands, August 29-31, 2012.

Rapports de recherche y compris rapports de contrat de recherche (RAPP)

- Ledoux, V., **Saint Pierre, G.** (2016) « *Usage des véhicules sondes pour la mesure des vitesses pratiquées en France* », rapport pour le groupe de travail DSCR « évolution de l'observatoire des vitesses ».
- **Saint Pierre G.** et al. (2016), « *Eco driving in the real-world: behavioural, environmental and safety impacts* », deliverable 43.1 of the ecoDriver project, EU FP7.
- Tattegrain, H., Letisserrand, D., **Saint Pierre, G.**, Val, C., Krishnakumar, R. (2015) « *Méthodes d'évaluation de dispositifs embarqués utilisant la conduite en situation naturelle* », livrable 2 de la convention N°: 2200683288 (DSCR-Ifsttar), Étude «IDACA» : Evaluation de l'Impact des Dispositifs d'Assistance à la Conduite Automobile.
- Kircher K. et al., (2014), « *D41_1 Performance indicators and ecoDriver test design* », public deliverable, projet ecoDriver, FP7.
- Woldeab, Z. et al., (2014), « *D32.1: Individual Use Cases and Test Scenarios Definition* », public deliverable, projet ecoDriver, FP7.
- Nouveliere L. et al, (2014), « *D21.3: Powertrain Model Validation* », public deliverable, projet ecoDriver, FP7.
- Seewald P., Orfila O., **Saint Pierre G.**, (2014), « *D23.2: Simulation and analysis document for on-line vehicle algorithms* », public deliverable, projet ecoDriver, FP7.
- Malta L., et al. (2012), « *Final results: Impacts on traffic safety* », livrable D6.4 du projet euroFOT (FP7), 215 pages.

- Faber F., et al. (2012), « *Final results: Impacts on traffic efficiency and environment* », livrable D6.5 du projet euroFOT (FP7), 141 pages.
- Sanchez D. et al. (2012), « *Final results: User acceptance and user-related aspects* », livrable D6.3 du projet euroFOT (FP7), 427 pages.
- **Saint Pierre G.** et al. (2010), « *Report of the results of adaptation, in-vehicle implementations and piloting* », livrable D2.2 du projet EuroFOT.
- Aron M., Bergel-Hayat R., Debaène J-Y., Dodet L., Le-Breton P., Lemaire E., Louah G., **Saint Pierre G.**, Violette E. (2006), « *Sur-risque par temps de pluie* », rapport du WBS 1.6 du projet PREDIT IRCAD-SARI.

6.3. Application : Systèmes régulateurs de vitesse

Cette section présente l'application de la méthodologie présentée dans la section précédente (section 6.2) au cas des systèmes longitudinaux agissant sur la vitesse (cf. Figure 29). Ce type de système a constitué mon premier cadre d'étude de l'impact des systèmes embarqués. Je me suis tout d'abord intéressé à l'analyse de l'impact d'un tel système sur la consommation de carburant (système Lavia), pour ensuite élargir cette évaluation à l'étude de la sécurité dans le cadre de l'usage du régulateur et du limiteur de vitesse.

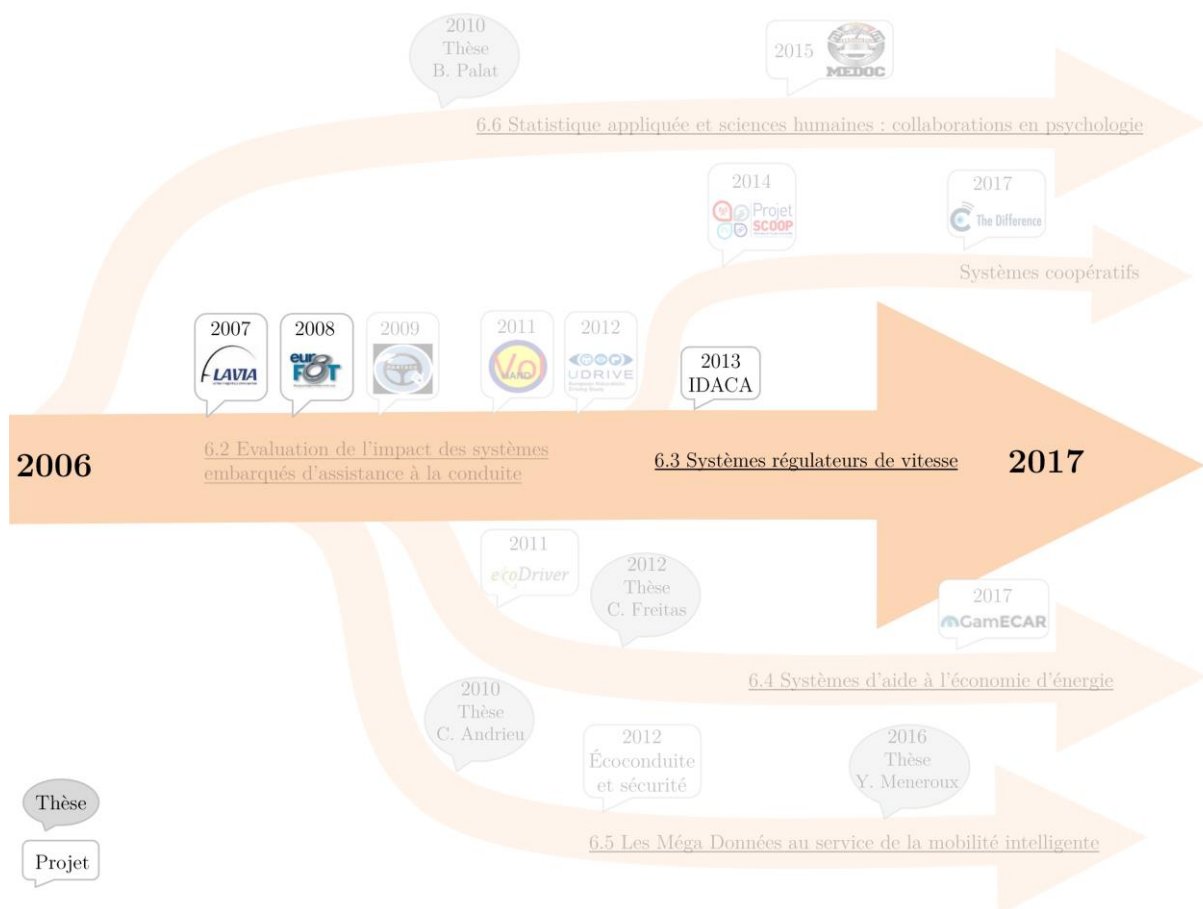


Figure 29 : Mise en évidence des axes de recherche développés dans cette section.

Cette section présente donc les résultats des projets Lavia et EuroFOT, obtenus par l'application des méthodes citées en 6.2. A la date de ces travaux, les études portant sur la consommation de carburant induite par les systèmes embarqués étaient rares et scientifiquement mal établies. Mon apport principal dans ces études, au-delà de la maîtrise des outils statistiques de gestion de grandes masses de données, a consisté à introduire l'usage des modèles statistiques pour données catégorielles non indépendantes (GEE ou GLMM) décrits à la section 6.2. Une illustration de

l'emploi de ce type de modèles pour l'évaluation de la consommation de carburant est présentée à la section 6.3.1.2.

6.3.1. Impact d'un limiteur de vitesse adaptatif sur la consommation de carburant (Projet LAVIA)

6.3.1.1. Le projet LAVIA

En 2004, le ministère français des transports (Direction de la Sécurité et de la Circulation Routière-DSCR) a lancé un ambitieux programme d'expérimentation destiné à évaluer les effets d'un limiteur de vitesse intelligent sur la réduction des vitesses pratiquées, ainsi que son acceptabilité sociale.

Le système LAVIA (Limiteur s'Adaptant à la Vitesse Autorisée) a été testé dans le département des Yvelines suivant différents modes de fonctionnement : neutre, informatif, actif débrayable et actif non débrayable (cf. encadré). L'expérimentation d'une durée totale d'un an, a concerné 90 conducteurs qui ont chacun testé pendant 8 semaines un des 20 véhicules équipés du système LAVIA. Une importante quantité d'information de conduite en situation naturelle a été recueillie grâce aux systèmes d'acquisition de données embarqués.

Lavia : quand le véhicule s'adapte de lui-même à la vitesse réglementaire

Fonctionnant grâce à un système embarqué de localisation par satellite, le LAVIA s'assure périodiquement que la vitesse du véhicule est conforme à la vitesse réglementaire de la route sur laquelle il circule. Au-delà de cette vitesse limite, la pédale d'accélérateur reste sans effet car un dispositif électronique limite automatiquement l'injection de carburant.

Il peut fonctionner selon trois modes d'assistance très différenciés :

En mode informatif : le conducteur lit sur le tableau de bord la vitesse autorisée. En cas de dépassement, il est informé par un affichage clignotant.

En mode actif volontaire : Dans ce mode, le conducteur choisit de mettre le dispositif en service ou de le couper par un bouton « marche-arrêt ». Lorsque le dispositif est en marche la vitesse est limitée : au-delà de la vitesse réglementaire, la pédale d'accélérateur devient sans effet. Toutefois, le conducteur peut à tout moment neutraliser temporairement l'effet du LAVIA grâce à un dispositif « kick down » (un appui fort sur la pédale d'accélérateur).

En mode actif contraint : le LAVIA ne peut être débranché. Le kick down reste disponible.

Le projet LAVIA, a été conçu initialement avec l'unique objectif de mesurer les impacts du système sur la sécurité. Il est cependant légitime de s'attendre à ce qu'un système limiteur de vitesse, contraignant les utilisateurs à respecter la vitesse légale en toutes circonstances, entraîne une économie de carburant. C'est d'ailleurs l'un des arguments forts avancés lors des campagnes de la sécurité routière pour inciter les Français à respecter les limitations de vitesse. Cependant, les choses ne sont pas si simples... la majorité des moteurs conçus par les constructeurs ont un rendement optimal pour des vitesses moyennes comprises entre 60 et 100 km/h, et les émissions de GES sont les plus faibles pour des vitesses stables comprises entre 40 et 90 km/h. Il s'ensuit qu'une réduction de 10 km/h de la vitesse légale, par exemple, n'aurait pas le même effet selon les types de réseau : urbain, inter-urbain ou autoroute. De la même façon, une vitesse contrainte à 30 km/h n'est certainement pas économique avec les moteurs thermiques actuels. Dans ces conditions, il faut poser une question : un système limiteur de vitesse contraignant a-t-il réellement un effet de réduction de la consommation moyenne ? La réponse intuitive est oui mais peut s'avérer trompeuse, aussi la validation d'une telle hypothèse nécessite quelques approfondissements. Lors de mon arrivée au LIVIC en 2007, j'ai donc souhaité évaluer scientifiquement cette hypothèse, par le biais d'un travail de recherche basé sur les données obtenues lors de l'expérience LAVIA.

L'objectif du projet LAVIA était relativement ambitieux par rapport aux expérimentations de l'époque, en général cantonnées à des zones limitées le plus souvent urbaines, et à des solutions technologiques parfois rudimentaires. En effet, le projet LAVIA préfigurait déjà les projets de type "Field Operational Test" (FOT), c'est à dire des évaluations des systèmes d'aides à la conduite à grande échelle et en conditions naturelles. Le meilleur exemple de ce type d'approche est le projet EuroFOT dans lequel j'ai été fortement impliqué, et dont on peut trouver une description détaillée sur le site web qui lui est dédié (<http://www.eurofot-ip.eu>). La méthodologie de ces FOT est maintenant bien au point et a fait l'objet d'un projet préliminaire pour la mettre au point (cf. FESTA consortium [2008]). Cependant, lors du lancement du projet LAVIA, peu d'expériences de ce genre avaient été tentées en France

Le projet disposait d'une flotte de 20 véhicules composée de 10 véhicules Renault Laguna II et de 10 véhicules Peugeot 307. Ces véhicules ont été conduits par 100 personnes différentes, chacune pour une durée de 8 semaines. Au total, cela représente plus de 1000 jours d'expérimentation par mode de fonctionnement (4 modes différents).

Le système se trouve dans l'état « Limitation active » lorsque le système LAVIA limite l'injection de carburant. Dans cet état, le système n'agit pas sur la pédale de frein, seule la pédale d'accélération demeure sans effet. On peut distinguer deux types de situations amenant le système à se trouver à l'état « Limitation active » :

- la vitesse du véhicule est supérieure à la vitesse de consigne suite à une accélération du véhicule. Cette accélération peut avoir deux origines : volontaire avec appui sur la pédale d'accélération, involontaire dans les descentes par exemple.
- le véhicule se trouve en infraction de vitesse suite à une transition de vitesse de consigne vers une limitation plus basse.

Etant donné que la consommation instantanée n'avait pas été mesurée, nous avons retenu des méthodes d'estimation et fait appel aux connaissances des constructeurs partenaires du projet LAVIA pour disposer d'un modèle de calcul de la consommation instantanée. Ce modèle est fourni par une cartographie, sorte de « carte d'identité » du moteur, obtenue en faisant tourner celui-ci à divers régimes et pour des pressions de pédales d'accélérateur variées. On obtient ainsi précisément la consommation d'essence en grammes/secondes pour l'ensemble des trajets disponibles (cf. Figure 27). Mais comme chacun le sait, la consommation de carburant s'avère très variable en fonction des situations de conduite. Le contexte est pour une grande part responsable du type de conduite, il est donc indispensable de bien distinguer l'influence du système étudié, de celle induite par les interactions avec l'environnement. Les données disponibles étant géoréférencées, elles permettent de connaître avec précision un des facteurs les plus influents sur la conduite : l'infrastructure. La consommation a donc été calculée et comparée selon que la conduite était urbaine, inter-urbaine ou autoroutière, ainsi que selon les limitations de vitesses (30 km/h, 50 km/h, 70 km/h, 90 km/h, 110 km/h).

Tableau 6 : Moyennes de la consommation de carburants en litres/100km pour chacune des zones étudiées, et selon le mode d'utilisation du LAVIA. La valeur de la statistique F de l'anova et sa p-value sont données dans les deux dernières colonnes.

Zone	Nombre de trajets	Consommation (litres/100km)	Neutre	Informatif	Débrayable	Contraint	F	p-value
30km/h	1898	Moyenne	16.79	17.24	19.07	18.29	20.91	<.0001
50km/h	6830	Moyenne	9.02	9.21	9.58	9.23	4.0	0.0075
70km/h	2660	Moyenne	13.59	13.62	14.13	13.61	7.92	<.0001
90km/h	2680	Moyenne	8.57	8.75	9.06	8.83	2.48	0.0592
110km/h	1870	Moyenne	7.58	7.66	7.72	7.50	0.98	0.4013

La méthodologie employée pour étudier les données LAVIA, ainsi que l'ensemble des résultats obtenus est présentée dans **Saint Pierre G.** (2010), « *Le système LAVIA et la consommation de carburant : méthodes exploratoires pour données de conduite en situation naturelle* » édité dans la collection « études et recherches » du laboratoire central des ponts et chaussées. Le chapitre 2 de ce document propose une description détaillée des différentes étapes du traitement des données. Dans le cas du LAVIA, l'approche consiste à découper les trajets selon les limitations de vitesse (seule variable situationnelle utilisée), ce qui permet d'obtenir des séquences de conduite comparables entre elles. Une analyse descriptive (Tableau 6) de la consommation de carburant lors de ces séquences, permet de constater des différences significatives selon le mode

d'utilisation du LAVIA et selon les limitations. A première vue, la consommation moyenne minimale est presque toujours obtenue dans le mode neutre (LAVIA non activé) testé au début de l'expérimentation. Le système LAVIA semble donc induire une augmentation de la consommation. Néanmoins, une telle analyse ne prend pas en compte un certain nombre de facteurs :

- Il peut exister un effet d'apprentissage du système. Il est possible par exemple, qu'un usage efficace du système ne soit obtenu qu'après un certain temps de pratique, et que la durée de l'expérimentation n'ait pas permis de le percevoir.
- Certains conducteurs peuvent ne pas avoir compris le système, ou peuvent ne pas avoir voulu adapter leur conduite. Si ces conducteurs sont des gros rouleurs, ils peuvent avoir biaisé les résultats.

Par ailleurs, le plan d'expérience utilisé a consisté à tester le système séquentiellement, dans un ordre fixe et prédéterminé, ce qui a pu renforcer la présence d'un effet d'apprentissage. Le second biais est courant dans les expériences à mesures répétées. Les modèles statistiques adaptés au contrôle de ces biais ont été détaillés à la section 6.2.6 (GEE et GLMM).

Une approche statistique plus pertinente a donc été menée pour analyser plus finement l'impact du système sur la consommation. Cette analyse est proposée à la section suivante afin de servir d'illustration de l'intérêt des modèles GEE et GLMM.

6.3.1.2. Impact sur la consommation

Si l'on observe la Figure 30 représentant les boîtes à moustache des valeurs de consommation pour chaque trajet limité à 50km/h en fonction de l'ordre des jours de l'expérience, on observe une tendance à la baisse qui n'est pas perceptible dans les résultats d'une anova (Tableau 6). La régression sur les quantiles à 95% montre une tendance globale à la baisse, tandis qu'un lissage adaptatif (LOESS) fait apparaître une augmentation de la consommation en mode débrayable qui semble disparaître ensuite. Nous souhaitons donc tester un modèle spécifiant que la consommation est modifiée dans chacun des modes par rapport au mode neutre selon des paramètres inconnus à estimer. La valeur estimée par le modèle de ces paramètres nous dira si il y a augmentation ou diminution significative.

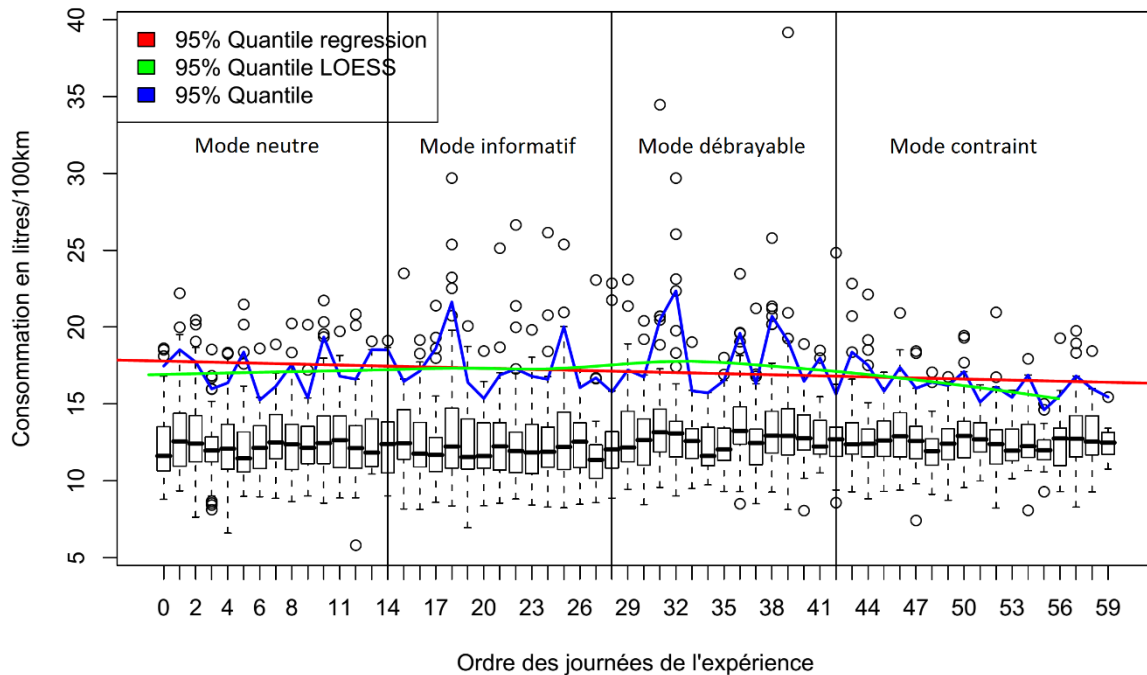


Figure 30 : Boîtes à moustache des consommations mesurées lors des trajets limités à 50km/h du LAVIA selon l'ordre des jours de l'expérience. La courbe reliant les quantiles à 95% est en bleu, la régression linéaire reliant ces mêmes quantiles est en rouge, tandis que la régression polynomiale par morceaux est en vert.

Etant donné que notre axe temporel d'évolution est celui correspondant à l'ordre d'usage des différents modes, il faut, pour chaque conducteur, calculer la consommation moyenne lors de chacun des modes, de façon à obtenir une unique valeur moyenne par mode. En effet, il n'y a pas lieu de modéliser un effet temporel plus fin dans la mesure où le passage d'une journée à l'autre n'a pas d'influence particulière sur la conduite.

On note y_{ct} la consommation de carburant moyenne du sujet c au temps t , où t représente l'ordre des différents modes du LAVIA testés (Neutre (1), Informatif (2), Débrayable (3) et Contraint (4)). Les valeurs y_{ct} correspondent donc aux moyennes de la consommation de carburant du conducteur c pour chacun des modes du LAVIA ($y_{c1}, y_{c2}, y_{c3}, y_{c4}$). On note $\mu_{ct} = E(y_{ct})$ l'espérance marginale de la consommation de carburant. Par la suite, l'indice c est omis de manière à simplifier l'écriture. Le modèle proposé est détaillé au Tableau 7.

Tableau 7 : Modèle proposé pour l'évolution de la consommation selon les modes du LAVIA.

$\mu_t = \beta_0 + \beta_1(t-1)$	si $1 \leq t \leq 2$
$\mu_t = \beta_0 + \beta_1 + \beta_2$	si $t = 3$
$\mu_t = \beta_0 + \beta_1 + \beta_2 + \beta_3$	si $t = 4$

Ce modèle s'écrit, en fonction de 3 variables t_1, t_2 et t_3 créés à partir de t sous la forme équivalente suivante :

$$\mu_t = \beta_0 + \beta_1 t_1 + \beta_2 t_2 + \beta_3 t_3,$$

où les valeurs de t_1, t_2 et t_3 en fonction de t sont données au Tableau 8. Le paramètre β_0 représente la consommation de carburant moyenne en mode neutre, β_1, β_2 et β_3 représentent respectivement les augmentations moyennes des consommations de carburant pour les modes informatif, débrayable et contraint.

Tableau 8 : valeurs de t_1, t_2 et t_3 en fonction de t .

t	1	2	3	4
t_1	0	1	1	1
t_2	0	0	1	1
t_3	0	0	0	1

Une étape importante est la modélisation de la variance. Dans les modèles marginaux il n'est pas nécessaire de définir la matrice des variances-covariances : la définition d'une matrice de corrélation de travail la plus proche possible de la réalité est suffisante. La matrice de corrélation choisie ne dépend que d'un paramètre ρ , signifiant que la corrélation entre deux trajets d'un même conducteur est constante (i.e. la corrélation provient soit du style de conduite, soit du type de trajet habituels). Ce paramètre représente donc la corrélation entre deux observations réalisées sur un même sujet. Les résultats sur l'ensemble des trajets sont fournis au Tableau 9.

$$\text{Var}(y_{ct}) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}.$$

Tableau 9 : Paramètres estimés par le modèle marginal, donnant l'effet des modes du LAVIA sur la consommation de carburant pour l'ensemble des trajets.

Paramètre	Estimation	IC 95%		p-value
β_0	9.8162	9.3825	10.2499	<.0001
β_1	0.0705	-0.2894	0.4304	0.7010
β_2	0.3184	-0.0613	0.6981	0.1002
β_3	-0.4900	-0.8139	-0.1661	0.0030

La consommation moyenne de carburant en mode neutre est de 9.82 litres aux 100km. La consommation moyenne reste stable pendant l'utilisation du LAVIA en mode informatif car le paramètre estimé (β_1) n'est pas significativement différent de 0 ($p=0.7$). Lors de l'usage du mode débrayable, la valeur estimée est positive (i.e. la consommation augmente) mais non significative ($p=0.1002$). Cela suggère cependant, la présence potentielle d'un effet d'apprentissage du système pendant le mode débrayable qui implique un mauvais usage du système et donc une augmentation de la consommation. La valeur du paramètre β_3 est quand à elle négative et significative. On peut donc affirmer que l'effet du mode contraint sur la consommation de carburant est significativement positif, c'est à dire qu'il diminue la consommation d'un demi litre pour 100km, par rapport au mode débrayable. Cette diminution de la consommation est supérieure à l'augmentation constatée pendant le mode débrayable, ce qui conduit à une diminution globale de l'ordre de 0,2 litres de carburant due au mode LAVIA. Il s'agit là d'un résultat valable sur l'ensemble du réseau parcouru lors de l'expérience LAVIA (i.e. le département des Yvelines).

D'autres analyses identiques ont été menées pour chacune des limitations de vitesses (30 km/h, 50 km/h, 70 km/h, 90 km/h, 110 km/h) qui permettent d'inférer les conclusions suivantes :

- Il n'y a pas d'effet significatif dans les zones limitées à 30km/h.
- Il existe bien un effet d'apprentissage du LAVIA pendant le mode débrayable, celui-ci induisant une sur-consommation dans les zones limitées à 50km/h, 70km/h et 90km/h. Cet effet est de moins en moins significatif à mesure que la vitesse limite est élevée. Cela peut signifier que le système LAVIA est plus difficile à maîtriser dans ces zones.
- Le mode contraint est efficace pour réduire la consommation dans les zones limitées à 110km/h. Dans les zones limitées à 50km/h l'effet positif est seulement dû à une meilleure gestion du système qui évite les erreurs de conduite observées lors du mode débrayable

6.3.1.3. Impact sur la sécurité (excès de vitesse)

Le projet LAVIA a fait l'objet de plusieurs études liées à son impact potentiel sur la sécurité. Pour le mode contraint les résultats obtenus (cf. Driscoll et al. 2007), font état d'une réduction potentielle du nombre de morts allant de 8 à 17% selon le type d'accident et de réseau, et pour un taux d'équipement de 100%. Ces résultats positifs sont obtenus en étudiant la distribution des vitesses, et sont pour l'essentiel dûs à une diminution de la variabilité. Un meilleur respect des limitations permet aussi de limiter les conséquences d'un accident. Le système LAVIA en mode contraint favorise un tel comportement comme l'atteste la Figure 31.

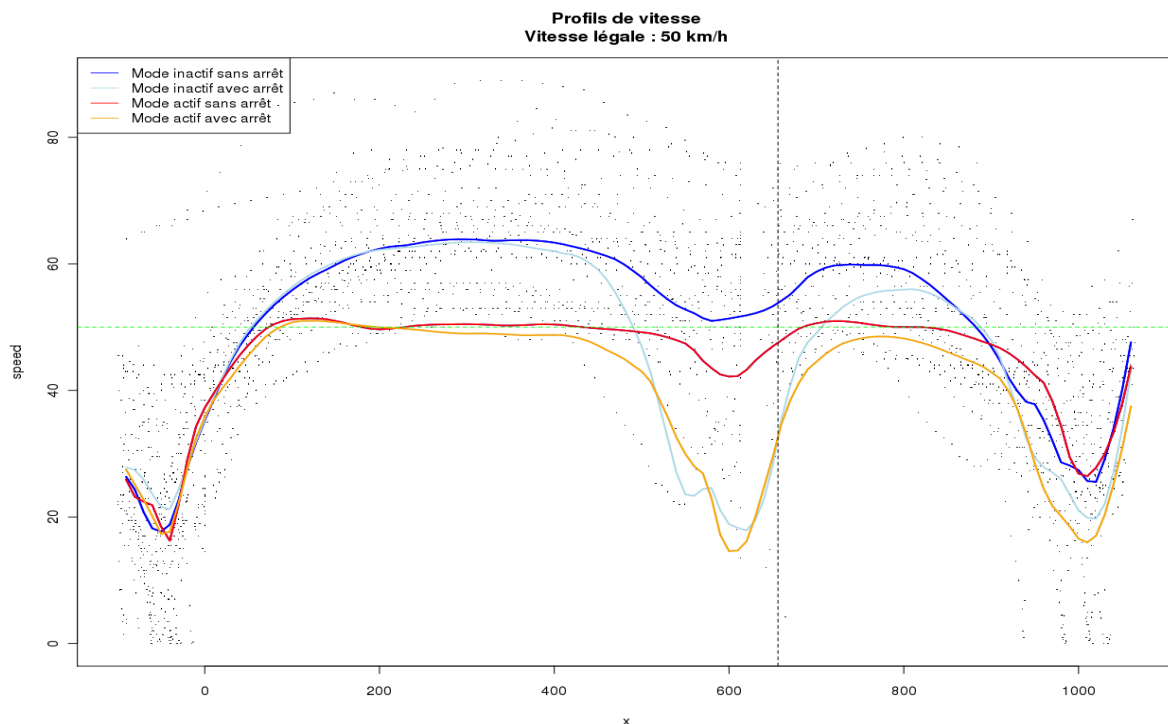


Figure 31 : Exemple de profils de vitesse moyens sur une section limitée à 50km/h avec présence d'un feu tricolore selon le mode du LAVIA et selon la présence d'un arrêt au feu.

Lors de ces travaux sur le LAVIA, j'ai souhaité explorer son impact sur la sécurité de façon détournée, en analysant l'influence de ce dernier sur la probabilité d'occurrence d'un excès de vitesse. En effet, il faut rappeler que les excès de vitesse sont toujours possible même lors du mode contraint pour deux raisons :

- Le passage à une limitation de vitesse inférieure à la précédente, entraînant une coupure de l'injection et une réduction progressive de la vitesse (on peut parler dans ce cas d'un excès de vitesse involotaire).
- La violation volontaire de la limitation avec l'utilisation du « kick-down ».

Une approche de type GEE (cf. section 6.2.6.2) permet de modéliser la probabilité d'occurrence d'un événement en fonction de différentes variables qualitatives. Une telle approche relève de la méthodologie détaillée à la section 6.2, et était novatrice pour cette application lorsque je l'ai menée en 2010.

Lors de la conduite avec un véhicule équipé du LAVIA, le pourcentage de temps passé en excès de vitesse reste compris entre 15 et 25%. Il a été décidé de classer les événements correspondant à des dépassements de vitesse dans deux catégories :

- Dépassement involontaire (type 1) : Dépassement de plus de 4km/h de la vitesse autorisée, et d'une durée inférieure à 10sec.
- Dépassement volontaire (type 2) : Dépassement de plus de 4km/h de la vitesse autorisée, et d'une durée supérieure à 10sec.
- Tout dépassement de moins de 4km/h correspond à la marge d'erreur du système et n'est donc pas considéré.

En effet, une étude approfondie des dépassements de vitesse et de leur durée, combinée avec une analyse des kick-downs (pression forte sur la pédale d'accélérateur pour débloquer le système) a permis de déduire qu'une durée de 10 secondes permettait une bonne discrimination entre les deux types de dépassement. La méthodologie employée pour construire notre jeu de données est similaire à la méthode d'agrégation numéro 2 (cf. section 6.2.5.2.4) et se détaille ci dessous :

- 1. Détecter et étiqueter les dépassements de vitesse dans le jeu de données complet.
- 2. Tirer au hasard 10000 sections de 30 secondes de conduite pour chacun des modes neutre et contraint (20000 sections au total).
- 3. Pour chaque section, noter la limitation de vitesse (retirer une nouvelle section si plusieurs limites de vitesses sont présentes).
- 4. Pour chaque section, mettre un 0 si il n'y a pas de dépassement de vitesse dans l'intervalle de 30 secondes, et un 1 si on est en présence d'un dépassement volontaire (type 2).
- 5. Appliquer la méthodologie des GEE pour déduire la probabilité d'apparition d'un dépassement volontaire en fonction du mode du LAVIA, et de la limitation de vitesse.

Tableau 10 : Paramètres estimés par le modèle marginal, donnant l'effet du mode du LAVIA et des limitations de vitesse sur la probabilité d'un dépassement de vitesse volontaire (type 2).

Paramètre	Estimation	IC 95%		p-value
Intercept	-1.6714	-2.0525	-1.2903	<.0001
ConsigneLimiteVitesse_30	0.7296	0.2451	1.2140	0.0032
ConsigneLimiteVitesse_50	0.4710	0.1010	0.8411	0.0126
ConsigneLimiteVitesse_70	1.0027	0.5789	1.4265	<.0001

ConsigneLimiteVitesse_90	0.6285	0.2561	1.0009	0.0009
ConsigneLimiteVitesse_110	0.0000			
ModeLavia_contraint	-0.5632	-0.7306	-0.3959	<.0001
ModeLavia_neutre	0.0000			

Les résultats de cette approche appliquée au seul mode contraint, sont présentés au Tableau 10, mais d'autres analyses ont été faites pour les autres modes.

Les lignes correspondant à une consigne de 110km/h et au mode neutre sont les références, à partir de laquelle est calculée la probabilité d'occurrence d'un dépassement. Ainsi, comparé à cette référence et toutes choses égales par ailleurs, le mode contraint entraîne une diminution significative de la probabilité d'occurrence d'un excès de vitesse. Le dépassement de vitesse pouvant être considéré comme un événement rare (par rapport à la fréquence de conduite respectant la limitation de vitesse), l'interprétation de l'exponentielle de ce coefficient peut se traduire par une multiplication par 0.569 ($\exp(-0.5632)$) de la probabilité d'effectuer un dépassement de vitesse lors de la conduite avec le mode contraint.

A l'inverse, concernant l'effet des limitations de vitesse, elles augmentent toutes la probabilité d'un dépassement de vitesse volontaire (paramètres positifs et significativement différents de 0) par rapport aux zones limitées à 110km/h. Les excès de vitesse volontaires sont plus probables dans les zones limitées à 70km/h (paramètre plus élevé). Nous venons donc d'apporter la preuve statistique que le système LAVIA en mode contraint réduit significativement la probabilité d'un excès de vitesse volontaire. Cette hypothèse a été testée de différentes façons (non détaillées ici) qui confirment ce résultat. D'après des analyses approfondies, l'effet du LAVIA sur l'occurrence des dépassements de vitesse serait plus marqué dans les zones limitées à 110km/h, en favorisant un meilleur respect des consignes.

6.3.1.4. Résumé des conclusions sur le projet LAVIA

Les résultats de l'analyse que j'ai effectuée mettent tout d'abord en avant un effet d'apprentissage : les conducteurs découvrant le système ont une conduite plus consommatrice que ceux qui l'expérimentent depuis déjà quelques semaines. Cet effet induit donc une surconsommation par rapport à la conduite normale (sans le système) dans les premières semaines d'usage, surtout pour les zones limitées à 50km/h, 70km/h et 90km/h. Puis, après quelques semaines, la consommation redevient comparable à celle de la conduite normale (sans le système). Une réduction de la consommation notable est observée pour les limitations de vitesse élevées (110km/h). Sur l'ensemble du réseau parisien, c'est une réduction de l'ordre de de l'ordre de 0.2 litres/100km qui peut être espérée. De même la réduction des épisodes d'excès de vitesse est plus flagrante sur les zones limitées à 110km/h, mais tout de même observée sur l'ensemble du réseau.

Gardons à l'esprit que le LAVIA a été pensé et conçu au départ pour réduire les vitesses et donc les risques d'accident. De ce point de vue, il remplit parfaitement ses objectifs (cf. Romon 2006). Mais la réduction de consommation de carburant ne semble pas être un sous-produit systématique de la réduction des vitesses pratiquées. En effet, de multiples facteurs impactent la consommation, comme par exemple les stratégies de changement de rapports de boîte de vitesse ou les accélérations et les freinages brusques. Que peut-on en conclure ? Une aide à la conduite complète et efficace ne peut plus se contenter d'intervenir uniquement sur le risque. Elle doit s'accompagner de la prise en compte des questions environnementales et offrir au conducteur de véritables conseils de conduite pour réduire la consommation mais seulement lorsque cela est possible. Il ne faut pas faire de l'écologie au détriment du risque routier... et inversement.

Pour aller dans ce sens, j'ai orienté mes recherches vers l'étude de l'écoconduite, et j'ai en particulier cherché à étudier dans le détail les relations entre le risque et l'éco-conduite, afin de concevoir à terme une aide à la conduite complète et fiable. Ces travaux seront présentés dans la section 6.4.

6.3.2. Impact du limiteur et du régulateur de vitesse (Projet EuroFOT)

Le projet européen EuroFOT (European Field Operational Test on Active Safety Systems) a débuté en 2009, et a bénéficié d'une subvention européenne de 14M€ partagés entre 28 partenaires académiques et industriel. Ce projet a permis la mise en place de tests opérationnels en conditions réelles, destinés à évaluer statistiquement l'impact de huit systèmes avancés d'assistance au conducteur. J'ai pris une part très active dans ce travail en assumant la responsabilité d'un « work package » dédié à la coordination des expérimentations pilotes, et en participant à l'analyse d'un des systèmes étudiés : le limiteur / régulateur de vitesse.

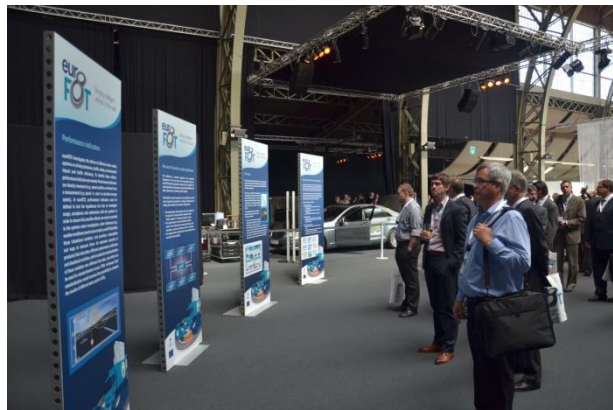


Figure 32 : Séminaire de clôture du projet euroFOT (<http://www.eurofot-ip.eu/>).

Au total, c'est presque 1000 véhicules personnels instrumentés et équipés de ces systèmes d'assistance qui ont parcourus l'Europe pendant une période d'un an. L'IFSTTAR et ses partenaires ont ainsi collecté des données précieuses pour évaluer l'impact de tels systèmes sur la sécurité, l'efficacité et le confort du conducteur. La France avait pour objectif d'évaluer le limiteur et le régulateur de vitesse (Speed Limiter et Cruise Control) à l'aide de 35 véhicules instrumentés. Les résultats de ce projet de 4 ans ont été présentés les 26 et 27 juin 2012 à Bruxelles, en présence des principaux acteurs européens de la recherche sur le transport.

EuroFOT a été une occasion unique pour moi de contribuer à améliorer les véhicules et la sécurité sur nos routes. Les données recueillies apportent en effet une connaissance inédite sur l'usage réel des véhicules et de leurs équipements. Cela m'a permis de mieux comprendre les conditions d'usage du limiteur/régulateur et d'en évaluer le bénéfice.

Les données utilisées pour l'évaluation du limiteur/régulateur ont été collectées par le CEESAR (Centre Européen Etudes de Sécurité et Analyse des Risques, <http://www.ceesar.fr/>) qui a eu la responsabilité de l'instrumentation de 35 véhicules de particuliers volontaires (Renault Clio et Renault Laguna III). La conduite de ces personnes a été enregistrée pendant 3 mois sans que le système ne soit utilisé, puis pendant 9 mois en les laissant libre d'utiliser le système selon leurs désirs. Ce protocole a conduit à l'enregistrement des données décrites à la table suivante :

Tableau 11 : Bilan des données de conduite recueillies dans euroFOT pour l'analyse du limiteur et du régulateur de vitesse.

Distance parcourue [km]	545 340
Heures de conduite [h]	12 590
Nombre de conducteurs	35

Outre l'enregistrement de données objectives, les conducteurs ont rempli trois questionnaires différents destinés aux analyses des partenaires en sciences humaines.

La méthodologie d'analyse utilisée a été décrite à la section 6.2. Le premier travail a consisté à rédiger une série de questions de recherche, chacune d'entre elle étant associée à un indicateur permettant d'inférer une réponse statistiquement validée. Les hypothèses de recherche que j'ai analysées sont détaillées au Tableau 12.

Tableau 12 : Liste des hypothèses de recherche testées lors du projet euroFOT pour l'étude du système régulateur de vitesse (Speed regulation system, SRS) composé du limiteur (speed limiter, SL) et du régulateur (cruise control, CC). Le chunk désigne un intervalle de temps de 30 secondes de conduite, homogène selon la limitation de vitesse (cf. section 6.2.5.2.4).

	Hypothesis	Performance indicator
Safety	Using SRS, the number of incidents will decrease	Number of incidents per 100 driven km
	SRC decreases the number of critical time gaps to the leading vehicle	Presence of a critical time gap event in the chunk
	SL/CC use increases over time	Average per month of the driver's % of SL or CC usage per day
	Using SL/CC, the number of harsh braking will decrease.	Presence of a hard braking event in the chunk
	Using SL/CC, the number of strong jerk will decrease.	Presence of a strong jerk event in the chunk
	Using SL/CC, reduces speeding occurrences	Presence of an over speeding event in the chunk
Safety	SL/CC decreases average speed.	Average speed per chunk
	The number of trips made will increase	Number of trips per month

	The number of vehicle km travelled will increase	Number of km travelled per month
Environ	SL/CC reduces the average fuel consumption	Average fuel consumption per chunk
Driver Behaviour	CC will be used more on roads with few curves or intersection	Presence of a CC-active event in the chunk, and relation with intersection or curve density
	SL will be used more on roads with few curves or intersection	Presence of a SL-active event in the chunk, and relation with intersection or curve density
	The SL/CC selected speed will be below legal speed for non-sensation seekers drivers and will be above legal speed for sensation seekers drivers	Differences between SL or CC selected speed and legal speed limit

6.3.2.1. Impact sur la sécurité

6.3.2.1.1. Analyse par évènements

Cette section présente une mise en œuvre de l'analyse de l'impact des systèmes régulateurs de vitesse sur la sécurité, par le biais de l'étude de l'occurrence des événements suivants :

- Excès de vitesse (+de 4km/h pendant au moins 10 secondes).
- Excès de jerk ($\text{Jerk} > 1 \text{ m/s}^3$).
- Temps inter-véhiculaire critique ($\text{THW} < 2.5\text{s}$).
- Freinage brusque (accélération $< -6\text{m/s}^2$).

Le système de régulation de vitesse étudié possède deux composants avec des objectifs différents : Le limiteur de vitesse dont l'utilité est de permettre au conducteur de définir volontairement un seuil de vitesse à ne pas dépasser, et le régulateur de vitesse dont l'utilité est de maintenir le véhicule à une vitesse constante définie par le conducteur.

Ces deux composantes ne peuvent être activées en même temps, et le choix d'activation par l'utilisateur de l'un ou l'autre dépend de nombreux facteurs liés aux conditions de conduite qu'il est difficile de contrôler. L'étude de l'influence de certains de ces facteurs sur les probabilités (odds ratios) d'usage d'un de ces systèmes mène aux conclusions suivantes :

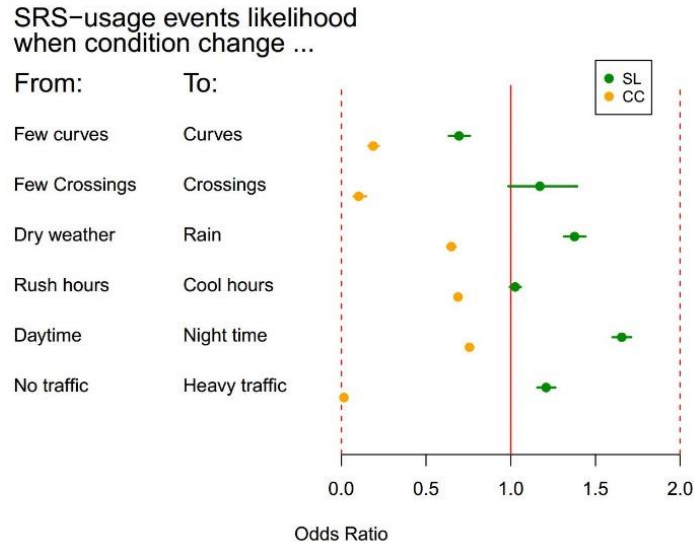


Figure 33 : Odds-ratios de l'usage du limiteur (SL) et du régulateur (CC) selon les conditions de conduite (SRS=Speed regulation system).

Des conditions de conduite contraintes conduisent à diminuer l'emploi du régulateur, particulièrement la présence d'intersections, de virages fréquents, ou un trafic élevé. Le limiteur de vitesse est quant à lui plus fréquemment utilisé la nuit ou en présence de pluie.

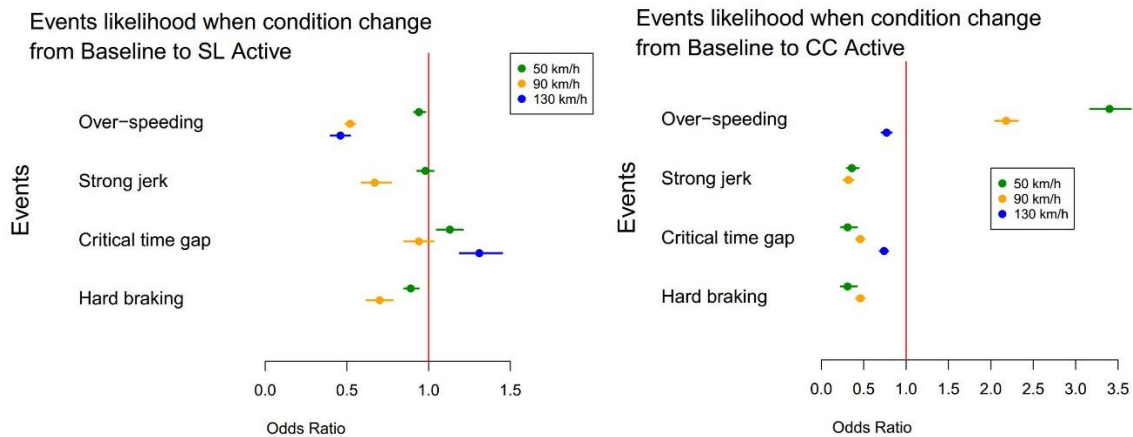


Figure 34 : Odds ratios de certains événements de la conduite lors de l'usage du limiteur (SL, gauche) et du régulateur (CC, droite).

En outre, j'ai pu mettre en évidence que lorsque le limiteur est activé, l'occurrence d'un dépassement de vitesse est divisée par deux sur les infrastructures avec une limitation de vitesse supérieure ou égale à 90 km/h, tandis que les freinages brusques sont diminués de 30% (cf. Figure 34). Cependant, il n'y a pas d'effet significatif sur l'occurrence des jerks excessifs et des temps inter-véhiculaires critiques. En ce qui concerne le régulateur de vitesse, son utilisation entraîne à l'inverse une forte augmentation de l'occurrence des excès de vitesse sur les zones limitées à 50 et 90 km/h, mais celle des jerks excessifs, des temps inter-véhiculaire critiques, et des freinages

brusques, est quant à elle divisée par trois (cf. Figure 34). Cela n'est pas surprenant dans la mesure où le régulateur est rarement utilisé en heures de pointes.

En ce qui concerne le type de réseau routier favorable à ces systèmes, autant le limiteur est utilisé sur l'ensemble du réseau, autant le régulateur est surtout utilisé pour les routes à limitation de vitesse élevée (110 et 130 km/h) (cf. Figure 35).

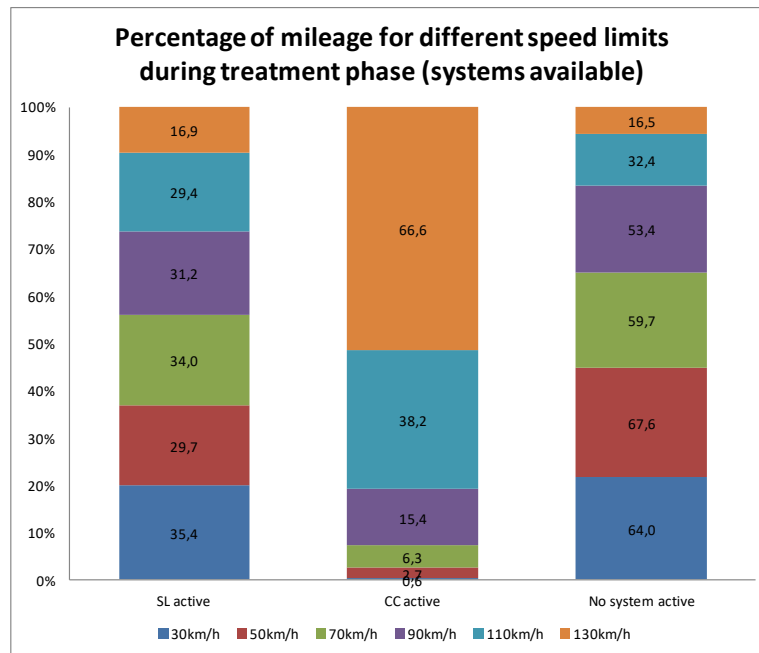


Figure 35 : Répartition des kilomètres parcourus selon le système (Speed limiter SL, cruise control CC) et les limitations de vitesse.

L'ensemble des résultats obtenus lors de ce projet (non reproduits ici dans leur intégralité) a permis d'émettre les suppositions suivantes quant aux caractéristiques de ces systèmes.

- Le limiteur de vitesse est utilisé sur tout type de réseau, plutôt dans des conditions où le respect des limitations semble prioritaire pour assurer une conduite en sécurité (faible limitation de vitesse, pluie, nuit).
- Le régulateur est plutôt utilisé sur les routes à limitation de vitesse élevée, ou avec des conditions de faible trafic. On peut conjecturer que ces situations sont perçues comme permettant une vitesse stable et plutôt élevée sans détériorer le niveau de sécurité.

Les deux systèmes diminuent l'occurrence d'événements à risque, malgré l'augmentation de l'occurrence des épisodes d'excès de vitesse constatés lors de l'usage du régulateur. Le système de régulation de la vitesse est donc globalement positif en termes de sécurité routière, même s'il est difficile d'évaluer précisément cet effet.

6.3.2.1.2. Analyse globale et extrapolation des résultats

L'extrapolation des résultats d'une telle étude est délicate pour plusieurs raisons. D'une part, les deux systèmes ne peuvent être utilisés en même temps, et le choix de l'un d'entre eux dépend de plusieurs facteurs non contrôlables. Il est ainsi très délicat d'extrapoler les effets d'un seul des deux systèmes pris séparément (quel aurait le choix de vitesse du conducteur si celui-ci ne disposait que d'un unique système ?). D'autre part, les données recueillies ne concernent que 35 conducteurs français, et la quantité de données recueillie ne permet pas de prétendre à une bonne représentativité de notre échantillon.

Comme on l'a vu dans les sections précédentes, l'extrapolation du risque d'accident peut se faire à partir de l'évolution de la vitesse moyenne. Dans notre cas, la vitesse moyenne ne peut pas se déduire uniquement des instants où le système est activé, car il nous manque dans ce cas un élément de comparaison (la vitesse moyenne dans des conditions de conduite équivalentes, c.a.d. la même route, mais sans aucun système activé). Il a donc été choisi d'utiliser la vitesse moyenne pendant la période où le système était disponible (mais activé selon la volonté du conducteur), comparativement à la période où le système n'était pas disponible (désactivé). Comme indiqué au Tableau 13, une telle analyse globale permet de mettre en évidence une augmentation significative de la vitesse moyenne. L'application du modèle puissance de Nielsson (cf. section 6.2.7.2) conduirait donc à une dégradation de la sécurité lorsque les véhicules sont équipés du limiteur et du régulateur. Cependant, la simple considération de la vitesse moyenne ne suffit pas à prouver un effet négatif des systèmes régulateurs de vitesse sur la sécurité. En effet, les données euroFOT ont aussi permis de mettre en évidence un resserrement de la distribution des vitesses instantanées autour de la valeur moyenne. La variabilité des vitesses est donc réduite par le système, ce qui a un effet positif sur la sécurité.

Tableau 13 : évaluation des effets globaux du limiteur et du régulateur de vitesse.

Evolution du temps de parcours moyen	Urban	-1.4%
	Rural	-0.8%
	Motorway	-2.4%
Evolution de la mobilité (kilomètres parcourus)		0%
Evolution de la vitesse moyenne	Urban	+1.4%
	Rural	+0.8%
	Motorway	+2.4%

Les systèmes de régulation de la vitesse sont des systèmes améliorant le confort de conduite et n'ont pas été conçus dans un but d'amélioration de la sécurité. Le projet euroFOT a permis de mettre en évidence des effets peu marqués et susceptibles de se compenser. Il n'a pas non plus été constaté de présence d'effets pervers ou négatifs sur la conduite. Il est donc raisonnable, au vu de ces informations, de prétendre à des effets inexistantes sur la sécurité lorsque les véhicules sont équipés d'un système de régulation de la vitesse. Il est néanmoins possible d'affirmer que le limiteur de vitesse est un outil pertinent pour aider le conducteur à maîtriser son niveau de sécurité en l'aidant à respecter les limitations, tandis que le régulateur est un outil de confort utilisé en situation de faible risque engendrant parfois une conduite plus rapide et plus risquée.

Étant donné la faible représentativité de l'échantillon et les faibles effets observés sur la sécurité, le projet euroFOT n'a pas procédé à une extrapolation à l'échelle européenne.

6.3.2.2. Impact sur la consommation

Mon travail dans ce projet a aussi consisté à évaluer l'impact des systèmes régulateurs de vitesse (SL & CC) sur la consommation de carburant, selon la méthodologie détaillée à la section 6.2.

En ce qui concerne le limiteur de vitesse, la réduction de consommation observée est toujours significative, allant de 1.55% sur autoroute, 3.75% en milieu rural, jusqu'à 5.19% en milieu urbain. Ces effets positifs peuvent être attribués à une plus faible variabilité des vitesses lorsque le système est activé. L'extrapolation des effets à l'échelle européenne n'a été faite que pour les autoroutes, ce qui, une fois combiné au taux observé d'usage du système, conduit à une diminution globale de la consommation de carburant (et donc des émissions de CO₂) de 0.26% si l'ensemble des véhicules européens était équipé (cf. Figure 36).

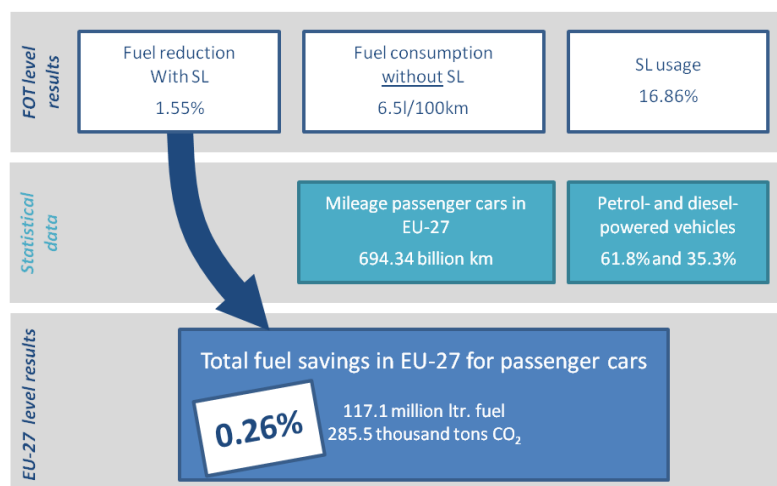


Figure 36 : économies de carburant potentielles en cas d'équipement complet de la flotte des véhicules personnels avec le limiteur de vitesse dans l'Europe à 27.

Le régulateur de vitesse, produit quant à lui une réduction des émissions de 1% sur le milieu autoroutier, 13.2% en milieu rural et 36.3% en milieu urbain. Ces deux dernières estimations, bien que significatives, sont fortement biaisées par les situations dans lesquelles les conducteurs choisissent d'utiliser le système (conduite plutôt de nuit et sans trafic). Par ailleurs, le faible taux d'usage du régulateur en milieu urbain et rural (2.7%) conduit à rejeter toute conclusion fondée sur ces données. En conséquence, seul le milieu autoroutier a été extrapolé à l'échelle européenne. Combiné à un fort taux d'usage de 66%, cela conduit à une réduction globale de la consommation de 0.7% (cf. Figure 37).

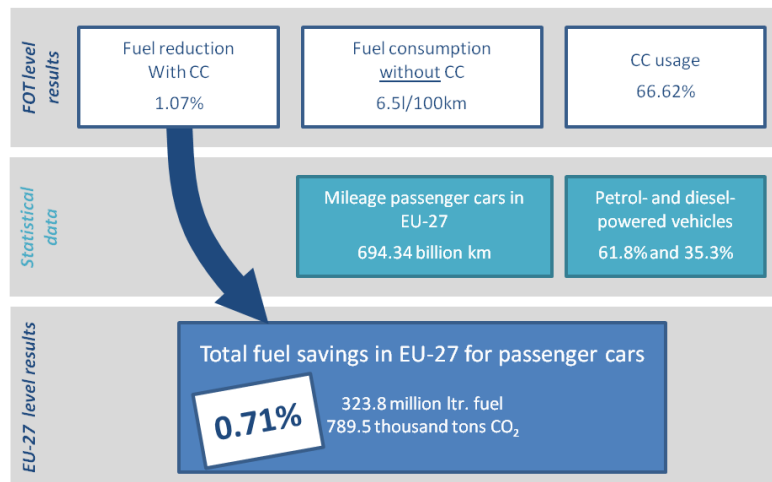


Figure 37 : Economies de carburant potentielles en cas d'équipement complet de la flotte des véhicules personnels avec le régulateur de vitesse dans l'Europe à 27.

6.3.3. Publications associées à cette section

Ouvrages scientifiques (OS)

- **Saint Pierre G.** (2010), « *Le système LAVIA et la consommation de carburant : méthodes exploratoires pour données de conduite en situation naturelle* », collection "études et recherches du laboratoire des ponts et chaussées", CR54, 120 pages.

Articles dans des revues à comité de lecture répertoriées (ACL)

- Popa I. S., Kharrat, A., Zeitouni K., **Saint Pierre G.** (2009), « *Base de données de capteurs à localisation mobile* », Ingénierie des Systèmes d'information, 14, Lavoisier, pp 35-58, DOI: 10.3166/ISI.14.5.35-58 (*Scopus, Intelligence artificielle*).

Communications avec actes dans congrès international (ACTI)

- **Saint Pierre G.**, Tattegrain H., Val C. (2014), « *Impact evaluation of speed regulation systems using naturalistic driving data: The EuroFOT example* », Transport Research Arena 2014, Paris, La Défense.

- Romon S., **Saint Pierre G.**, Khoudour L. (2014), « *Fuel consumption and speed profiles comparisons according to infrastructure using probe vehicle data* », Poster au Transport Research Arena 2014, Paris, La Défense.
- Benmimoun M., Ljung Aust M., Faber F., **Saint Pierre G.**, Zlocki A., (2011), « *Safety analysis method for assessing the impacts of advanced driver assistance systems within the European large scale field test “euroFOT”* », European congress on intelligent transport systems, ITS2011, Lyon.
- **Saint Pierre G.**, Ehrlich J. (2008), “*Impact of Intelligent Speed Adaptation systems on fuel consumption and driver behaviour*”, 15th World Congress on Intelligent Transport Systems.

Communications invitées sans actes (INV)

- **Saint Pierre G.** (2013), « *Statistique et sécurité routière : L'observation de la conduite en situation naturelle* », conférencier invité aux 45ème journées de Statistique, Toulouse, mai 2013.
- **Saint Pierre G.** (2012), Expert invité au FOT-Net International Workshop, « *round table on Impact Assessment and Scaling up* », Vienne, Autriche, Octobre 2012.
- **Saint Pierre G.** (2012), « *Statistical modeling of safety events in naturalistic driving data: An application to speed regulation systems* », conférencier invité à la "5th International Conference on Traffic and Transport Psychology", ICTTP 2012, Groningen, The Netherlands, August 29-31, 2012.

Rapports de recherche y compris rapports de contrat de recherche (RAPP)

- Tattegrain, H., Letisserrand, D., **Saint Pierre, G.**, Val, C., Krishnakumar, R. (2015) « *Méthodes d'évaluation de dispositifs embarqués utilisant la conduite en situation naturelle* », livrable 2 de la convention N°: 2200683288 (DSCR-Ifsttar), Étude «IDACA» : Evaluation de l'Impact des Dispositifs d'Assistance à la Conduite Automobile.
- Malta L., et al. (2012), « *Final results: Impacts on traffic safety* », livrable D6.4 du projet euroFOT (FP7), 215 pages.
- Faber F., et al. (2012), « *Final results: Impacts on traffic efficiency and environment* », livrable D6.5 du projet euroFOT (FP7), 141 pages.
- Sanchez D. et al. (2012), « *Final results: User acceptance and user-related aspects* », livrable D6.3 du projet euroFOT (FP7), 427 pages.
- **Saint Pierre G.** et al. (2010), « *Report of the results of adaptation, in-vehicle implementations and piloting* », livrable D2.2 du projet EuroFOT.

6.4. Application : Systèmes d'aide à l'économie d'énergie

6.4.1. Introduction

Mes travaux sur les systèmes embarqués agissant sur la vitesse m'ont permis de comprendre que le lien entre la vitesse pratiquée et la consommation de carburant n'était pas si clair. Dès lors, m'intéressant aux économies d'énergies par sensibilité personnelle, j'ai cherché à comprendre quels comportements de conduite pouvait générer les économies d'énergies espérées. Je me suis donc intéressé à la pratique de l'écoconduite dès 2009 et ai cherché à en quantifier expérimentalement les bénéfices. A cette date, très peu d'études indépendantes existaient sur ce sujet, et certains constructeurs automobiles annonçaient des performances exceptionnelles et assez peu crédibles (jusqu'à 20% de réduction de consommation). J'ai donc mené personnellement des campagnes d'études à petite échelle, qui m'ont valu une expertise dans ce domaine me permettant d'intégrer plus tard de grands projets européens (ecoDriver), ou de coordonner une partie des recherches du ministère de l'écologie (opération de recherche SERRES).

Les résultats présentés dans cette section ont été obtenus lors des projets et études suivants :

- 2009 : étude expérimentale personnelle sur l'écoconduite (non financée).
- 2009-2013 : coordination de l'action 5 de l'opération de recherche SERRES (solutions pour une exploitation de la route respectueuse de l'environnement et de la sécurité) conjointe entre le Cerema et l'IFSTTAR.
- 2012-2013 : Ecoconduite et sécurité en milieu urbain (financement DSCR, 30k€) en partenariat avec le laboratoire de psychologie de la conduite.
- 2011-2016 : Projet ecoDriver (financement UE-FP7, 14M€).

Le positionnement de cet axe de recherche par rapport à l'ensemble de mes travaux est présenté à la Figure 38, qui met en évidence les projets financés et les thèses encadrées dans ce contexte. Il faut noter que le projet européen GamECAR indiqué pour l'année 2017 a été obtenu lors de ma dernière année de présence au LIVIC et que j'ai été forcé de m'en retirer, mon nouvel établissement (le Cerema) ne faisant pas partie du consortium. Un des objectifs de ce projet était d'explorer l'emploi des méthodes de ludification (gamification pour les anglo-saxons) pour améliorer les performances des systèmes d'aide aux économies d'énergies.

Cette section propose tout d'abord une introduction (6.4.2) au concept d'écoconduite et une revue des paramètres de la conduite ou de l'infrastructure influençant la consommation de carburant (6.4.3 et 6.4.5) ainsi que les liens entre vitesse pratiquée et consommation de carburant (6.4.4). Les informations présentées sont issues de travaux personnels menés dans différents contextes, que j'ai synthétisés dans le cadre d'une formation donnée pendant plusieurs années à l'école des Ponts et Chaussées.

La sous-section 6.4.6 présente quelques résultats publiés concernant la caractérisation de l'écoconduite et son impact sur le comportement et la consommation de carburant. Ce travail se fonde sur l'analyse d'une série d'indicateurs associés aux règles d'or de l'écoconduite, que j'ai proposés pour évaluer le niveau de mise en œuvre de l'écoconduite.

La sous-section 6.4.7 présente l'application de ces résultats à la conception d'un indicateur agrégé reflétant le niveau d'écoconduite en temps réel, et sa mise en œuvre concrète dans le cadre de la conception d'une application pour smartphone d'aide à la conduite économe (projet européen ecoDriver, sous-section 6.4.8).

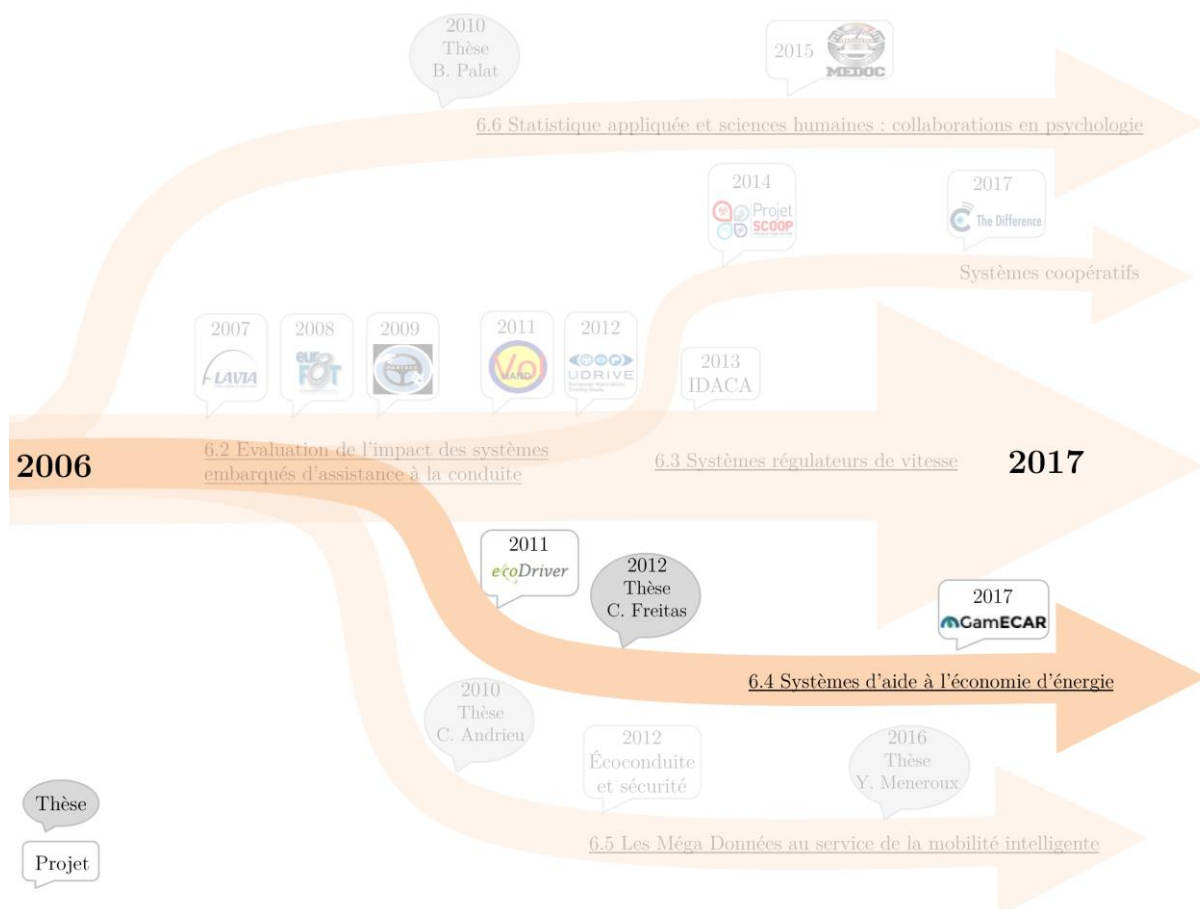


Figure 38 : Mise en évidence des axes de recherche développés dans cette section.

6.4.2. La conduite éco-responsable (écoconduite)

D'après la définition adoptée au ministère de l'écologie, l'écoconduite est un comportement de conduite citoyen permettant de réduire sa consommation de carburant. Suite à la prise de conscience du changement climatique, ce comportement a été revalorisé, au point qu'il est parfois référencé comme le « nouveau style de conduite ».

Pourtant, cette technique de conduite est connue et encouragée depuis longtemps. En France, on peut retrouver la trace de vidéos promotionnelles éditées par l'agence pour les économies d'énergie dès 1976 (http://www.dailymotion.com/video/x3cdr4_1976-agence-pour-les-economies-d-en_news). Hélas, les tentatives d'adoption de ce comportement vertueux se sont heurtées à de fortes résistances :

- Un prix du carburant trop faible pour être vraiment incitatif.
- Une vision sociale de la voiture associée au progrès et à la rentabilité, mais aussi au confort et au plaisir.
- La place de la voiture n'a fait que croître... aidée par un urbanisme conciliant.

La question centrale du réchauffement climatique et le contexte de crise économique, a accru la préoccupation pour les questions d'économies de carburant. L'écoconduite s'est donc naturellement imposée comme un moyen de concilier sécurité, économie de carburant, et temps de parcours. Cependant, dès lors qu'il s'agit de rajouter à la conduite un critère d'économie aux critères d'efficacité et de sécurité, la tâche de conduite devient un processus de plus en plus complexe. En effet, les paramètres influant sur la consommation énergétique d'un véhicule sont multiples, et l'écoconduite est une pratique complexe nécessitant entraînement et réflexion.

Lors de mes recherches, ma volonté de participer à l'encouragement de cette pratique m'a conduit à me poser les questions suivantes :

- Quel est l'impact des formations sur le long terme ?
- Peut-on pratiquer l'écoconduite efficacement sans formation spécifique ?
- Agir sur les vitesses pratiquées suffit-il à garantir une conduite sûre et économe ?
- Quels sont les liens entre conduite éco-responsable, respect des vitesses, sécurité, et économies d'énergies ?
- Comment construire un système embarqué d'aide à l'écoconduite ?

Cette section propose un résumé de l'ensemble des recherches que j'ai menées et coordonnées autour de ces questions. Les résultats présentés ont été obtenus lors de différents projets financés (nationaux et internationaux) mais aussi d'études sur fonds propres. Lors de ces différentes études, j'ai coordonné et organisé plusieurs expérimentations en conditions de conduite réelles grâce aux moyens expérimentaux du LIVIC. En particulier, j'ai pu disposer de deux véhicules Clio III instrumentés pour effectuer du recueil de données (Figure 39).



Figure 39 : Clio III du LIVIC et système de recueil de données du véhicule et des capteurs additionnels (radar et caméras).

Tout au long de ces études expérimentales, j'ai pu bénéficier d'un soutien technique exceptionnel de la part de mes collègues ingénieurs au LIVIC Benoît Lusetti et Jean-Christophe Smal, ainsi que de l'assistance de personnels du Cerema Didier Dominois et Christian Leverger. Cela m'a permis de mener à bien des expériences ambitieuses impliquant des dizaines de conducteurs différents, et de recueillir plusieurs milliers de kilomètres de conduite dans différentes conditions. Les résultats présentés dans cette section sont fondés sur ces données ainsi que sur certains jeux de données externes (Données du projet LAVIA, ainsi que du projet ANR Gerico notamment).

6.4.3. Facteurs de réduction de consommation et règles d'or de l'écoconduite

Il existe de nombreuses études détaillées concernant la consommation de carburant et les gains susceptibles d'être obtenus par diverses actions. Des programmes de formation des conducteurs et d'apprentissage de l'écoconduite peuvent permettre d'espérer une réduction de la consommation moyenne entre 5 à 20% selon les conditions (impact immédiat ou de long terme, formation spécifique ou assistance embarquée, conducteurs novices ou professionnels).

La variabilité très grande des effets observés, est en partie dû à la complexité et au nombre de facteurs pouvant influencer sur la consommation de carburant. De manière globale, l'état de l'art identifie les principaux facteurs pouvant avoir une influence sur la consommation de carburant. On peut schématiquement distinguer :

Les paramètres (fixes) relatifs au véhicule proprement dit.

- Silhouette du véhicule.
- Normes de pollution.
- Aérodynamique.
- Roulement et résistance au roulement (Pneumatiques plus ou moins usés, taille basse, rechapés, pression des pneus, ...)
- Electronique moteur (Régulateur, pompe à injection, ...).

- Equipements complémentaires (Climatisation et chauffage notamment).

Les paramètres (variables) relatifs aux conditions opérationnelles d'exploitation.

- Charge.
- Entretien.
- Adaptation du véhicule à l'utilisation.
- Conditions de route et profils de parcours.
 - Il s'agit du type de routes (autoroute, milieu urbain ou inter-urbain) et d'infrastructures (notamment revêtements). Mais aussi des déclivités, tracés et contraintes, conditions de trafic, trajets de nuit, ...
- Conditions climatiques.
- Contraintes de conduite et vitesse commerciale.

Il est intéressant de noter que l'incidence de ces paramètres "sur-déterminants" peut représenter jusqu'à + ou - 20% des valeurs de consommation lorsque sont additionnés plusieurs facteurs de surconsommation (ou aussi + ou - 10 litres/100km).

La mise en œuvre d'une conduite éco-responsable repose sur le suivi de certaines règles, relatives soit au comportement de conduite (paramètres variables) soit aux conditions opérationnelles (paramètres fixes). Au fil de mes recherches, j'ai mené une étude approfondie de la littérature professionnelle, et rencontré plusieurs enseignants en écoconduite. Je me suis bien entendu moi-même formé à ce type de conduite, afin d'être en mesure de pouvoir coordonner une série d'expérimentations.

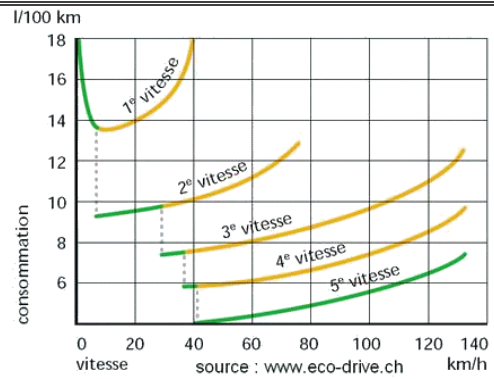
L'écoconduite est enseignée de manière très variable en Europe, mais il est possible d'extraire une série de comportements recommandés et unanimement considérés comme réducteurs de consommation. J'ai appelé ces comportements les règles d'or de l'écoconduite.

Tableau 14 : règles d'or de l'éco-conduite.

N°	Règles d'or de l'écoconduite
1	Engagez le rapport supérieur le plus rapidement possible

N° Règles d'or de l'écoconduite

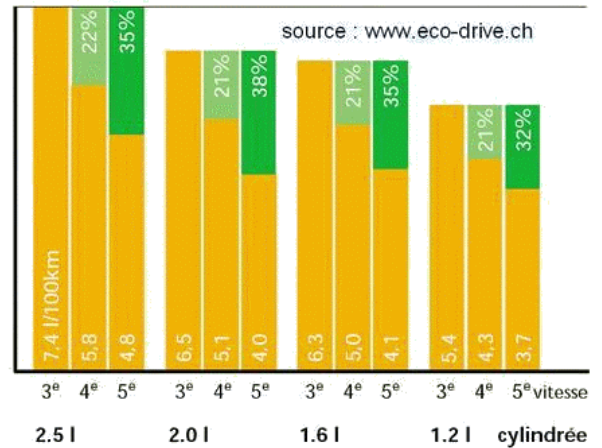
Passer à la vitesse supérieure entre 1.500 et 2.000 t/min



Courbes de consommation en fonction de la vitesse et du rapport de boîte

2 Maintenez une vitesse stable en utilisant le rapport le plus élevé possible

Enclencher la plus haute vitesse possible et conduire avec un régime moteur faible, en cherchant à stabiliser sa vitesse



Consommation moyenne en fonction du rapport de boîte pour différentes cylindrées.

3 Anticiper le trafic afin d'éviter d'accélérer ou de freiner inutilement

Regarder le plus loin possible (un terrain de football) et anticiper le trafic environnant

4 Décélérez doucement en relâchant l'accélérateur à temps tout en laissant une vitesse engagée

S'il faut ralentir ou s'arrêter, décélérer progressivement en relâchant l'accélérateur à temps et en laissant la voiture en prise (utilisation du frein moteur)

Néanmoins, les règles de l'écoconduite peuvent être raffinées aux paramètres fixes de la consommation, et on retrouve régulièrement les conseils suivants.

Tableau 15 : Autres conseils pour économiser le carburant.

N°	Autres conseils pour économiser le carburant																				
1	<p data-bbox="331 304 919 338">Adopter une vitesse adaptée à l'infrastructure</p> <p data-bbox="331 383 1455 510">-En descente, quand le moteur n'est pas sollicité, et qu'une vitesse est enclenchée, les véhicules avec un moteur à injection ne consomment pas d'essence car l'arrivée du carburant est coupée.</p> <p data-bbox="331 524 1455 651">-En montée, sur autoroute avec une pente de 4%, rouler à 120 km/h peut représenter une consommation de 19 litres de carburant au 100 km contre moins de 15 litres au 100 si le véhicule roule à 100 km/h.</p> <p data-bbox="331 665 1455 748">-Si la pente est de 6%, à 120 km/h, la consommation du véhicule peut atteindre 26 litres / 100 km contre 20 litres à la vitesse de 100km/h.</p> <p data-bbox="331 761 1455 844">-Idée reçue : il est inutile de passer au point mort dans les descentes ou ralentissements en espérant consommer moins d'essence.</p> <p data-bbox="331 857 1455 940">-Au point mort, le moteur consomme du carburant pour maintenir son ralenti alors que si vous utilisez le frein moteur, il ne consomme rien.</p>																				
2	<p data-bbox="331 983 812 1016">Entretien régulièrement son véhicule</p> <p data-bbox="331 1061 1455 1189">-Un véhicule mal entretenu peut entraîner une surconsommation de carburant pouvant aller jusqu'à 25%. Par exemple, un filtre à air encrassé fait consommer 3% de plus et un moteur mal entretenu, encore plus.</p>																				
3	<p data-bbox="331 1234 810 1267">Préchauffer son moteur ne sert à rien</p> <p data-bbox="331 1312 1455 1395">-Les moteurs modernes permettent de rouler juste après le démarrage et ainsi de réduire votre consommation de carburant.</p>																				
4	<p data-bbox="331 1435 895 1469">Vérifier régulièrement la pression des pneus.</p> <p data-bbox="331 1514 823 1641">-Une pression 25 % trop faible accroît la résistance au roulement de 10 % et la consommation de carburant, de 2 %.</p> <div data-bbox="922 1529 1378 1816"> <table border="1"> <caption>Données du graphique</caption> <thead> <tr> <th>Pression des pneus (bar)</th> <th>Consommation supplémentaire (%)</th> </tr> </thead> <tbody> <tr><td>2.5</td><td>0</td></tr> <tr><td>2.4</td><td>0.5</td></tr> <tr><td>2.3</td><td>1.0</td></tr> <tr><td>2.2</td><td>1.5</td></tr> <tr><td>2.1</td><td>2.0</td></tr> <tr><td>2.0</td><td>2.5</td></tr> <tr><td>1.9</td><td>3.0</td></tr> <tr><td>1.8</td><td>3.5</td></tr> <tr><td>1.7</td><td>4.0</td></tr> </tbody> </table> <p data-bbox="1129 1794 1334 1816">source : www.eco-drive.ch</p> </div> <p data-bbox="863 1861 1437 1944" style="text-align: right;"><i>Surconsommation du véhicule en fonction de la pression des pneus.</i></p>	Pression des pneus (bar)	Consommation supplémentaire (%)	2.5	0	2.4	0.5	2.3	1.0	2.2	1.5	2.1	2.0	2.0	2.5	1.9	3.0	1.8	3.5	1.7	4.0
Pression des pneus (bar)	Consommation supplémentaire (%)																				
2.5	0																				
2.4	0.5																				
2.3	1.0																				
2.2	1.5																				
2.1	2.0																				
2.0	2.5																				
1.9	3.0																				
1.8	3.5																				
1.7	4.0																				
5	<p data-bbox="331 1962 954 1995">Aérer la voiture avant de mettre la climatisation</p>																				

N° Autres conseils pour économiser le carburant

-Selon une étude de la Société des ingénieurs automobiles (SAE), dans une berline avec fenêtres baissées la consommation de carburant peut augmenter de 8% à 20%, la différence étant plus marquée au-delà des 80km/h.

-Cependant, la surconsommation moyenne avec la climatisation est de l'ordre de 20%, et selon les chiffres de l'ADEME, à 30° C avec un fort ensoleillement, la surconsommation peut atteindre 30% à 40% en ville, et 15% à 25% sur route-autoroute, si la température de consigne est de 26° C.

6 Ne pas laisser de charges inutiles à bord

-Le poids additionnel sur un véhicule entraîne un accroissement de la résistance au roulement du véhicule (friction entre la roue et la route plus importante). Il faut donc éviter de surcharger son véhicule inutilement

7 Utiliser les systèmes d'aide à la conduite

-Un système de navigation vous aidera à optimiser votre trajet, tandis que le régulateur de vitesse réduit la consommation

-Le dispositif "start and go" permet d'éteindre automatiquement le moteur lors des arrêts.

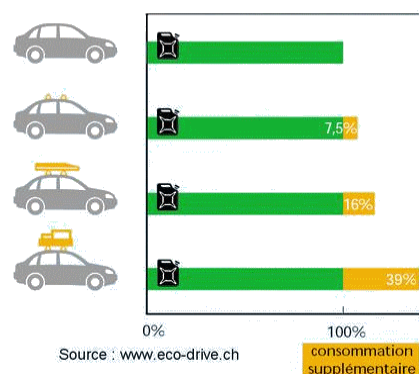
8 Eteignez le moteur de votre véhicule lorsqu'il est en stationnement ou dans une file d'attente

-Roulez dès que vous avez démarré votre moteur

-Et arrêtez-le pour tout stationnement supérieur à 30 secondes.

9 Enlever les éléments nuisant à l'aérodynamisme du véhicule

-Le graphique ci-après donne la consommation supplémentaire pour une vitesse de 120 km



Surconsommation du véhicule en fonction du chargement d'objets sur le toit.

Laissez votre véhicule au garage !

N°	Autres conseils pour économiser le carburant
10	<ul style="list-style-type: none"> -Évitez la voiture sur courts trajets -La marche à pied, le vélo ou les transports en commun sont souvent plus efficaces ou rapides en centre urbain que les courses en voiture. -Prendre tous les jours le vélo pour faire 10 km, plutôt que la voiture ne prend pas beaucoup plus de temps (et en gagne dans les grandes villes !) -Cela permet d'éviter de rejeter 700 kg de CO₂ par an dans l'atmosphère

Dans le contexte de mes recherches, je me suis d'abord intéressé à l'impact de l'infrastructure sur la consommation, pour ensuite chercher à évaluer l'influence du comportement sur les économies de carburant et étudier comment créer les conditions de l'émergence de ce comportement vertueux par le biais des assistances embarquées. Les stratégies majeures de réduction de la consommation de carburant des véhicules sont présentées à la Figure 40, l'écoconduite ayant le potentiel pour agir sur l'ensemble des facteurs.

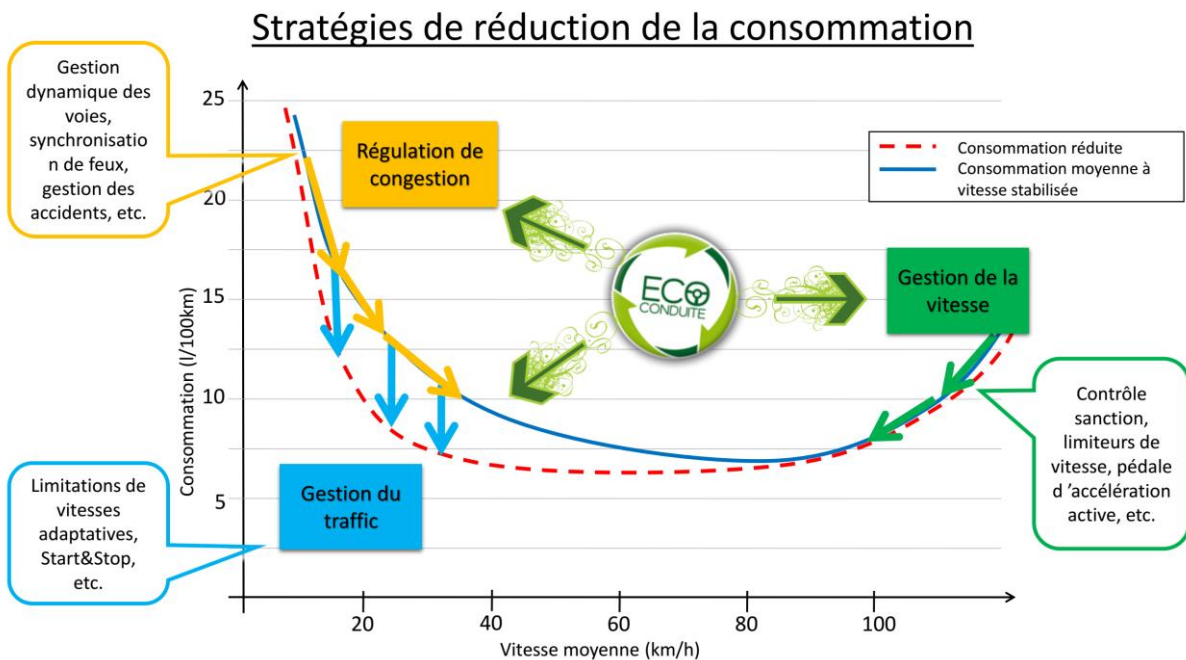


Figure 40 : Vue d'ensemble des principales stratégies de réduction de la consommation de carburant.

6.4.4. Lien entre vitesse et consommation

La première constatation concerne le lien entre la vitesse moyenne et la consommation. Lors de l'étude des données du projet LAVIA, j'ai pu observer que le lien était contraire à l'intuition : les trajets effectués avec une vitesse moyenne faible sont les plus consommateurs, tandis que ceux avec des vitesses moyennes élevées sont associés à des consommations réduites. Une illustration

de cette constatation est donnée à la Figure 41, et confirme l'état de l'art (cf. par exemple Rakha 2003). Bien que surprenant, cela s'explique aisément par le fait que la vitesse moyenne d'un trajet est la résultante d'un grand nombre de paramètres : la présence de congestion, les caractéristiques du moteur, le type d'infrastructure etc. Les moteurs thermiques ne sont en effet pas optimisés pour de faibles vitesses, mais plutôt pour un usage au-delà de 30 km/h, et la simple considération de la vitesse moyenne ne suffit pas à comprendre les mécanismes de l'économie d'énergie.

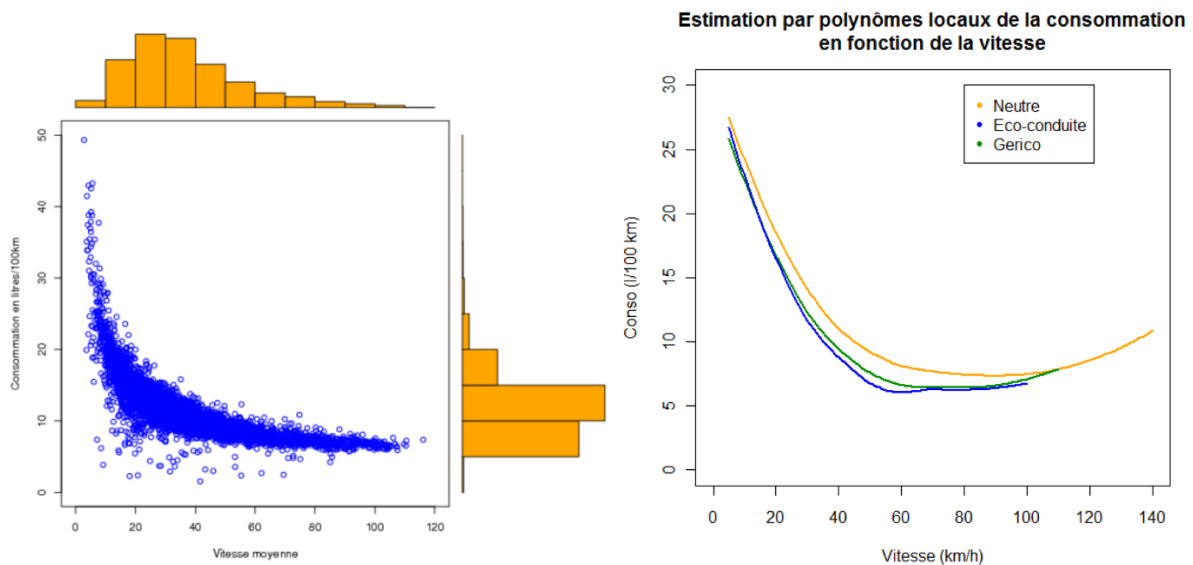


Figure 41 : Nuage des vitesses et consommations moyennes de 6465 trajets réels (projet LAVIA), et estimation par polynômes locaux de la relation entre vitesse et consommation moyenne selon le type de conduite (conduite neutre, après formation à l'écoconduite, et en utilisant une assistance embarquée d'aide à l'écoconduite, données issues du projet GERICO).

Il est donc clair qu'agir simplement sur la vitesse, comme le font les systèmes régulateurs de vitesse (cf. section 6.3, système LAVIA et limiteur et régulateur) ne suffit pas à obtenir des gains de consommation, et peuvent même produire des effets négatifs aux faibles vitesses (<40km/h).

En explorant cette question plus en détail à l'aide des données dont je disposais, j'ai pu mettre en évidence l'importance du choix du rapport de boîte de vitesse (cf. Figure 42). En effet, pour une même vitesse stabilisée, il est possible d'opter pour différents rapports de boîtes, le plus petit induisant une plus forte consommation de carburant. Il est donc primordial de sélectionner avec justesse le rapport de boîte le plus adapté à sa vitesse, une vitesse trop faible ou trop forte induisant une sur-consommation élevée, particulièrement en phase d'accélération (cf. Figure 42).

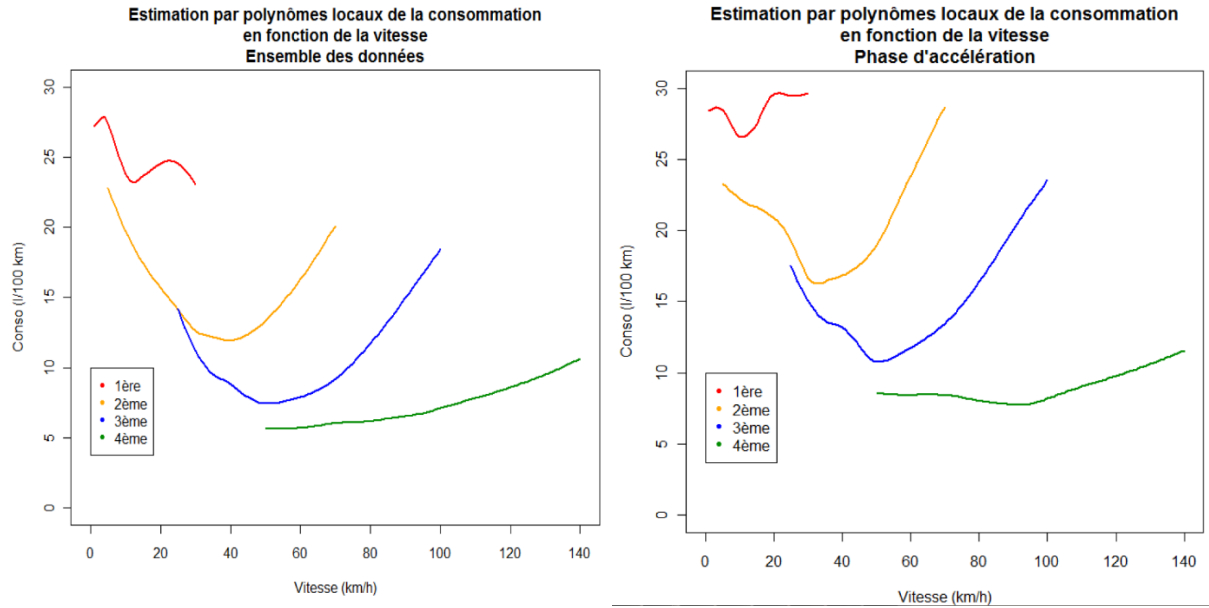


Figure 42 : Estimation par polynômes locaux du lien entre consommation et vitesse instantanée pour différents rapports de boîte (données GERICO), pour l'ensemble de la conduite (gauche) et pour les seules phases d'accélération (droite).

Lors du projet GERICO, 39 conducteurs ont effectué un trajet de 70 km/h en conditions normales, puis la moitié d'entre eux a suivi un cours d'écoconduite, tandis que l'autre moitié apprenait l'usage d'un système embarqué d'aide à l'écoconduite. Le second trajet effectué par chacun d'eux a permis de comparer les stratégies de changement de rapport de boîte mises en place ainsi que les régimes moteurs résultants. Un exemple de résultat en situation de traversée d'agglomération (cf. Figure 43) permet de constater le changement très marqué de stratégie de choix de rapport de boîte de vitesse, et ses conséquences sur le régime moteur, et donc la consommation de carburant.

Plus qu'une bonne gestion de la vitesse, une bonne gestion des changements de rapports de boîte est à privilégier pour réduire sa consommation et maintenir le moteur dans son régime d'efficacité maximale.

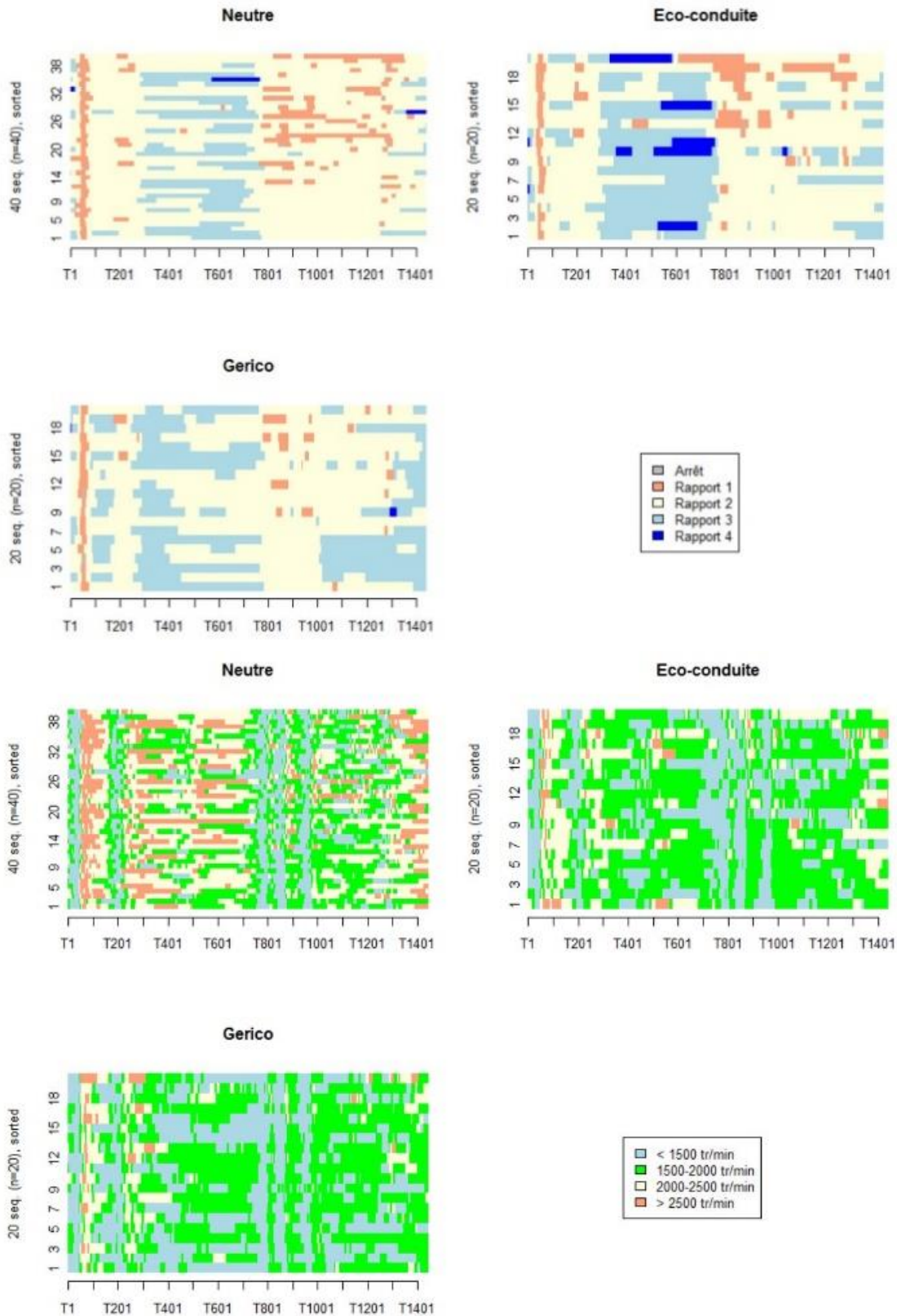


Figure 43 : Analyse descriptive des séquences de rapports de boîtes (colonnes de gauche) et de classes de régimes moteurs (colonnes de droite) pour 3 catégories de conducteurs sur une distance de 1500 mètres pendant la traversée d'une agglomération.

Cependant, la recherche effrénée d'une consommation minimale sans considération de vitesse instantanée peut conduire à un comportement certes économe mais avec un niveau de sécurité dégradé (cf. Chapnick, N. (2007)). L'écoconduite ne doit donc pas se pratiquer au détriment de la sécurité, et la vitesse adoptée doit être en harmonie avec les infrastructures rencontrées et les contraintes qui leurs sont associées.

Les expériences que j'ai menées ont permis de mettre en évidence certains comportements « à risque » associés à la pratique de l'écoconduite. Ainsi, la Figure 44 permet de constater une meilleure adéquation générale des vitesses adoptées par les écoconducteurs avec les limitations en vigueur, excepté pour la phase finale. Le dernier tronçon étant en forte pente, la vitesse plus élevée des écoconducteurs correspond à la volonté de préservation de l'énergie cinétique emmagasinée et donc à un souci d'économie d'énergie. Néanmoins, ce comportement induit un non-respect des règles de sécurité auquel il faut veiller lorsque l'on cherche à encourager l'écoconduite par le biais d'un système embarqué.

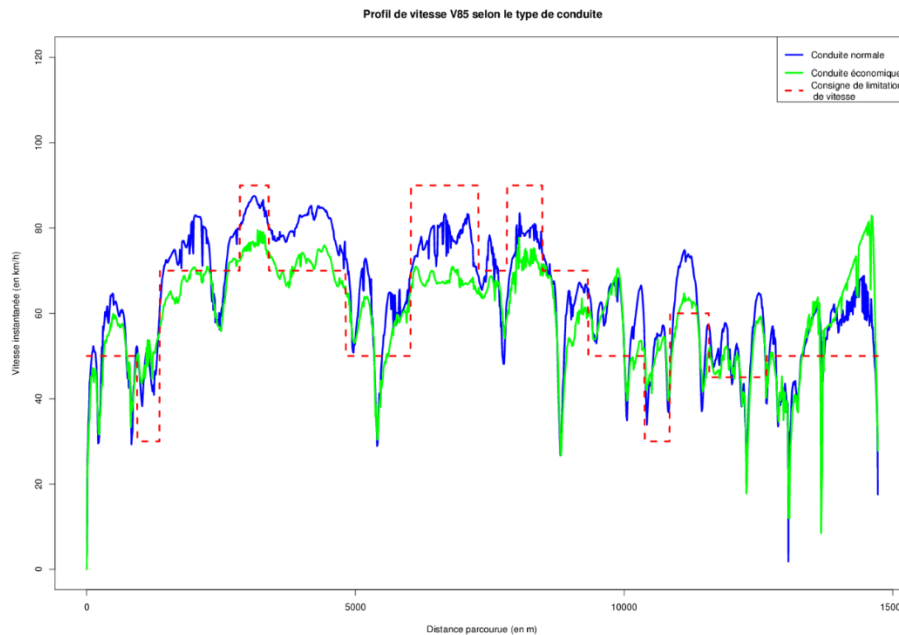


Figure 44 : Profils de vitesses moyens (30 conducteurs) selon le type de conduite, mis en rapport avec la limitation de vitesse en vigueur.

6.4.5. Lien entre infrastructure et consommation

Il m'a semblé intéressant d'approfondir l'étude de l'influence de l'infrastructure sur la consommation. En effet, le conducteur n'est pas seul responsable d'une consommation excessive, et l'écoconduite est parfois impraticable dans certaines situations. De même, certaines infrastructures nécessitent des adaptations spécifiques du comportement de conduite.

Dans le cadre d'un travail en collaboration avec le Cerema, j'ai cherché à quantifier statistiquement la sur-consommation induite par la présence d'une infrastructure particulière comme un rond-point, un feu rouge, ou une chicane.

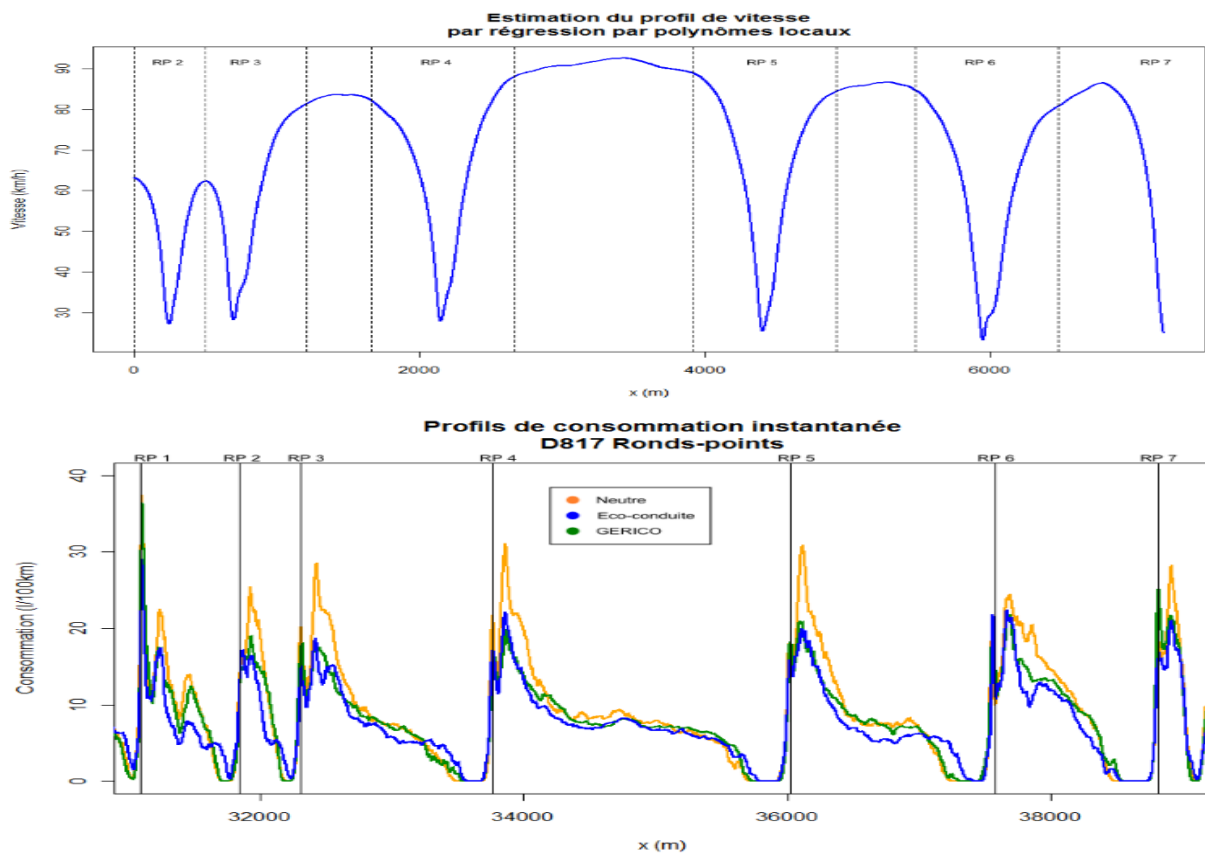
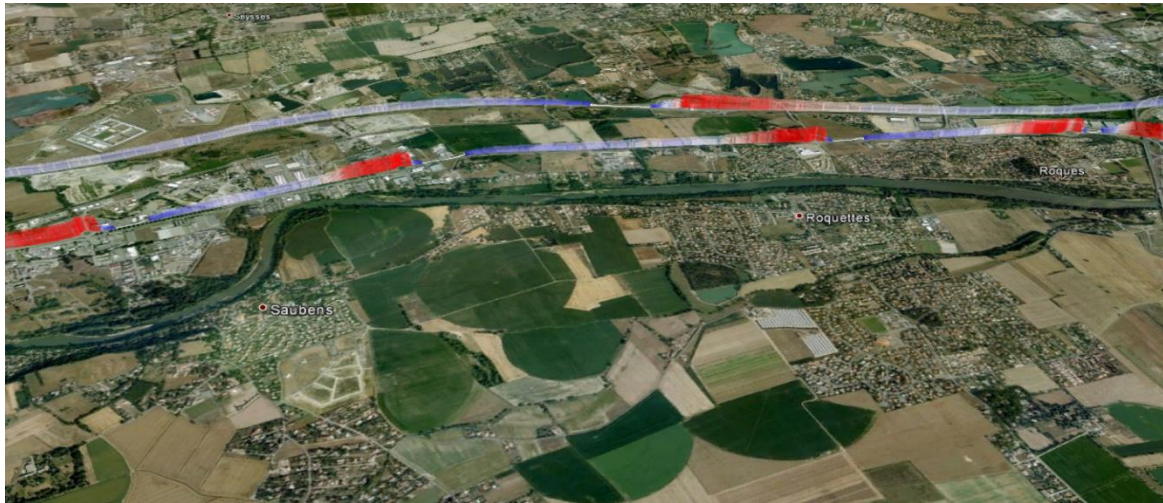


Figure 45 : Traversée de cinq ronds-points. Visualisation du profil de consommation moyen (haut), du profil de vitesse moyen associé (milieu), et du profil de consommation instantanée selon le type de conduite (bas).

La figure précédente (Figure 45) permet de constater la très grande influence d'une série de ronds-points sur le choix des vitesses et sur la consommation associée. Bien que très utiles pour

fluidifier le trafic, ces infrastructures entraînent aussi des émissions de polluants localisées à l'endroit des pics de consommation observés (phases d'accélération).

Le travail effectué a tiré profit de l'utilisation des modèles mixtes (GLMM, cf. section 6.2.6.3) pour quantifier le surcroît de consommation observé. La sur-consommation obtenue pour un rond-point est de 20ml de carburant par véhicule. Un tel modèle est en adéquation avec la consommation cumulée moyenne observée (cf. Figure 46).

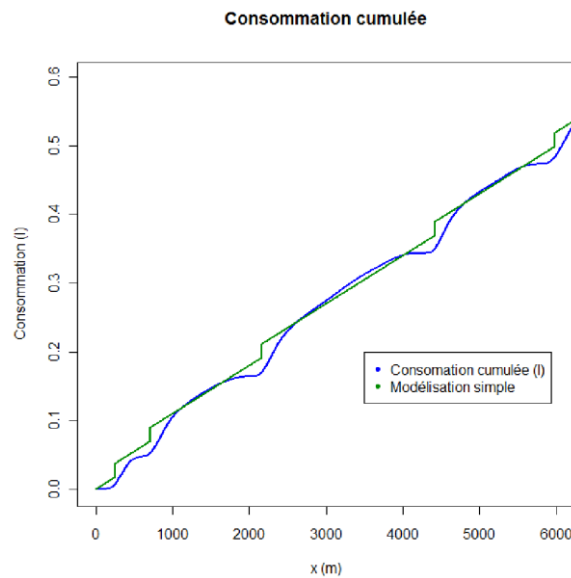


Figure 46 : Consommation cumulée lors de la traversée de 5 ronds-points, comparée au modèle linéaire simple obtenu.

Si l'écoconduite peut avoir un effet sur la gestion des phases d'accélération, son impact est trop contraint pour être significatif dans les phases de décélération contraintes comme une série de ronds-points. De même, un feu rouge induit un arrêt forcé, et donc une sur-consommation, que l'on soit écoconducteur ou non. Si le feu est encombré, plusieurs arrêts sont alors nécessaires et la surconsommation est encore accrue. Une approche similaire au cas des ronds-points a été menée pour les feux rouges (cf. Figure 47), en tenant compte du nombre d'arrêts. Lorsque le conducteur est arrêté une seule fois, il consomme 22ml de carburant de plus que lors du feu vert. Si il est arrêté deux fois, il consomme 36ml de plus, et 65ml de plus dans le cas de plus de deux arrêts.

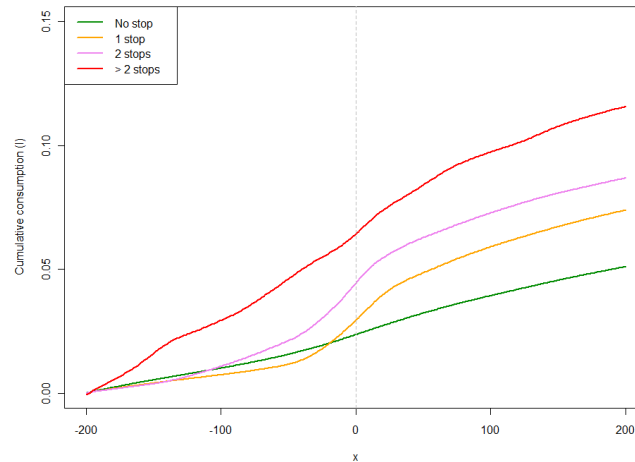


Figure 47 : Profils de consommation cumulée selon le nombre d'arrêts au feu tricolore.

6.4.6. Caractérisation et impact de l'écoconduite

Les liens complexes reliant les actions de conduite avec la consommation énergétique, ainsi que les contraintes d'infrastructure et de sécurité, m'ont conduit à rechercher comment évaluer au mieux les effets d'une conduite apaisée, qu'elle soit pratiquée de façon autonome, ou bien avec le support d'un système d'assistance embarquée. Ainsi, j'ai choisi de développer des indicateurs de performance spécifiques et innovants, associés aux différentes règles d'or de l'écoconduite (cf. Tableau 16). Ce procédé a l'avantage d'associer un type de comportement reconnu par les experts et aisément compréhensible avec un indicateur quantitatif aisément mesurable.

Tableau 16 : Règles d'or de l'écoconduite et indicateurs associés.

Règles d'or	Indicateur de performance	Nom de l'indicateur
Règle N°1 : Engagez le rapport supérieur le plus rapidement possible	Moyenne du régime moteur auquel s'effectue le changement de vitesse vers un rapport de boîte supérieur	AvgRPMSHiftUp
Règle N°2 : Maintenez une vitesse stable en utilisant le rapport le plus élevé possible	Indice de distribution des rapports de boîte et de régimes moteurs moyens associés à ces rapports	IndexGearRPM
Règle N° 3 : Anticiper le trafic afin d'éviter d'accélérer ou de freiner inutilement	Positive Kinetic Energy	PKE
Règle N° 4 : Décélérez doucement en relâchant l'accélérateur à temps	Pourcentage du temps passé avec le frein moteur	TimeEngineBrake

Règles d'or	Indicateur de performance	Nom de l'indicateur
tout en laissant une vitesse engagée		

La première règle incite à changer de rapport de boîte rapidement, il est donc naturel de caractériser ce comportement par la moyenne des régimes moteurs lors des changements de rapports. La seconde règle concerne autant le régime moteur que le rapport de boîte engagé, et il est donc naturel d'utiliser un indicateur rassemblant ces deux notions. J'ai utilisé une moyenne des régimes moteurs pondérés selon le temps passé dans chacun des rapports de boîte. Le second indicateur s'écrit donc de la façon suivante :

$$IndexGearRPM = \frac{1}{3500} (TimeNeut \times AvgRPMNeut + TimeGear1 \times AvgRPMGear1 + \dots + TimeGear5 \times AvgRPMGear5) \quad (1.18.)$$

Avec $TimeNeut$ (resp. $TimeGear1$, ..., $TimeGear5$) le pourcentage de temps au point mort, et $AvgRPMNeut$ (resp. $AvgRPMGear1$, ..., $AvgRPMGear5$) la moyenne des régimes moteurs enregistrés au point mort (respectivement $Gear1$, ..., $Gear5$). Le tout est normalisé par une valeur arbitraire représentant la valeur maximale du régime moteur (variable selon les véhicules). Pour la règle n°3, l'indicateur est issu de la littérature et représente une notion de d'énergie cinétique du profil de vitesse. Il s'écrit de la façon suivante :

$$PKE = \frac{\sum (v_f^2 - v_i^2)}{x} \text{ when } \frac{dv}{dt} > 0 \quad (1.19.)$$

Avec v_f et v_i respectivement les vitesses finales et initiales (en m/s) lors d'une phase croissance ($dv/dt > 0$) et x la distance parcourue totale en mètres. Un conducteur nerveux avec de fortes variations de vitesse aura un PKE élevé, tandis qu'un conducteur ayant une vitesse stable aura un PKE proche de 0. Enfin, la règle n°4 est tout simplement associée au niveau d'utilisation de frein moteur.

6.4.6.1. Mise en œuvre de l'écoconduite

Afin de permettre l'évaluation de l'impact environnemental des systèmes embarqués d'assistance à la conduite, j'ai tout d'abord cherché à évaluer les gains potentiels lors de la mise en œuvre de l'écoconduite en conditions optimales, c'est-à-dire juste après une sensibilisation spécifique, ou immédiatement après une formation adaptée. J'ai donc mené une série d'expérimentations en conditions réelles, consistant à demander à des conducteurs de conduire deux fois le même trajet avec un véhicule instrumenté, le second trajet étant effectué après une sensibilisation (présentation des règles d'or de l'écoconduite) ou une formation prodiguée par des professionnels. Une synthèse des résultats de ces deux expérimentations est présentée au Tableau 17.

Tableau 17 : Evolution des indicateurs de performance selon le style de conduite, pour chacune des expérimentations effectuées, et test de significativité (T-test).

Paramètre	Description	Expérience 1 (sensibilisation)			Expérience 2 (formation)		
		Moyenne "Normal"	Moyenne "Eco"	Variation (%)	Moyenne "Normal"	Moyenne "Eco"	Variation (%)
AvgFuelConsum	Consommation de carburant moyenne (l/100km).	6.86	6.00	-12.5***	9.01	7.99	-11.3***
AvgRPMShiftUp	Moyenne du régime moteur auquel s'effectue le changement de vitesse vers un rapport de boîte supérieur (règle 1).	2737.5	2232.8	-18.4***	3177.3	2465.6	-22.4***
IndexGearRPM	Indice de distribution des rapports de boîte et de régimes moteurs moyens associés à ces rapports (règle 2).	61.0	52.9	-13.3***	70.8	60	-15.3***
PKE	Positive Kinetic Energy (règle 3).	0.343	0.243	-29.2***	0.293	0.197	-32.8***
TimeEngineBrake	Pourcentage du temps passé avec le frein moteur (règle 4).	20.3	26.3	+29.6**	16.2	16.8	+ 0.04
AvgSpeed	Vitesse moyenne (km/h)	50.85	47.89	-5.8**	61.45	55.22	-10.1***
AvgAccel	Accélération moyenne (ms ⁻²)	0.498	0.387	-22.3***	0.596	0.473	-20.6***
AvgDecel	Décélération moyenne (ms ⁻²)	-0.619	-0.523	-15.5***	-0.672	-0.599	-10.9***
AvgRPM	Régime moteur moyen (rpm)	2097.4	1835.5	-12.5***	2379.6	2009.6	-15.5***
TimeNonLegalSpeed	Pourcentage de temps en excès de vitesse	37.9	26.5	-30.1***	28.5	18.2	-36.1***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

La première constatation consiste à remarquer qu'une réduction de consommation de carburant de plus de 10% est observée dans les deux cas, ce qui incite à penser qu'une conduite économe est possible sans une longue formation. Il apparaît clairement que la conduite économe peut produire un effet positif sur la sécurité (diminution de la vitesse moyenne, et du temps en excès de vitesse) et contribuer à une fluidification du trafic (accélérations et décélérations moins marquées). Ces résultats descriptifs sont en accord avec de nombreuses autres études autour de l'écoconduite.

Une question essentielle consiste à comprendre comment créer les conditions de l'émergence d'un tel comportement de conduite en conditions normales de conduite (un expérimentateur accompagnait le conducteur lors des deux expériences précédentes).

J'ai donc cherché à caractériser les différences observées dans la pratique de l'écoconduite selon le niveau d'apprentissage (sensibilisation ou formation par des professionnels). En fondant mon analyse sur les 4 indicateurs associés aux règles d'or de l'écoconduite (lignes 2 à 5) il apparaît des effets similaires pour les deux expériences lors de l'application des règles 1 à 3, mais l'application de la règle 4 (usage du frein moteur) semble sujette à une différence de mise en œuvre. Une évaluation utilisant des critères plus rigoureux (Nagelkerke R^2 dans le cadre d'une régression logistique), et selon les limitations de vitesse, a été menée et publiée (Andrieu C., Saint Pierre G. (2014)). Ce travail a mené à la conclusion que certaines règles d'or sont plus difficiles à mettre en œuvre (utilisation du frein moteur), et que cela peut être variable selon le type d'infrastructure et le niveau de sensibilisation à l'écoconduite (sensibilisation ou formation).

Il m'a donc semblé pertinent de chercher à construire un indicateur agrégé rassemblant les différentes dimensions de la conduite écologique dans une unique valeur. Un tel indicateur peut servir par exemple, à générer un signal informant l'utilisateur de son « niveau » d'écoconduite actuel, pouvant être incorporé dans un système embarqué d'assistance à la conduite.

6.4.7. Conception et mise en œuvre d'un éco-index

Résumer une notion aussi complexe que l'écoconduite en un seul indicateur est en soit délicat, car cela peut contribuer à une trop grande simplification et produire des effets indésirables. Il est cependant envisageable de fonder un indicateur simple, une jauge de couleur variable par exemple, sur cet indicateur agrégé, et d'offrir la possibilité de désagréger l'information en permettant d'accéder aux différents sous-indicateurs.

L'approche que j'ai utilisée a consisté à exploiter les multiples expérimentations effectuées, pour paramétrer un modèle de prédiction de la situation de conduite économe. Le modèle utilisé est un modèle linéaire de type « régression logistique ». Cette approche permet de déterminer la probabilité qu'une section de trajet ait été effectuée en conditions d'écoconduite, en fonction des valeurs des 4 paramètres associés aux règles d'or de l'écoconduite. Cette probabilité peut ensuite

être convertie sous une forme aisément perceptible par le conducteur (couleurs, étoiles, note sur 100, etc.).

L'éco-index proposé peut être construit à partir d'une combinaison linéaire des 4 indicateurs principaux, obtenue par le biais d'une régression logistique. La probabilité qu'une section de trajet Y corresponde à une conduite économe peut donc être exprimée de la façon suivante :

$$F(Y) = \frac{e^Y}{1 + e^Y} \quad \text{avec } Y = G(PIS) \text{ une combinaison linéaire des 4 indicateurs principaux}$$

(PIs=Performance Indicators). Cette approche présentée dans Saint Pierre & Andrieu (2010) et Andrieu & Saint Pierre (2012), est pertinente pour quantifier le niveau d'écoconduite, mais n'est pas adaptée à la construction d'un feedback utilisateur pertinent. En effet, certains indicateurs associés aux règles d'or ne sont calculables qu'en disposant de certaines informations (régime moteur) et ne peuvent donc être utilisés pour une application nomade. De plus, une approche agrégeant les 4 indicateurs rend l'auto-apprentissage complexe pour l'utilisateur.

Dans le cadre du projet ecoDriver, l'enjeu consistait à développer une application nomade pour smartphones, avec un fort enjeu d'acceptabilité et d'ergonomie. L'approche initiale a donc été « éclatée » en scindant l'éco-index global initialement mis au point, en 4 sous éco-index associés à chacune des règles d'or. L'éco-index global devenant ainsi une simple moyenne des 4 sous-indicateurs. Chacun de ces sous-indicateurs restant, étant lui-même obtenu très simplement à l'aide d'une régression logistique simple :

$$F(Y) = \frac{e^Y}{1 + e^Y} \quad \text{avec } Y = a \times PI + b \text{ et PI un des 4 indicateurs principaux. Cela permet d'associer}$$

une note à chacune des règles d'or de l'éco-conduite, et aide le conducteur à comprendre de lui-même quel aspect doit être amélioré.

De même, afin d'augmenter l'acceptabilité du système et améliorer l'expérience utilisateur, il est intéressant de prévoir un mécanisme ludique (ludification ou gamification) incrémentant la difficulté en fonction des résultats obtenus par le conducteur. Dans notre cas, une approche à 3 niveaux de difficulté a été choisie. Un conducteur qui obtient un score supérieur à 70% sur plus de la moitié de ses 10 derniers trajets passera au niveau suivant, tandis qu'un conducteur ayant un score inférieur à 40% pour au moins 8 de ses 10 derniers trajets verra son niveau diminuer.

Un conducteur de niveau élevé devra atteindre de meilleures performances pour conserver un bon score. Par exemple, au niveau 1, le conducteur aura un score de 100 dès que son indicateur *timeenginebrake* atteindra 80%, tandis qu'il faudra atteindre 95% pour obtenir ce même score au niveau 3.

La mise en œuvre de cette approche par niveaux m'a conduit à « caler » les différents modèles logistiques sur les différentes valeurs observées pour les quantiles des différents indicateurs (cf. Tableau 18).

Tableau 18 : Percentiles des valeurs observées pour les 4 indicateurs associés aux règles d'or de l'écoconduite (conduite normale uniquement).

		Percentiles					
		Q25	Q40	Q50	Q90	Q95	Q97.5
Performance Indicators (PI)	pke	0.05	0.1	0.16	1.1	1.5	1.9
	indexgearrpm	20.60	30.0	40.80	60.5	64.4	69.8
	timeenginebrake	10.00	20.0	30.00	80.0	90.0	95.0
	avgrpmshiftup	1387.00	1500.0	1647.00	2214.0	2393.0	2544.0

La courbe logistique étant une courbe asymptotique, il est nécessaire de fixer $\epsilon=0.01$ la précision désirée. En notant $Max_{eco}=1-\epsilon$ et $Min_{eco}=0+\epsilon$, on peut imposer la condition $F(Y) = Max_{eco}$ lorsque l'indicateur associé atteint la valeur maximale désirée (Q50 au niveau 1, Q40 au niveau 2, Q25 au niveau 3), et $F(Y) = Min_{eco}$ lorsque l'indicateur atteint la valeur minimale (Q97.5 au niveau 1, Q95 au niveau 2, Q90 au niveau 3). Ainsi, pour le niveau 1, on détermine les constantes a et b en résolvant les équations suivantes :

$$\begin{aligned}
 a \times Q50 + b &= \text{Log} \left(\frac{Max_{eco}}{Min_{eco}} \right) \\
 a \times Q97.5 + b &= \text{Log} \left(\frac{Min_{eco}}{Max_{eco}} \right)
 \end{aligned}
 \tag{1.20.}$$

On comprendra aisément que le processus est inversé pour la règle n°4 dans la mesure où un indicateur élevé correspond à une meilleure performance d'écoconduite, contrairement aux règles 1 à 3. La Figure 48 présente graphiquement le lien entre les valeurs observées des indicateurs et le score selon le niveau.

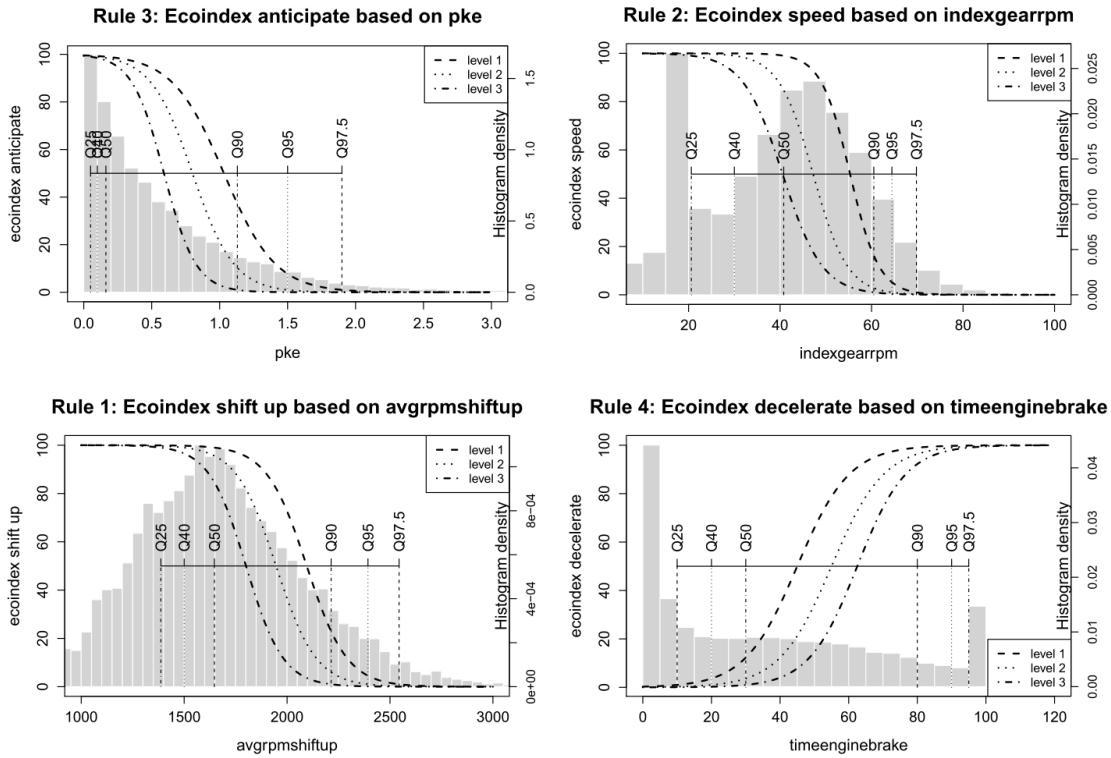


Figure 48 : Liens entre valeurs observées des indicateurs associés aux règles d'or et l'éco-index associé, selon le niveau du conducteur.

L'implémentation concrète de ce mécanisme a conduit à la présentation d'un score indiqué avec des étoiles (de 1 à 5) pour chacune des 4 règles d'or de l'écoconduite, desquels sont déduit le score global (moyenne des sous-scores). Un exemple d'évolution du score global d'écoconduite selon chacun des 3 niveaux est fourni à la Figure 49. L'ensemble de ces travaux ont été publiés dans Orfila, Saint Pierre et Messias (2015).

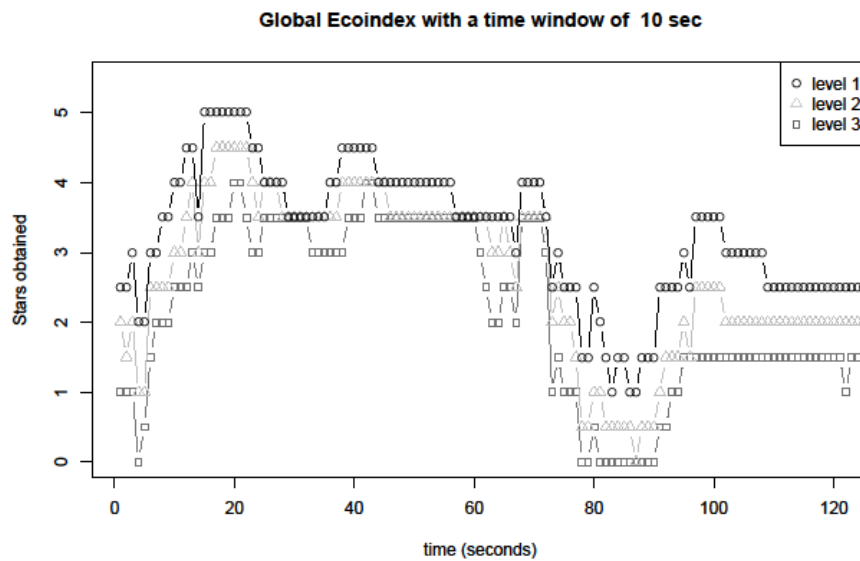


Figure 49 : Ecoindex global selon le niveau et pour chaque seconde d'un trajet réel.

6.4.8. Impact d'un système d'aide à l'écoconduite : ecoDriver

Le projet européen ecoDriver (<http://www.ecodriver-project.eu/>) a été très important pour moi à plusieurs titres. J'ai été responsable de l'ensemble du sous-projet consacré à l'évaluation des impacts (sub-project leader), et j'ai pu concrétiser plusieurs années de recherche en ayant la charge, avec mes collègues de l'Ifsttar, de la création de l'application pour smartphones Android. Cette application, entièrement conçue et développée intégralement par l'Ifsttar, a été construite en intégrant plusieurs recherches menées par notre équipe, notamment les éco-index détaillées à la section précédente. Cette application inclut par ailleurs un modèle d'estimation de la vitesse optimale et un indicateur de changement de rapport de boîte de vitesse (Orfila & Saint Pierre & Andrieu, 2012).

6.4.8.1. Conception d'une application Android d'assistance à l'écoconduite

Outre les différents éco-index destinés à proposer au conducteur un retour instantané sur sa capacité d'écoconduite, cette application tirait profit de notre expérience de l'évaluation de la conduite pour présenter différentes alertes au conducteur en temps réel, selon les paramètres de sa conduite. L'application comporte 4 types d'alertes « négatives » (signalement de comportement à risque)

- Décélération forte : Alerte déclenchée si, sur une fenêtre de temps fixée (5 secs par défaut), la décélération est inférieure à un seuil prédéfini (variable selon la vitesse) pendant au moins 0.5 secondes.
- Accélération forte : Alerte déclenchée si, sur une fenêtre de temps fixée (5 secs par défaut), l'accélération est supérieure à un seuil prédéfini (variable selon la vitesse) pendant au moins 0.5 secondes.
- Excès de vitesse : Alerte déclenchée si le conducteur est en excès de vitesse pendant 10 secondes sans interruption.
- Changement de rapport de boîte supérieur trop lent : Alerte déclenchée si le changement de rapport est effectué 10 secondes après la recommandation du système.

L'application comporte aussi 4 types d'évènements positifs.

- Bon usage du frein moteur : Message déclenché si l'usage du frein moteur dépasse 5 secondes sur une fenêtre glissante de 10 secondes.
- Respect des limitations : Message envoyé si le pourcentage de temps en excès de vitesse est inférieur à un seuil (5%) sur une période de 10 minutes.

- Bon changement de rapport supérieur (resp. inférieur) : Message déclenché si le conducteur effectue le bon changement de rapport 2 secondes avant ou après la recommandation du système.

Grâce à un modèle permettant de calculer la vitesse instantanée optimale tenant compte de l'horizon électronique (infrastructure routière en amont selon la cartographie), l'application comporte aussi des annonces « préventives » destinées à améliorer la sécurité de conduite :

- Approche de carrefour (avec indication du type de carrefour)
- Approche de sommet de côte / de descente
- Approche de côte
- Approche de limitation de vitesse.

Chaque événement ou alerte est signalée en fonction de la distance le séparant de l'événement. Un système de priorité a été mis en place afin de ne présenter qu'une seule alerte ou événement à la fois au conducteur. L'interface graphique a été mise au point par le partenaire Espagnol (CTAG) et est présentée à la Figure 50.

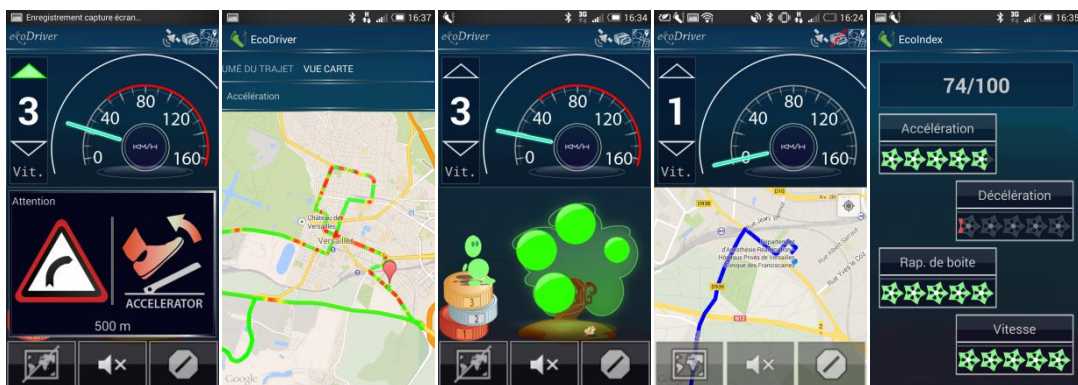


Figure 50 : Interface de l'application android ecoDriver mise au point par le CTAG.

Il est important de signaler que le projet ecoDriver a permis de construire différents types d'assistances embarquées :

- 7 systèmes conçus par des partenaires industriels (FIAT, BMW, Daimler, TomTom).
- Un système complet embarqué et connecté au véhicule.
- Une application pour smartphones Android, connectée ou non au véhicule.

Bien que j'ai supervisé l'évaluation de l'ensemble de ces systèmes dans le cadre du projet ecoDriver, je n'évoquerai dans ce manuscrit que la seule application pour smartphones Android. Le lecteur intéressé pourra se référer au livrable public D6.4 disponible sur le site du projet (<http://www.ecodriver-project.eu/>).

6.4.8.2. Impact du système

Le système a été évalué selon la méthodologie détaillée à la section 6.2, en utilisant deux types d'approche : Une approche classique en conditions contrôlées (circuit et véhicule identiques), et une approche utilisant la conduite en situation naturelle (trajets variables et véhicules personnels). J'ai assumé personnellement la coordination internationale de ces travaux impliquant une dizaine de partenaires européens autour de l'évaluation des sept systèmes conçus.

Un grand nombre d'hypothèses de recherche ont été testées, chacune étant associée à un aspect spécifique de la conduite : gestion de l'énergie, vitesses pratiquées, vitesse et interdistances dans certaines situations, règles d'or de l'écoconduite, et gestion de l'accélération dans certaines situations. Le tableau suivant présente un résumé des résultats obtenus.

Tableau 19 : Principaux résultats du projet ecoDriver concernant l'impact de l'application nomade Android, en conditions contrôlées et en conditions naturelles.

Effect sizes in percentages (differences in % from relevant baseline)				
Hyp. Number & cat.	PI abbreviated	Road type	Controlled Android App	Naturalistic Android App
1 Energy	% of reduction for fuel consumption & CO2	Urban	1.54 (N.S.)	-1.57 (N.S.)
		Rural	3.15 (N.S.)	-2.49 (N.S.)
		Motorway	-	0.3 (N.S.)
		All road types	2.54	-0.8 (N.S.)
3 Energy	% of NoX reduction compared to resp. baseline	Urban	-0.28 (N.S.)	-1.07 (N.S.)
		Rural	2.35 (N.S.)	-0.9 (N.S.)
		Motorway	-	3.44
		All road types	1.34 (N.S.)	0.97 (N.S.)
9 Speed	Average speed when cruising	Urban	-8.86 (N.S.)	3.28
		Rural	2.17 (N.S.)	0.03 (N.S.)
		Motorway	-	1.25
		All road types	-	1.24
10 Speed	Average speed when freely driving	Urban	0.45 (N.S.)	-
		Rural	1.31 (N.S.)	-
		Motorway	-	-
		All road types	1.18 (N.S.)	-
11 Speed Situations	avg_speed_distance_based before intersections	Urban	-1.4 (N.S.)	-
		Rural	1.78	-
		Motorway	-	-
		All road types	-0.61 (N.S.)	-
12 Speed Situations	avg_speed_distance_based before zebra crossings	Urban	0.49 (N.S.)	-
		Rural	3.18	-
		Motorway	-	-
		All road types	0.59 (N.S.)	-
13 Speed Situations	avg_speed_distance_based before speedbumps	Urban	0.6 (N.S.)	-
		Rural	-0.12 (N.S.)	-
		Motorway	-	-
		All road types	-0.12 (N.S.)	-
14 Speed Situations	avg_speed_distance_based before sharp curves	Urban	3.26 (N.S.)	-
		Rural	1.35 (N.S.)	-
		Motorway	-	-
		All road types	-0.79 (N.S.)	-

15 Speed Situations	avg_speed_distance_based at crests	Urban	2.25 (N.S.)	-
		Rural	2.16	-
		Motorway	-	-
		All road types	2.21	-
16 Speed Situations	avg_speed_distance_based before speed limit changes	Urban	3.08 (N.S.)	-
		Rural	0.74 (N.S.)	-
		Motorway	-	-
		All road types	1.45 (N.S.)	-
17 THW Situations	Average time headway	Urban	3.97 (N.S.)	-
		Rural	-1.88 (N.S.)	-
		Motorway	-	-
		All road types	-0.33 (N.S.)	-
18 THW Situations	Average time headway before intersections	Urban	3.67 (N.S.)	-
		Rural	-7.63 (N.S.)	-
		Motorway	-	-
		All road types	-1 (N.S.)	-
19 THW Situations	Average time headway before zebra crossings	Urban	-2.11 (N.S.)	-
		Rural	4.45 (N.S.)	-
		Motorway	-	-
		All road types	-1.74 (N.S.)	-
20 THW Situations	Average time headway before speed bumps	Urban	6.47 (N.S.)	-
		Rural	8.06 (N.S.)	-
		Motorway	-	-
		All road types	6.33 (N.S.)	-
21 THW Situations	Average time headway before sharp curves	Urban	0.56 (N.S.)	-
		Rural	8.37 (N.S.)	-
		Motorway	-	-
		All road types	1.32 (N.S.)	-
23 THW Situations	Average time headway before speed limit changes	Urban	5.59 (N.S.)	-
		Rural	-0.36 (N.S.)	-
		Motorway	-	-
		All road types	1.33 (N.S.)	-
27 Golden rules	Average rpm when shifting gear up	Urban	7.34	3.85
		Rural	7.43	8.29
		Motorway	-	2.19
		All road types	8.03	2.97
28 Golden rules	weighted average engine rpm	Urban	7.70	7.13
		Rural	6.00	9.12
		Motorway	-	2.24 (N.S.)
		All road types	7.03	5.00
29 Golden rules	Positive kinetic energy	Urban	1.45 (N.S.)	1.56 (N.S.)
		Rural	0 (N.S.)	0 (N.S.)
		Motorway	-	0 (N.S.)
		All road types	0 (N.S.)	1.69
30 Golden rules	Percentage of driving time with engine brake	Urban	1.96 (N.S.)	-0.71 (N.S.)
		Rural	6.38	3.73 (N.S.)
		Motorway	-	-4.73
		All road types	4.90	-0.54 (N.S.)
31 Accel Decel	95th percentile positive acceleration	Urban	2.11 (N.S.)	4.77
		Rural	1.61 (N.S.)	3.06 (N.S.)
		Motorway	-	7.44 (N.S.)
		All road types	1.12 (N.S.)	4.57
32 Accel Decel	5th percentile negative acceleration	Urban	0.65 (N.S.)	3.88
		Rural	-1.54 (N.S.)	3.28 (N.S.)
		Motorway	-	7.38
		All road types	-1.05 (N.S.)	4.31
		Urban	1.77	-

33 Accel Decel	maximum acceleration after stationary	Rural	-	-
		Motorway	-	-
		All road types	-	-
34 Accel Decel Situation	95th percentile of the negative acceleration before intersections	Urban	-0.09 (N.S.)	-
		Rural	-1.42 (N.S.)	-
		Motorway	-	-
		All road types	-	-
35 Accel Decel Situation	95th percentile of the negative acceleration before zebra crossings	Urban	0.76 (N.S.)	-
		Rural	-7.55 (N.S.)	-
		Motorway	-	-
		All road types	0.56 (N.S.)	-
36 Accel Decel Situation	95th percentile of the negative acceleration before speed bumps	Urban	4.49	-
		Rural	11.98	-
		Motorway	-	-
		All road types	4.91	-
37 Accel Decel Situation	95th percentile of the negative acceleration before sharp curves	Urban	-0.7 (N.S.)	-
		Rural	0.78 (N.S.)	-
		Motorway	-	-
		All road types	0.18 (N.S.)	-
38 Accel Decel Situation	95th percentile of the negative acceleration at crests	Urban	-	-
		Rural	-1.57 (N.S.)	-
		Motorway	-	-
		All road types	-	-
39 Accel Decel Situation	95th percentile of the negative acceleration before speed limit changes	Urban	-2.01 (N.S.)	-
		Rural	0.21 (N.S.)	-
		Motorway	-	-
		All road types	-0.4 (N.S.)	-

Ainsi, la réduction de consommation de carburant associée à l'usage d'une application nomade d'assistance à l'écoconduite est non significative en conditions naturelles, alors qu'elle est de 2.5% en conditions contrôlées (présence d'un expérimentateur). Cela confirme la difficulté à maintenir le comportement vertueux du conducteur sur le long terme.

Néanmoins, l'effet se fait sentir sur certaines "externalités" de l'écoconduite :

- Les utilisateurs « naturels » adoptent des vitesses moyennes plus faibles et sont moins sujets à des accélérations ou décélérations fortes (hypothèses 9, 31, 32).
- L'ensemble des utilisateurs ont une meilleure gestion de leur régime moteur (hypothèses 27 et 28).
- L'usage du frein moteur est accru en conditions contrôlées, mais peu impacté lors d'un usage « naturel » (pas de message lié à la gestion du régime moteur dans l'application, cf. hypothèse 30).
- Les décélérations sont mieux anticipées à l'approche des passages piétons en milieu urbain (hypothèse 36).

Le lecteur est renvoyé à la lecture des rapports de recherche du projet ecoDriver pour des précisions sur l'analyse et l'interprétation de ces résultats, ainsi que pour les aspects techniques.

6.4.9. Publications associées à cette section

Articles dans des revues à comité de lecture répertoriées (ACL)

- Orfila O., **Saint Pierre G.**, Messias M., (2015), « *An Android Based Ecodriving Assistance System to Improve Safety and Efficiency of Internal Combustion Engine Passenger Cars* », Transportation research part C: Emerging technologies, Volume 58, Part D, Pages 772–782 (*JCR, Transport science, IF=2.818*).
- Andrieu C., **Saint Pierre G.** (2014), « *Evaluation of ecodriving performances and teaching method: comparing training and simple advices* », European Journal of Transport and Infrastructure Research, 14(3), pp. 201-213, ISSN: 1567-7141 (*JCR, Transportation, IF=0.818*).

Communications avec actes dans congrès international (ACTI)

- Freitas-Salgueiredo C., Orfila O., **Saint Pierre G.**, Glaser S., Coiret A., Vandanjon P.O. (2014), « *Overview of solutions for reducing CO2 emissions during the use phase of passenger cars* », Transport Research Arena 2014, Paris, La Défense.
- Orfila O., **Saint Pierre G.**, Messias M. (2014), « *Development of an ecodriving assistance application for nomadic devices performing real-time and post trip coaching for road vehicles* », Transport Research Arena 2014, Paris, La Défense.
- Romon S., **Saint Pierre G.**, Khoudour L. (2014), « *Fuel consumption and speed profiles comparisons according to infrastructure using probe vehicle data* », Poster au Transport Research Arena 2014, Paris, La Défense.
- Orfila O., **Saint Pierre G.**, Andrieu C. (2012), « *Gear Shifting Behavior Model for Ecodriving Simulations Based on Experimental Data* », 15th meeting of the Euro Working Group on Transportation, EWGT 2012, 10-13 septembre, Paris, France.
- Andrieu C., **Saint Pierre G.** (2012), « *Comparing effects of ecodriving training and simple advices on driving behavior* », 15th meeting of the Euro Working Group on Transportation, EWGT2012, 10-13 septembre, Paris, France.
- Andrieu C., **Saint Pierre G.** (2012), « *Using statistical models to characterize eco-driving style with an aggregated indicator* », Intelligent Vehicles Symposium 2012, June 3-7, 2012, Alcalá de Henares, Spain, DOI: 10.1109/IVS.2012.6232197.
- Rakotonirainy A., Haworth N., **Saint Pierre G.**, Delhomme P. (2011), « *Research issues in Eco-driving* », 22nd International Technical Conference on the Enhanced Safety of Vehicles, Washington DC, June.
- **Saint Pierre G.**, Andrieu C. (2010). « *Caractérisation de l'éco-conduite et construction d'un indicateur dynamique pour véhicules thermiques* ». PRAC 2010.

Communications invitées sans actes (INV)

- **Saint Pierre G.** (2015), « *Quantifying ecodriving performances: efficient feedback for efficient driving* », exposé invité au workshop « Human factors of green driving », TRB 94th Annual Meeting, Washington, 11 janvier 2015.

- **Saint Pierre G.** (2012), « *Design and development of mobile service for ecodriving* », conférencier invité au Workshop DECOMOBIL "Nomadic transport services for multimodal mobility", organisé par le réseau Européen HUMANIST, (<http://www.youtube.com/watch?v=-xwM56ndyNU>), décembre 2012.

Rapports de recherche y compris rapports de contrat de recherche (RAPP)

- **Saint Pierre G.** et al. (2016), « Eco driving in the real-world: behavioural, environmental and safety impacts », deliverable 43.1 of the ecoDriver project, EU FP7.
- H. Tattegrain, D. Letisserrand, **G. Saint Pierre**, C. Val, R. Krishnakumar, « Méthodes d'évaluation de dispositifs embarqués utilisant la conduite en situation naturelle », livrable 2 de la convention N°: 2200683288 (DSCR-Ifsttar), Étude «IDACA» : Evaluation de l'Impact des Dispositifs d'Assistance à la Conduite Automobile.
- Kircher K. et al., (2014), « *D41_1 Performance indicators and ecoDriver test design* », public deliverable, projet ecoDriver, FP7.
- Woldeab, Z. et al., (2014), « *D32.1: Individual Use Cases and Test Scenarios Definition* », public deliverable, projet ecoDriver, FP7.
- Nouveliere L. et al, (2014), « *D21.3: Powertrain Model Validation* », public deliverable, projet ecoDriver, FP7.
- Seewald P., Orfila O., **Saint Pierre G.**, (2014), « *D23.2: Simulation and analysis document for on-line vehicle algorithms* », public deliverable, projet ecoDriver, FP7.

6.5. Les Méga Données au service de la mobilité intelligente

6.5.1. Introduction

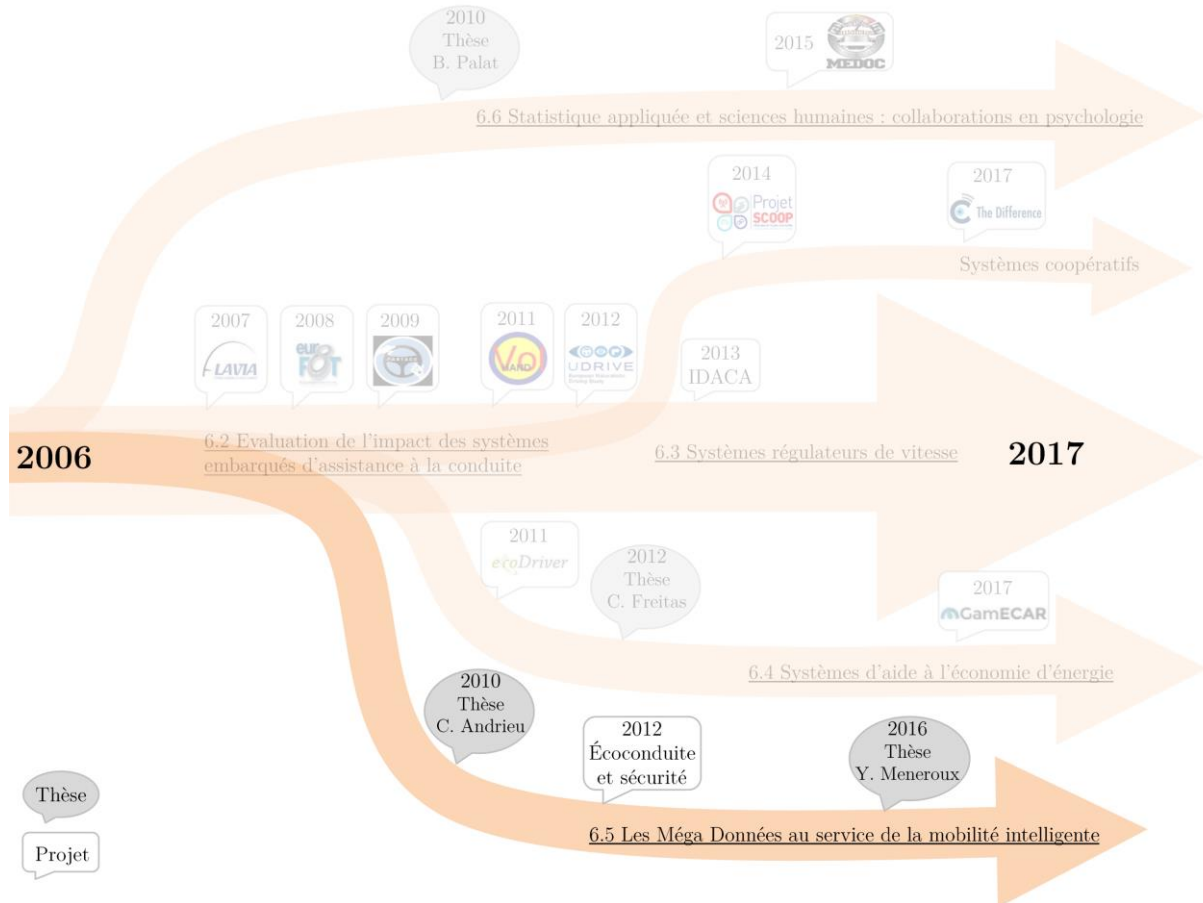


Figure 51 : Mise en évidence des axes de recherche développés dans cette section.

L'émergence d'une mobilité numérique et l'accès de plus en plus aisé aux données générées permet d'envisager de multiples applications dans le domaine de la gestion et de la sécurité des transports. Une simple trace GPS permet d'accéder à des informations fines sur le déplacement concernant à la fois le temps et l'espace, et par déduction la vitesse. Lorsqu'on les considère en grand nombre, ces traces numériques représentent une source d'information prodigieuse.

Les vitesses pratiquées sont une caractéristique essentielle du comportement des conducteurs et de leur usage du réseau routier. La connaissance de ces vitesses et leur évolution au fil du temps représente un enjeu de sécurité routière au cœur des préoccupations. Au-delà des enjeux de sécurité et de temps de parcours, problématiques jusqu'ici traitées par d'autres moyens, la collection de profils de vitesse (dédiés de traces GPS) peut permettre de produire des connaissances fines sur des phénomènes encore peu étudiés. Par exemple, il est déjà possible de fabriquer des cartes routières automatiquement à partir des traces GPS, comme il est possible de

détecter après coup de situations de conduite défaillante comme des freinages intempestifs répétés.

Mon expérience des traces numériques dans le contexte de l'évaluation des systèmes embarqués m'a conduit à imaginer des usages innovants pour ces informations. L'arrivée imminente des véhicules autonomes conduit à se poser la question de la vitesse de navigation adéquate selon le contexte routier (type d'infrastructure, présence de piétons, visibilité etc.). Il m'a semblé pertinent de considérer que la collection des vitesses pratiquées par les humains ayant préalablement emprunté un trajet précis, peut constituer un « a priori » pertinent pour le véhicule autonome en l'absence d'autre information. La construction d'un tel « a priori » a été l'objet d'une thèse que j'ai co-encadrée de 2010 à 2013 concernant la modélisation fonctionnelle des profils de vitesse.

De même, les cartographies numériques devenant un élément indispensable au déplacement motorisé individuel (déplacement conventionnel, mais aussi et surtout déplacement de véhicules autonomes), la quantité et la précision des éléments qu'elles contiennent devra augmenter. Il m'est donc apparu utile de chercher à exploiter des traces numériques pour créer, corriger, et/ou valider les informations contenues dans ces cartes. Le premier problème que j'ai souhaité aborder consiste donc à utiliser l'information de vitesse pour identifier les positions précises des éléments de signalisation verticale (feux rouges, stops, cédez le passage) ou horizontale (dos d'ânes, passages piétons). Faisant suite aux travaux théoriques engagés dans le cadre de la thèse citée précédemment, j'ai engagé en 2016 la co-direction d'une thèse d'un ingénieur de l'IGN autour de ces questions.

Ces idées, avancées dès 2010, commencent à se répandre dans le domaine industriel et opérationnel. Ainsi, le projet eHorizon de Continental, consiste à exploiter les données de leur flotte de véhicules connectés, pour construire une cartographie détaillée et dynamique de la route.

Cette section est consacrée à la présentation succincte de ces travaux menés dans le cadre de deux encadrements de thèse. La section 6.5.2 présente les principaux résultats obtenus par Cindie Andrieu (2010-2013), tandis que la section 6.5.3 s'intéresse aux travaux en cours menés par Yann Meneroux sous ma direction (2015-2018).

6.5.2. Modélisation des profils de vitesse

Les profils de vitesse spatialisés (vitesse en fonction de la distance parcourue, cf. Figure 52) sont des objets complexes, qui peuvent être vus comme des fonctions continues et non dérivables au point vitesse=0. La mise au point d'outils théoriques permettant leur manipulation en vue d'une extraction d'information, nécessite un travail méthodologique approfondi.

J'ai donc proposé en 2010 un sujet de thèse à l'Ifsttar qui a été financé, et dont la direction a été assurée par X. Bressaud, chercheur en probabilité et statistique à Toulouse III. J'ai eu le plaisir

de co-encadrer cette thèse avec une étudiante exceptionnelle, C. Andrieu, qui m'a accompagné lors d'une partie de mes recherches (écoconduite notamment). Pour ces travaux, j'ai souhaité employer une modélisation statistique fonctionnelle, afin de répondre à cette question concrète : comment, à partir d'une série de profils de vitesse, construire un profil moyen réalisable physiquement ?

En pratique, ces profils spatiaux de vitesse correspondent à une succession de mesures horodatées de vitesse et de position, et sont représentés par des vecteurs de \mathbf{R}^n . Cependant, le développement des technologies des capteurs permet de collecter des données avec des fréquences d'échantillonnage pouvant être très élevées, ce qui conduit à des vecteurs de grandes dimensions. Les méthodes usuelles de statistique multivariée ne sont alors plus adéquates (fléau de la dimension, corrélation entre les observations), et il devient nécessaire d'utiliser des méthodes plus appropriées au traitement des profils de vitesse.

Une solution consiste à représenter les profils spatiaux de vitesse par des fonctions et non par des vecteurs de \mathbf{R}^n . Cette approche est fondée sur l'Analyse des Données Fonctionnelles (en anglais, "Functional Data Analysis") qui s'est considérablement développée depuis une vingtaine d'années, notamment sous l'impulsion des travaux de J.O. Ramsay (Ramsay, 2015). Cette approche fonctionnelle est particulièrement appropriée au traitement des profils spatiaux de vitesse afin de préserver la cohérence physique entre vitesse et position (et implicitement temps). En effet, le fait de se placer dans un cadre fonctionnel permet de tenir compte de la structure sous-jacente continue de ce type de données, ainsi que de leurs caractéristiques fonctionnelles : calcul des dérivées, régularité, contraintes de forme...

Dans une première partie, nous proposons une modélisation fonctionnelle des profils spatiaux de vitesse. Puis dans une seconde partie, nous proposons une méthode de construction d'un profil agrégé (profil moyen) en procédant d'abord à une étape de lissage afin de se ramener à un cadre fonctionnel, puis à une étape d'alignement des profils afin d'obtenir un profil agrégé représentatif de l'ensemble des profils individuels. Enfin, dans une troisième partie, nous proposons la construction d'enveloppes de vitesse reflétant la dispersion des vitesses pratiquées, en utilisant une extension des boîtes à moustaches aux données fonctionnelles.

6.5.2.1. Modélisation fonctionnelle des profils spatiaux de vitesse

Afin de se placer dans le cadre de l'analyse des données fonctionnelles, nous devons définir l'espace fonctionnel des profils spatiaux de vitesse. En effet, toute fonction n'est pas un profil spatial de vitesse (ex: une fonction constante égale à zéro). En pratique, un profil spatial de vitesse est une succession de mesures horodatées de position et de vitesse, et peut donc être manipulé dans les 3 espaces suivants : distance×temps, vitesse×temps ou vitesse×distance (voir Figure 52).

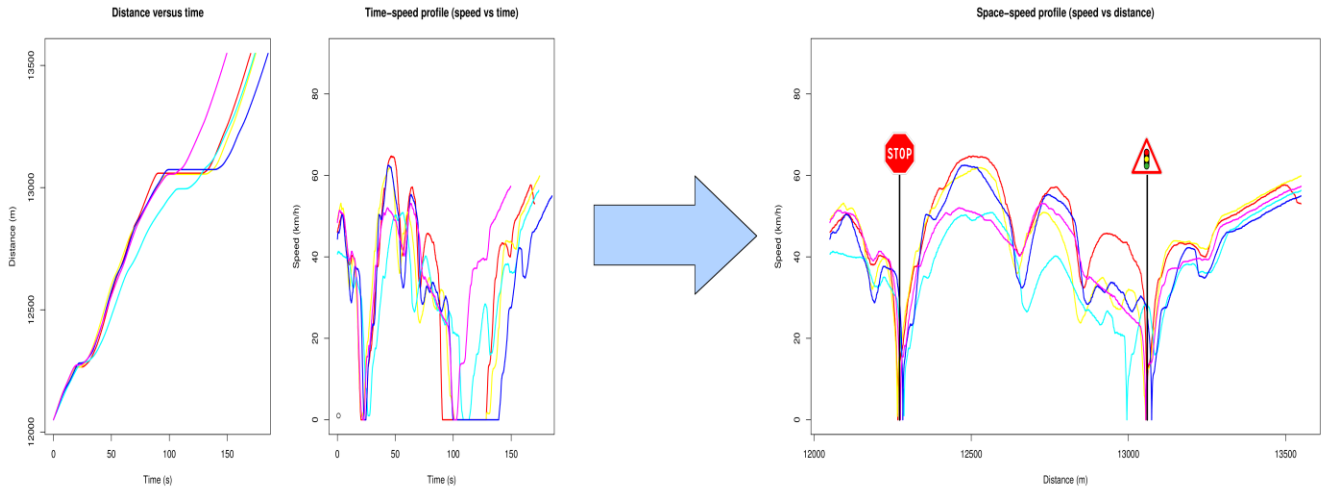


Figure 52 : Lien entre les trois espaces : [distance temps, vitesse temps] et [vitesse distance].

Les fonctions définies dans chacun de ces trois espaces sont définies mathématiquement de la manière suivante : si on note $F(t)$ une fonction de l'espace distance×temps représentant la distance parcourue en fonction du temps, alors sa dérivée $F'(t)$ représentant la vitesse en fonction du temps est définie dans l'espace vitesse×temps. Par définition, la fonction F doit être croissante et au moins de classe \mathbf{C}^1 . Nous proposons de définir l'espace fonctionnel des profils spatiaux de vitesse (fonctions de l'espace vitesse×distance) de la manière suivante :

Definition 1 Soit $x_f \in \mathbf{R}^+$. Alors l'espace des profils spatiaux de vitesse sur $[0, x_f]$, noté \mathbf{E}_{SSP} , est défini de la manière suivante :

$\mathbf{E}_{SSP} = \{v_s : [0, x_f] \rightarrow \mathbf{R}^+ \text{ tel qu'il existe un réel positif } T \text{ et une fonction } \mathbf{C}^2 \text{ croissante } F : [0, T] \rightarrow [0, x_f] \text{ avec } F(0) = 0 \text{ tels que } v_s(x) = F' \circ F^{-1}(x), x \in [0, x_f]\}$, où F' est la dérivée de F par rapport à t , et F^{-1} est l'inverse généralisée de F définie par $F^{-1}(x) = \inf\{t \in [0, T], F(t) = x\}$ pour tout $x \in [0, x_f]$.

Dans cette définition, les réels positifs x_f et T représentent respectivement la longueur et le temps de parcours de la section étudiée. Le diagramme de la Figure 53 illustre le lien entre ces 3 espaces en utilisant les notations de la définition 1. Une étude des propriétés mathématiques des profils spatiaux de vitesse montrent que ces fonctions sont continues et non dérivables aux points où la vitesse est nulle (démonstrations dans Andrieu (2013)).

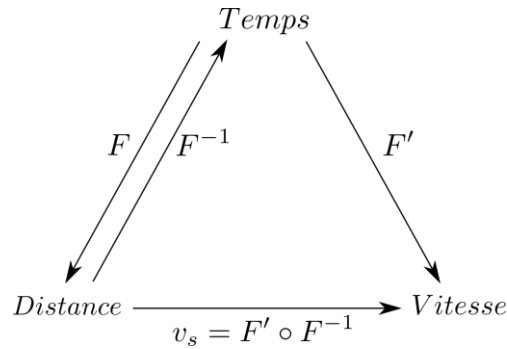


Figure 53 : Diagramme fonctionnel illustrant la définition des profils spatiaux de vitesse.

6.5.2.2. Construction d'un profil de vitesse de référence

Dans le cas de gros volume de données, on obtient un ensemble de profils de vitesse individuels qui sont riches d'information sur le comportement des usagers du réseau, mais qu'il est nécessaire de résumer par des profils agrégés (ex : moyenne, médiane) afin d'en faciliter l'étude et permettre, par exemple, d'estimer les vitesses pratiquées par les usagers de la route le long d'un itinéraire.

Dans la suite, l'application concrète des solutions proposées est faite à partir d'une expérimentation réalisée par le LIVIC entre septembre et octobre 2012 impliquant 39 participants ayant effectué deux fois un trajet de type urbain-résidentiel. Nous nous focalisons ici sur une section d'une longueur d'environ 1100 m comportant un stop, deux ronds-points, et un feu rouge. Les résultats présentés sont donc obtenus à partir de 78 profils spatiaux de vitesse issues de cette expérimentation, et constitués des mesures de position et de vitesse issues du GPS et rééchantillonnées à 1 Hz, soit une mesure par seconde (cf. Figure 54).

6.5.2.2.1. Des données brutes aux courbes : l'étape de lissage

La première étape d'une analyse de données fonctionnelles consiste à convertir les données brutes de nature vectorielle en objet fonctionnel. Les mesures de vitesse et de position issues de capteurs étant bruitées, cette étape nécessite l'utilisation d'une méthode de lissage. Plusieurs études ont été consacrées à l'estimation de profils de vitesse ou d'accélération. Rakha et al. (2001) proposent une comparaison de différentes méthodes de lissage telles que l'estimateur à noyau et le filtre de Kalman, mais montrent que ces méthodes ne sont pas robustes à la présence de valeurs aberrantes et ne sont pas appropriées au calcul des dérivées. Bratt and Ericsson (1999) proposent d'utiliser la méthode des polynômes locaux qui présente l'avantage d'être simple mais qui tend à sur-estimer ou sous-estimer les vitesses. Nous proposons une autre méthode basée sur les splines de lissage qui présente l'avantage de conserver la nature fonctionnelle des profils de vitesse et de faciliter ainsi le calcul des dérivées (ex : profils d'accélération).

L'estimation d'un profil spatial de vitesse à partir de mesures bruitées de position et de vitesse est un problème de régression non paramétrique complexe. En effet, d'une part, la variable réponse correspondant à la vitesse et la variable explicative correspondant à la position sont toutes les deux bruitées. D'autre part, la fonction de régression doit appartenir à l'espace des profils spatiaux de vitesse défini précédemment et doit donc vérifier ses propriétés, en particulier la non dérivabilité aux points où la vitesse est nulle. Afin de pallier ces difficultés, nous proposons de changer d'espace d'étude et d'estimer dans un premier temps la fonction F représentant la distance parcourue en fonction du temps (espace distance×temps de la figure 1). Le nouveau problème de régression non paramétrique est alors le suivant :

$$y_i = F(t_i) + \varepsilon_{x,i}, \quad i = 1, \dots, n, \quad (1)$$

où les y_i sont les mesures bruitées de la distance parcourue du véhicule (issues d'un odomètre ou d'un GPS après map-matching) obtenues en chaque instant t_i , les $\varepsilon_{x,i}$ sont des erreurs aléatoires non corrélées entre elles, de moyenne nulle et de variance σ_x^2 correspondant aux erreurs de mesures, et F est la fonction de régression que l'on cherche à estimer. Cependant, ce problème de régression nécessite la prise en compte de deux contraintes :

(C1). Utiliser l'information sur la dérivée, i.e. estimer la fonction de régression $F(t)$ à partir d'observations bruitées de cette fonction (mesures de position) et également d'observations bruitées de sa dérivée $F'(t)$ (mesures de vitesse).

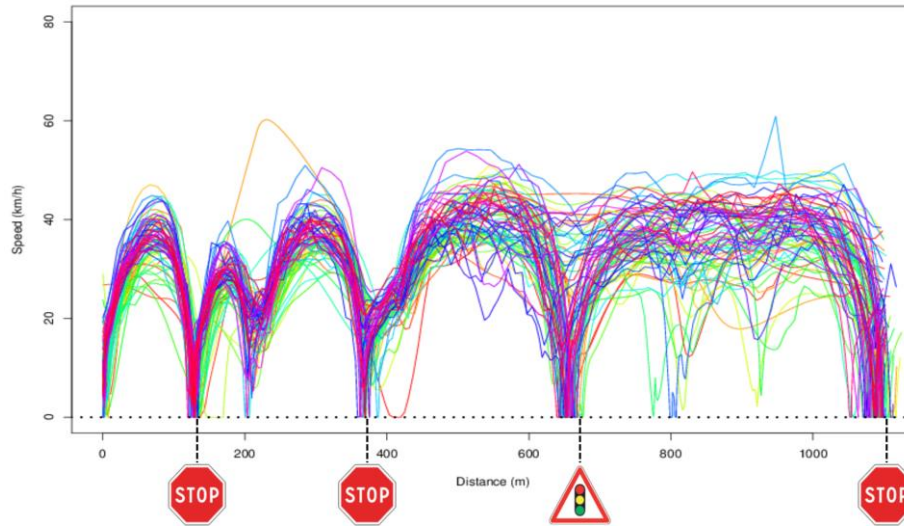
(C2). Une contrainte de monotonie puisque la fonction F représentant la distance parcourue en fonction du temps doit être croissante.

Afin de prendre en compte la 1ère contrainte, nous avons proposé de nous placer dans le cadre d'un problème de régression régularisée dans un espace de Hilbert à noyau reproduisant (RKHS), et de se ramener à un cas particulier du modèle général des splines de lissage (Wahba (1990), Berlinet & Thomas-Agnan (2004)). On obtient alors un estimateur de F dont le degré de lissage est contrôlé par un unique paramètre contrairement à la méthode des polynômes locaux. Puis la seconde contrainte a été prise en compte grâce à une étape de monotonisation de l'estimateur obtenu en prenant en compte la 1ère contrainte. On obtient ainsi un estimateur monotone \hat{F} de F , puis on en déduit par dérivation un estimateur \hat{F}' du profil de vitesse F' en fonction du temps (espace vitesse×temps), et enfin un estimateur $\hat{v}_s = \hat{F}' \circ \hat{F}^{-1}$ du profil spatial de vitesse v_s . Plus de détails sur cette procédure de lissage sont données dans le chapitre 4 de la thèse de Andrieu (2013).

Les résultats de la procédure de lissage appliquée aux 78 profils de vitesse sont illustrés à la Figure 54. On observe une atténuation de certains pics présents dans les données brutes (courbes orange et bleue) qui semblent correspondre à des valeurs aberrantes. Les valeurs

manquantes sont également comblées par ce lissage. Enfin, le principale avantage de la procédure de lissage est ramener l'étude des profils spatiaux de vitesse à un cadre fonctionnel, et ainsi de tenir compte de leurs caractéristiques fonctionnelles telles que les variations de phase.

a. Données brutes et infrastructure associée.



b. Profils fonctionnels obtenus après lissage.

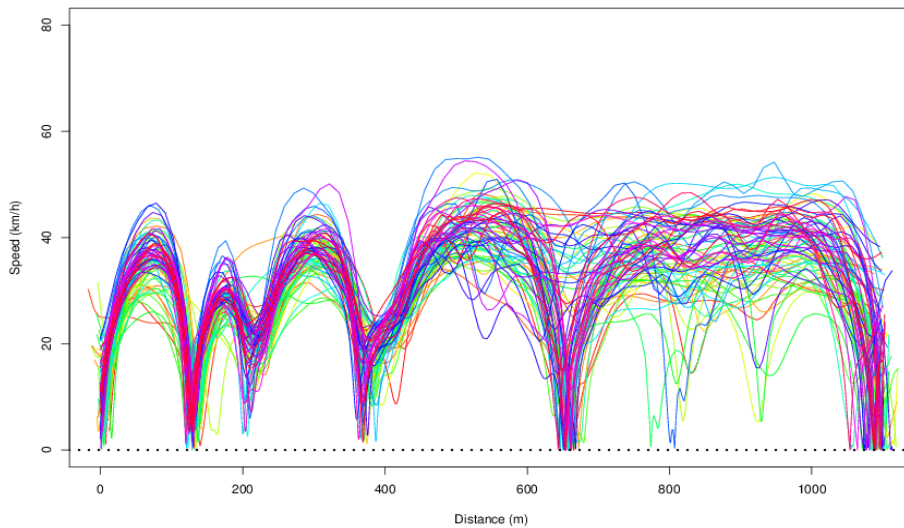


Figure 54 : Application de la procédure de lissage sur 78 profils spatiaux de vitesse.

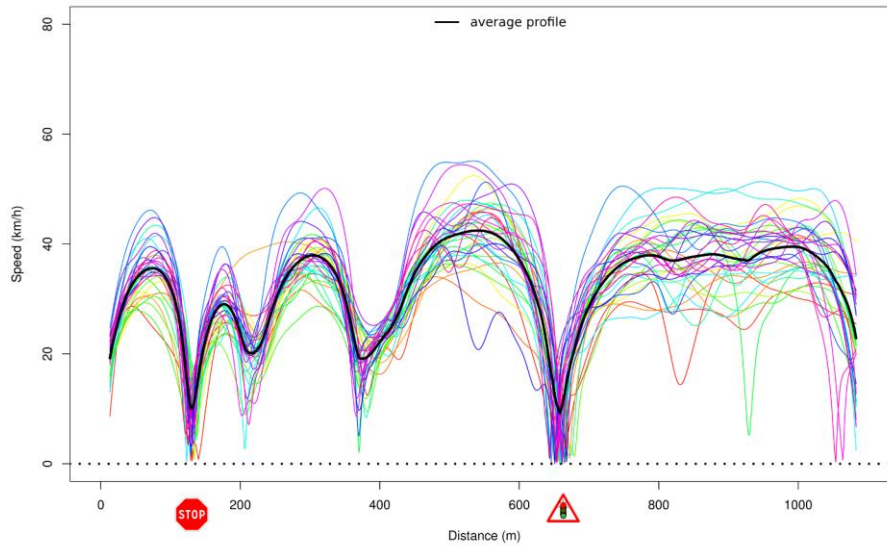
6.5.2.3. Alignement par la méthode des landmarks

Afin d'illustrer la procédure d'alignement des profils de vitesse lissés obtenus, nous utilisons un sous ensemble des 78 profils précédents. En effet, la section comportant un feu rouge, nous illustrerons la démarche avec le sous-ensemble des 36 profils pour lesquels le feu était rouge (arrêt du véhicule au droit du feu).

On observe que le profil moyen représenté en noir à la figure 5.a n'est pas représentatif de l'ensemble des profils. En effet, chaque véhicule ne s'arrête pas exactement au même endroit au feu et au stop (ex : file de véhicules), ce qui entraîne une variation locale de la position de l'arrêt et biaise le calcul d'un profil agrégé tel que le profil moyen. Il est donc nécessaire avant de calculer un profil moyen, de recalcr les profils de vitesse en certains points caractéristiques.

Le problème du recalage (ou de la synchronisation) d'un ensemble de signaux apparaît dans de nombreux domaines tels que la biologie, la météorologie, la reconnaissance de formes... (Ramsay (1998), Bigot (2003)). Dans le cadre de l'étude des profils de vitesse, le problème du recalage apparaît notamment dans Violette (2010) et Kerper (2012). Dans les deux cas, les auteurs proposent de translater les profils initiaux afin de les aligner ("*shift registration*"). Cependant si une simple translation est adéquate pour aligner des signaux en un point, celle-ci n'est pas adaptée dès que l'on cherche à aligner plusieurs points caractéristiques. Afin de pallier ce problème, nous proposons d'utiliser la méthode d'alignement par landmarks qui consiste à déterminer des transformations qui mettent en correspondance un ensemble de points caractéristiques appelés landmarks (ex : maxima, minima, passage par zéro...), qui sont communs à l'ensemble des signaux devant être recalés. Dans notre cas d'étude, nous avons choisi de définir comme landmarks les positions des deux éléments de l'infrastructure imposant un arrêt du véhicule, à savoir le stop et le feu rouge. Les profils de vitesse obtenus après recalage sont représentés à la Figure 55.b. On observe que le profil moyen (en noir) calculé à partir de ces nouveaux profils déformés est bien plus représentatif de l'ensemble des profils que celui obtenu sans recalage.

a. Profils de vitesse avant alignement.



b. Profils de vitesse après alignement.

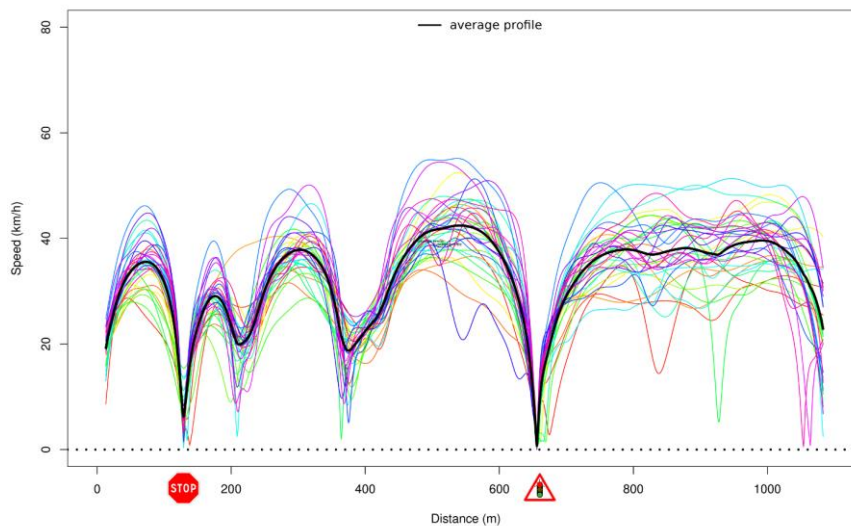


Figure 55 : Alignement des profils par la méthode des landmarks dans le cas où le feu est rouge (36 courbes). Le profil moyen est représenté en noir.

6.5.2.4. Extension à la construction d'enveloppes de vitesse

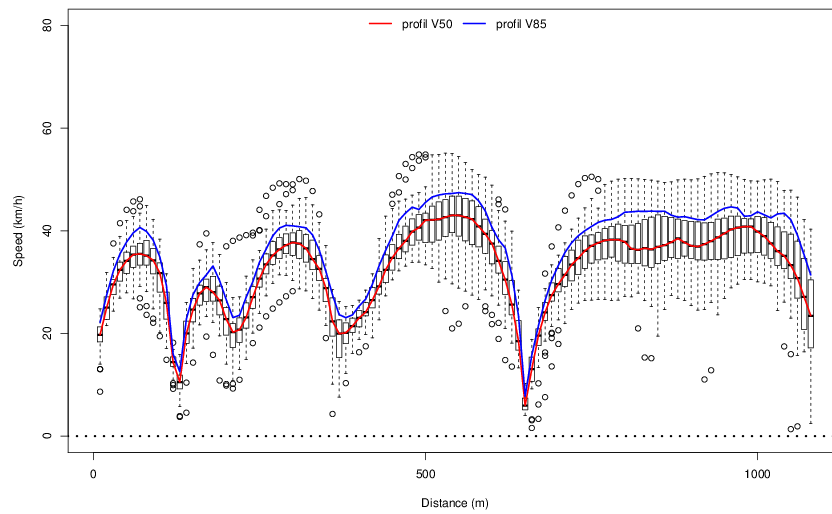
Si la construction d'un profil de vitesse agrégé tel que le profil moyen permet d'obtenir un profil de vitesse représentatif des vitesses moyennes pratiquées sur la section étudiée, un tel profil agrégé ne reflète pas la variabilité des vitesses entre les différents conducteurs. Un outil graphique classiquement utilisé en statistique univariée pour représenter la distribution des valeurs d'une série statistique est la boîte à moustaches (ou "boxplot" en anglais). Cet outil peut être utilisé pour représenter la dispersion des vitesses entre les individus en un point donné. La Figure 56.a représente un ensemble de boîtes à moustaches ponctuelles calculées tous les 10 m sur l'ensemble

de la section étudiée à partir des valeurs des profils de vitesse recalés au niveau des arrêts (cas où le feu est rouge). Une interpolation linéaire entre les points médians (50 centile) et les 85 centiles nous permet de représenter les profils V50 et V85. Cependant, cette représentation ne permet pas de garder le caractère fonctionnel des données, et les profils V50 et V85 ainsi obtenus ne sont pas de "vrai" profils de vitesse au sens de la définition 1.

Nous proposons d'utiliser un outil graphique appelé "functional boxplots", introduit par Sun et Genton (2011) qui permet d'étendre la notion classique de boîte à moustaches aux données fonctionnelles et de représenter ainsi un ensemble de courbes par des enveloppes caractérisant la distribution des courbes. Cet outil repose sur la notion de profondeur qui consiste à mesurer la centralité d'une courbe donnée par rapport à un ensemble de courbes, la courbe ayant la plus grande valeur de profondeur étant la courbe médiane. Nous avons choisi d'utiliser la profondeur modale (en abrégé, hMD pour "*h-mode depth*") introduite par Cuevas (2006), et basée sur le concept de mode (la notion de mode fonctionnel étant définie comme la fonction la plus entourée par les autres courbes du point de vue d'une certaine densité). Les résultats sont illustrés à la Figure 56.b avec en noir le profil médian, en rose foncé la région centrale à 25%, en magenta la région centrale à 50%, en rose pale la région centrale à 75%, et en bleu l'enveloppe maximale (région centrale à 50% augmentée d'un facteur de 1.5 comme dans le cas univariée). Les profils représentés en pointillés rouge sont les courbes situées en dehors de l'enveloppe maximale, et qui sont considérées comme atypiques (outliers).

L'avantage de cette représentation graphique est qu'elle permet de représenter la variabilité des vitesses pratiquées selon les individus sur une section de route donnée. Ces enveloppes de vitesse permettent de distinguer les zones pour lesquelles la variabilité des vitesses est la plus grande, et celles pour lesquelles les vitesses pratiquées sont plus homogènes. De plus, cet outil graphique permet également d'extraire un profil de vitesse représentatif de l'ensemble des profils étudiés (profil médian) au sens de la profondeur statistique choisie, et pouvant être utilisé comme alternative au profil moyen selon l'usage.

a. Boîtes à moustaches ponctuelles représentées tous les 10 m.



b. Boîtes à moustaches fonctionnelles.

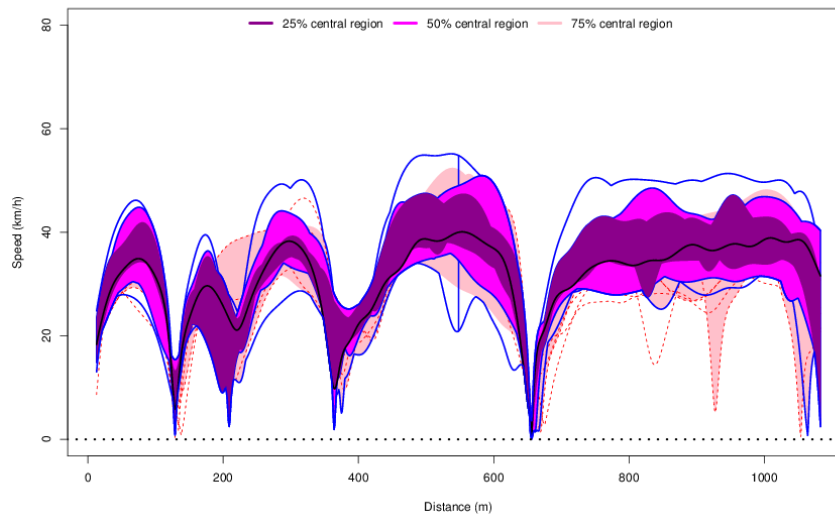


Figure 56 : Boîtes à moustaches ponctuelles et fonctionnelles dans le cas où le feu est rouge (36 courbes).

6.5.2.5. Conclusion

Les données produites par l'étude de la conduite en situation naturelle sont caractéristiques de ce qui sera bientôt accessible aux gestionnaires du réseau routier. Relevant de la catégorie des "big data", ces informations offrent de très nombreuses possibilités d'exploitation statistique.

Ce travail a permis de créer un cadre d'analyse théorique pertinent pour ce type de courbes intimement liées à l'infrastructure et a généré plusieurs publications. Ces travaux ont été

récompensés par deux fois, autant en conférence internationale, que par des acteurs privés (fondation Abertis France, cf. section 4.1.3). Le cadre théorique développé sert de support méthodologique à la poursuite de mes recherches sur ce thème, notamment l'exploitation des traces GPS pour l'ajout d'informations fines dans les cartographies numériques.

6.5.3. Détection de l'infrastructure

La disponibilité de grandes quantités de traces GPS issues de véhicules modernes a conduit à l'émergence de solutions pour leur exploitation dans le contexte de la création de cartographies (cf. par exemple Ahmed et al. 2015). Les traces numériques sont aussi exploitées par les agglomérations pour améliorer la gestion du trafic et la prédiction des temps de parcours (cf. par exemple Delhome et al. 2017), ou sont utilisées pour des applications plus variées, comme la prédiction de destination (cf. Besse et al. 2016) ou l'évaluation des vitesses pratiquées (cf. par exemple les travaux de l'institut belge de sécurité routière, Trotta 2017). L'idée d'utiliser les traces GPS pour construire, améliorer, et corriger des cartographies précises est donc bien répandue. En effet, la faisabilité de véhicules à haut niveau d'automatisation (niveaux 4 et 5) ne peut reposer sur une évolution rapide de l'infrastructure physique (évolution plus lente que la technologie des véhicules), mais s'appuiera très probablement sur une infrastructure digitale très détaillée, incluant des informations dynamiques (souvent appelées LDM, local dynamic map). Suivant la logique ayant mené à la construction de cartes à partir de traces GPS, j'ai cherché à poursuivre la démarche en étudiant la possibilité de détecter des éléments plus détaillés de l'infrastructure, comme par exemple certains panneaux de signalisation, ou les éléments destinés à réduire la vitesse (dos d'ânes par exemple).

Ainsi, en 2015, j'ai proposé une collaboration avec l'équipe COGIT (« Cartographie et Géomatique ») du LaSTIG (laboratoire en sciences et technologies de l'information géographique) de l'IGN autour de l'exploitation des traces GPS pour un usage lié à la cartographie. Il s'agit de développer des méthodes d'apprentissage permettant de détecter la présence de certaines infrastructures routières spécifiques (feux rouge, stop, passage piéton, dos d'âne) à partir d'une collection de traces GPS de véhicules. Cette proposition s'est concrétisée en 2016 par le co-encadrement d'une thèse d'un ingénieur de l'IGN (Yann Méneroux) sur ce sujet.

La cadre d'application des travaux menés peut s'illustrer à partir de la Figure 54-a. La présence de feux rouges et de stops induit des contraintes visibles sur les profils de vitesse, qui sont de nature différentes. Ainsi, un stop sera caractérisé par un arrêt systématique, tandis qu'un feu rouge sera caractérisé par des arrêts et des passages à vitesse normale. Il semble donc possible de faire automatiquement le travail effectué par le cerveau humain lors de l'analyse de ces courbes. Actuellement, les informations de signalisation verticale (feux rouge, stop, cédez le passage) ou horizontale (passages piétons, dos d'ânes) ne sont pas systématiquement incluses dans les cartes numériques. Un travail collaboratif mené par Open Street Map (OSM) et ses contributeurs

permet de recueillir petit à petit ces informations, mais aucune société privée ne propose à ce jour une information précise et fiable pour ces objets. Or, il semble primordial pour les futurs véhicules autonomes, de connaître l'emplacement précis de ces objets afin d'anticiper une prise de décision et une action de conduite, avant la détection par un des capteurs embarqués. Une méthode validée permettant d'exploiter les traces numériques et de détecter la position précise de ces éléments serait un atout considérable pour résoudre les problèmes posés par : l'enrichissement des cartes, la validation des informations qu'elles contiennent, et la correction de ces dernières en cas d'erreur.

La première phase de ces travaux a consisté à s'intéresser uniquement à la détection des feux rouges, en cherchant à les détecter de manière précise à l'aide d'algorithmes d'apprentissage statistique génériques.

On a cherché dans un premier temps à valider l'emploi de telles méthodes sur le cas de la détection de feux tricolores, l'objectif étant ici de répondre à quelques questions exploratoires :

→ Quels sont les espaces de descripteurs adaptés pour modéliser des paquets de profils et y rechercher la signature caractéristique de feux tricolores ?

→ Quels algorithmes d'apprentissage ?

→ Quelles performances peut-on espérer obtenir ?

Ce travail a été mené à partir d'un jeu de données réelles de 144 traces GPS (1Hz) recueillies sur un circuit mixte (urbain et périurbain) de 25 km, comportant 44 feux. Après une phase de prétraitement des données brutes (nettoyage, interpolation, mapmatching), nous avons procédé à la préparation d'un jeu de données adaptées à l'apprentissage statistique. Nous avons ainsi créé une base d'entraînement en procédant à l'extraction de fenêtres glissantes de taille 100m, annotées manuellement comme comportant un feu ($Y=1$) ou non ($Y=0$). Ces fenêtres sont extraites tous les 10 mètres, ce qui entraîne un recouvrement de 90% entre deux fenêtres consécutives (cf. Figure 57). Nous obtenons ainsi un total de 2505 fenêtres, dont 402 (16%) avec un feu rouge (cf. Figure 58).

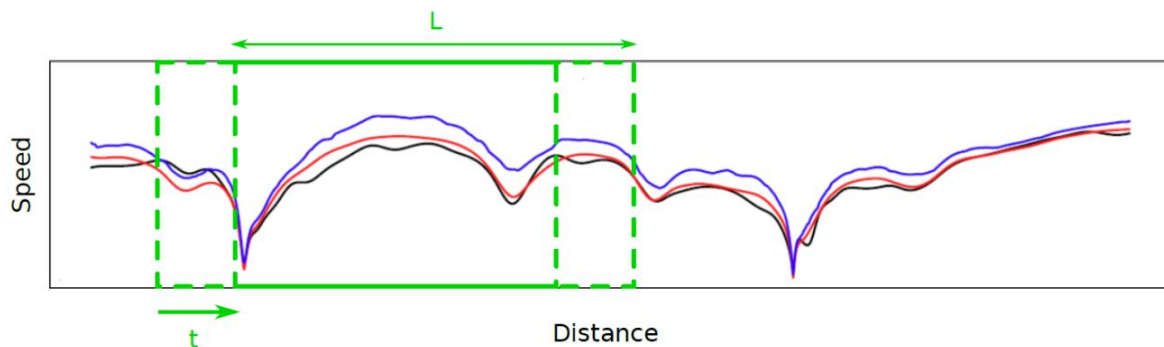


Figure 57 : Illustration de la méthode d'extraction par fenêtres glissantes. Dans notre cas, $L=100m$, et $t=10m$.

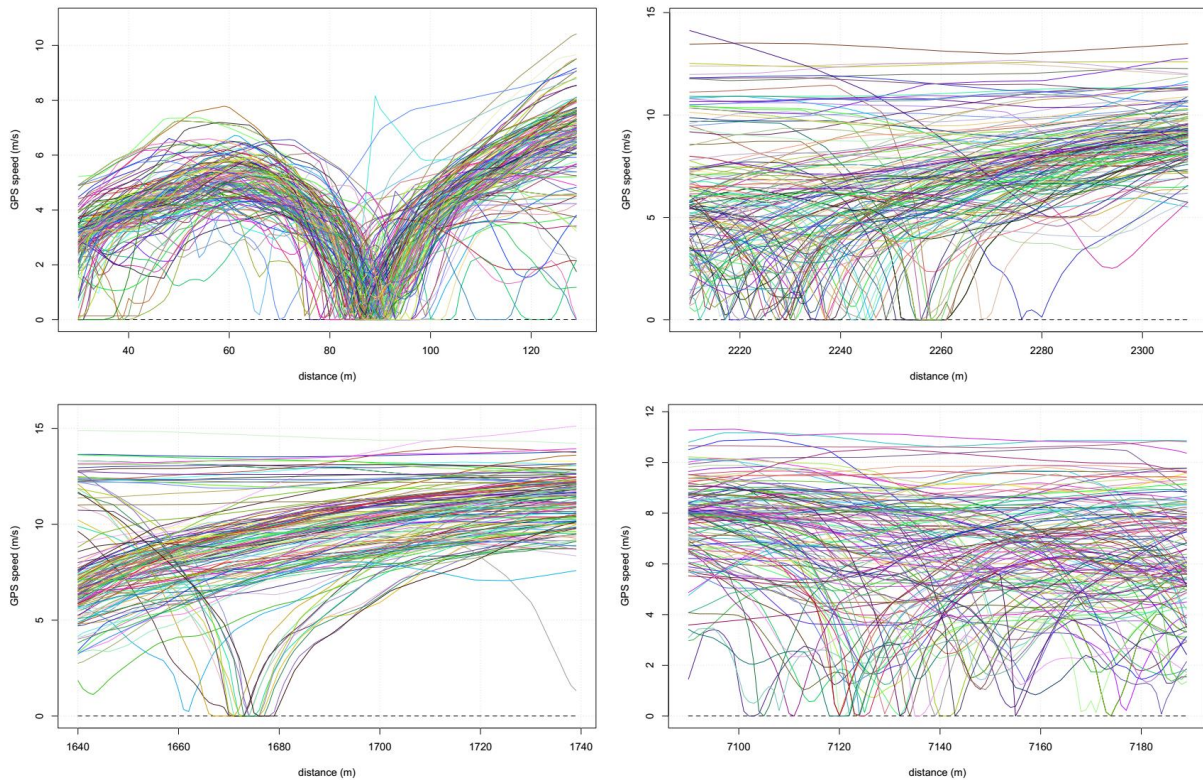


Figure 58 : Quatre exemples de fenêtres obtenues. Seules les deux fenêtres du bas contiennent un feu rouge.

Dans un premier temps, il s'agit de déterminer une modélisation pertinente de ces fenêtres. Quels sont les descripteurs les mieux adaptés pour conserver de manière parcimonieuse, l'information contenue dans ces fenêtres ? Plusieurs approches ont été testées :

- **Approche directe** : Chaque profil de vitesse est discrétisé en espace. Sur une section de 100m, pour un pas de discrétisation de 1m, il y a $144 \times 100 = 14400$ valeurs de vitesse contenues dans une fenêtre. L'approche directe consiste à considérer l'ensemble de ces 14400 valeurs ordonnées selon l'ordre lexicographique, comme un unique vecteur de descripteurs.
- **Approche image** : Consiste à traiter les fenêtres comme des images, en se fondant sur l'hypothèse que la reconnaissance humaine est principalement de nature visuelle. La méthode utilisée pour extraire des descripteurs de l'image est fondée sur la sélection aléatoire de couples de sections de l'image. Pour chaque couple, on calcule une distance X_j normalisée basée sur la différence du nombre de courbes contenues dans chacune des deux sections de l'image (Ozuysal et al (2007)). Les valeurs X_j constituent l'ensemble des descripteurs de l'image. Dans notre cas, nous avons extrait 15000 descripteurs par image.

- **Approche fonctionnelle** : Cette approche parait adaptée à la nature fonctionnelle des données. Afin de ne pas dépendre du nombre de courbes contenues dans la fenêtre, nous avons choisi de décrire les profils de vitesse par un nombre réduit de profils agrégés (profil moyen, médian, standard deviation, moments d'ordre 2 et 3, coefficient de Sarle, 15^{ème} et 85^{ème} percentile, minimum, maximum, etc.). Chacun de ces profils est ensuite projeté sur une base d'ondelettes (Haar), les coefficients de niveau supérieur à 3 étant mis à zéro, ainsi que le premier composant d'ordre 0 (pour ne pas prendre en compte la moyenne globale des vitesses observées). Les descripteurs de la fenêtre sont alors les coefficients non nuls.

Enfin, le processus de classification lui-même, est opéré par différents algorithmes adaptés à notre problème et issus de la littérature.

- **Classifieur bayésien naïf (Naïve Bayes)** : Ce classifieur très simple, appartenant à la famille des classifieurs linéaires, suppose l'indépendance des variables explicatives. Utilisé ici comme référence, ce classifieur reste populaire pour ses bonnes performances dans un contexte d'apprentissage supervisé avec peu de données d'entraînement.
- **Arbres de décision (Decision Tree)** : Connus sous le nom de CART, ce classifieur est très populaire car il est simple et aisément interprétable. Nous utilisons une profondeur maximale de 25 et l'homogénéité des classes est évaluée selon le critère de Gini.
- **K plus proches voisins (K Nearest Neighbors)** : D'usage très simple, cet algorithme ne nécessite comme paramètre que le nombre de classes K, et la définition d'une distance (K=7 et distance euclidienne dans notre cas).
- **Forêts aléatoires (Random Forest)** : Ce classifieur est composé d'un ensemble (forêt) d'arbres de classification, utilisant des sous-ensembles de descripteurs différents, et entraînés sur des sous-ensembles de données différents. Cet algorithme est très utilisé car il est simple, rapide, robuste au sur-apprentissage, et très efficace pour des problèmes de grandes dimensions. Nous utilisons 500 arbres, et 120 descripteurs aléatoires pour l'approche directe, et 13 pour l'approche fonctionnelle.
- **Fougères aléatoires (Random Ferns)** : Cet algorithme peut être considéré comme un classifieur Bayésien semi-naïf, consistant à supposer l'indépendance par groupe des descripteurs. Nous utilisons 1000 fougères (ferns) de 10 variables chacune.

Pour chaque combinaison [approche x algorithme] (15 combinaisons), on entraîne un classifieur et on le teste selon une procédure de 10-fold cross validation. Ce travail, ainsi que l'ensemble des résultats détaillés, est actuellement en cours de relecture par la revue « Data knowledge and discovery ». La Figure 59 présente une partie des résultats obtenus pour les approches images et fonctionnelle, sous la forme de courbes ROC.

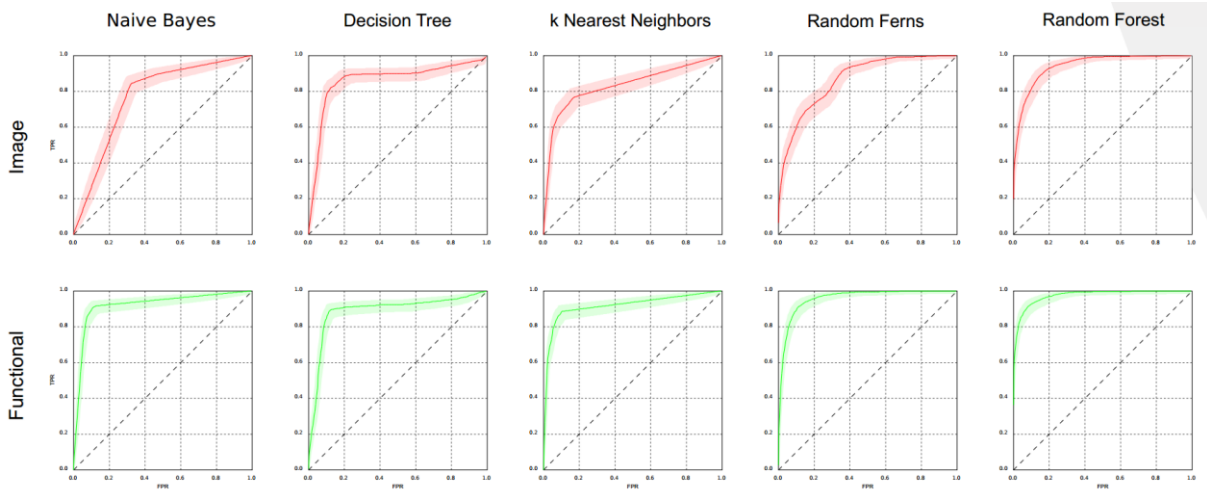


Figure 59 : Courbes ROC et bandes de confiance à 95% (méthode Threshold Averaging avec distribution binomiale) pour chaque combinaison [approche x algorithmique]. Ratio de faux positifs (FPR) en abscisse et vrais positifs (TPR) en ordonnée.

L'aire sous la courbe ROC est plus importante pour les classificateurs de type forêts aléatoires, ce qui indique de meilleures performances pour cette approche. De même, l'approche fonctionnelle donne de meilleurs résultats qu'une approche image. Dans le cas des forêts aléatoires, ces deux approches sont comparées à l'aide d'indices de performances adaptés, et sont présentés au Tableau 20.

L'expérimentation a montré que la méthode des forêts aléatoires semble la plus performante pour identifier la signature d'un feu tricolore dans une collection de profils de vitesse, quel que soit l'approche utilisée, avec un taux de confusion avec les stops marginal. Les résultats démontrent également la supériorité de l'approche fonctionnelle sur tous les algorithmes testés.

Tableau 20 : Performances des approches images et fonctionnelle pour la méthode Random Forest. Tous les écarts sur les indicateurs sont significatifs ($p < 0.01$). ¹ : Taux de confusion avec les signes stops, ² : Temps d'apprentissage (en secondes / fenêtre).

	Image	Fonctionnelle
Sensitivité	64.2 %	82.6 %
Spécificité	96.3 %	97.2 %
Précision	77.7 %	85.6 %
F-mesure	70.3 %	84.1 %
Roc Area	94.0 %	97.5 %
Confusion ¹	6.13 %	2.81 %
Temps ²	5.16"	0.32"

D'autres algorithmes d'apprentissage statistique (tels que les réseaux de neurones fonctionnels) pourront également être testés. L'objectif dans un second temps sera d'étendre les résultats obtenus à d'autres éléments de l'infrastructure routière (éventuellement à l'aide d'autres sources

de données enregistrées à bord du véhicule ou à l'aide des images satellitaires de l'IGN). La portée théorique de ces travaux devrait in fine permettre d'apporter des éléments de réponse aux interrogations suivantes :

Quelle robustesse de l'approche aux changements d'environnement ? Combien de données faut-il avoir à disposition dans la base d'entraînement ? Quelles influences (sur les précisions de détection et de localisation) de la fréquence d'acquisition des capteurs, des erreurs de mesures et du nombre de traces observées ?

Enfin, ces travaux se poursuivront par la mise au point de méthodes de détection d'autres type d'infrastructure comme les stops, les cédez le passage, ou les dos d'ânes. Afin de s'affranchir des jeux de données réduits utilisés actuellement pour valider cette approche, une collaboration avec le CSIS de Tokyo (Center for Spatial Information Science) est en cours afin d'appliquer ces méthodes à des données GPS issues d'une flotte de taxi, couvrant une large zone géographique. Pour le CSIS, les travaux effectués lors de cette collaboration, devraient s'inscrire dans la suite de l'étude DeepTransport (Song et al. 2016), visant à proposer un système intelligent, combinant techniques de machine learning et traces GPS (entre autres), dans le but de simuler et de prédire le trafic automobile, afin d'optimiser l'usage du réseau routier à l'échelle de la ville.

6.5.4. Conclusion

Les travaux de recherche présentés dans cette section, représentent une tentative de participation à l'effort global d'exploitation des données massives générées par la mobilité numérique. Ces données peuvent être de différentes sortes, et provenir de différentes sources, mais leur accès et leur utilisation sont pour le moment compliquées par l'absence de données publiques et la mainmise des entreprises privées sur les moyens de collecte (Coyote, TomTom, Google, etc.).

J'ai ainsi voulu développer des outils simples, destinés à exploiter les données les plus simples à recueillir que sont les données GPS. La généralisation d'outils communautaires d'aide au déplacement, et le développement d'une politique de l'open data en France (cf. par exemple le blog d'Etalab <https://www.etalab.gouv.fr/>) rendent possible à moyen terme l'accès à de grandes masses de traces numériques. Les outils que je développe grâce à ces recherches participent à la création d'un bien commun extrait de ces données brutes, qui se doit d'être restitué à ceux qui contribuent bénévolement à la création et au partage de ces données.

Il va sans dire que la mobilité numérique et l'exploitation des traces générées peut aider à améliorer les connaissances en sécurité routière. La connaissance des pratiques des automobilistes reste un sujet peu ou mal connu. Son étude approfondie pourrait être au cœur des travaux de recherche à venir en relation avec l'avènement de la société numérique. Une collaboration à l'initiative des pouvoirs publics et des organismes de recherche en relation avec les constructeurs automobiles, les sociétés de service (navigation, assurance) devrait être initiée. Cette

collaboration pourrait avantageusement être concomitante avec le développement des véhicules connectés, assistés et autonomes qui redéfinissent les responsabilités des acteurs. Dans une perspective de réduction drastique des accidents de la route (perspectives à 2020, vision zéro), c'est une alternative obligée pour pouvoir continuer à progresser dans le domaine de la sécurité routière.

J'ambitionne que mes recherches contribuent à la meilleure compréhension du comportement des conducteurs, et à ce titre je collabore régulièrement avec des chercheurs en sciences humaines pour associer des méthodes objectives et subjectives d'évaluation du risque routier. Cet aspect de mes recherches est présenté à la section suivante.

6.5.5. Publications associées à cette section

Communications avec actes dans congrès international (ACTI)

- Méneroux Y., Le Guilcher A., Orfila O., Lusetti B., **Saint Pierre G.**, and Mustière S. (2017), « Using Surrogate Road Network for Map-Matching: A Sensitivity Analysis of Positional Accuracy », International Conference on GeoComputation GeoComp2017, Leeds, UK, 2017-9-4.
- Andrieu C., **Saint Pierre G.**, Bressaud X. (2013), « *Estimation of space-speed profiles: A functional approach using smoothing splines* », Intelligent Vehicles Symposium 2013, June 23-26, 2013, Gold Coast, Australia, DOI: 10.1109/IVS.2013.6629594.
- Andrieu C., **Saint Pierre G.** (2012), « *Speed profiles functional modeling for driving behavior and infrastructure evaluation* », Young European Arena of Research, YEAR 2012 (Finalist and bronze medal), Athènes, Greece.

Communications avec actes dans congrès national (ACTN)

- **Saint Pierre G.**, Andrieu C. (2013), « *Statistique et sécurité routière : L'observation de la conduite en situation naturelle* », 45ème journées de Statistique, Mai 27-31, 2013, Toulouse, France.
- Andrieu C., **Saint Pierre G.**, Bressaud X. (2012), « *Modélisation fonctionnelle et lissage sous contraintes d'un profil spatial de vitesse* », 44èmes Journées de Statistique, Bruxelles, Belgique.

Autres publications (AP)

- Andrieu C., **Saint Pierre G.**, Bressaud X., (2013), « *A moving fixed-interval filter/smoother for estimation of vehicle position using odometer and map-matched GPS* » (lien HAL : <http://hal.archives-ouvertes.fr/hal-00915402>, lien ArXiv : <http://arxiv.org/abs/1312.2129>).
- Andrieu C., **Saint Pierre G.**, Bressaud X., (2013), « *A functional analysis of speed profiles: smoothing using derivative information, curve registration, and functional*

boxplot» (lien HAL : <http://hal.archives-ouvertes.fr/hal-00915475>, lien ArXiv : <http://arxiv.org/abs/1312.2252>).

6.6. Statistique appliquée et sciences humaines : collaborations en psychologie expérimentale et acoustique.

6.6.1. Introduction

La statistique mathématique est un domaine transversal à plusieurs disciplines, et les modèles détaillés à la section 6.2.6 sont régulièrement utilisés dans le contexte des sciences humaines. En outre, l'étude du triptyque véhicule-infrastructure-conducteur implique nécessairement de s'intéresser à l'objet situé « entre le siège et le volant » qu'est l'humain. C'est donc très naturellement que j'ai collaboré tout au long de mes recherches, avec des chercheurs en sciences humaines dans différents domaines.

Ma première expérience de cet ordre a eu lieu avant d'intégrer le LCPC puis l'IFSTTAR, dans le cadre d'un contrat de travail impliquant l'entreprise Renault et l'IRCAM (Institut de recherche et coordination acoustique/musique). Les travaux menés sont rapidement évoqués à la section 6.6.2.

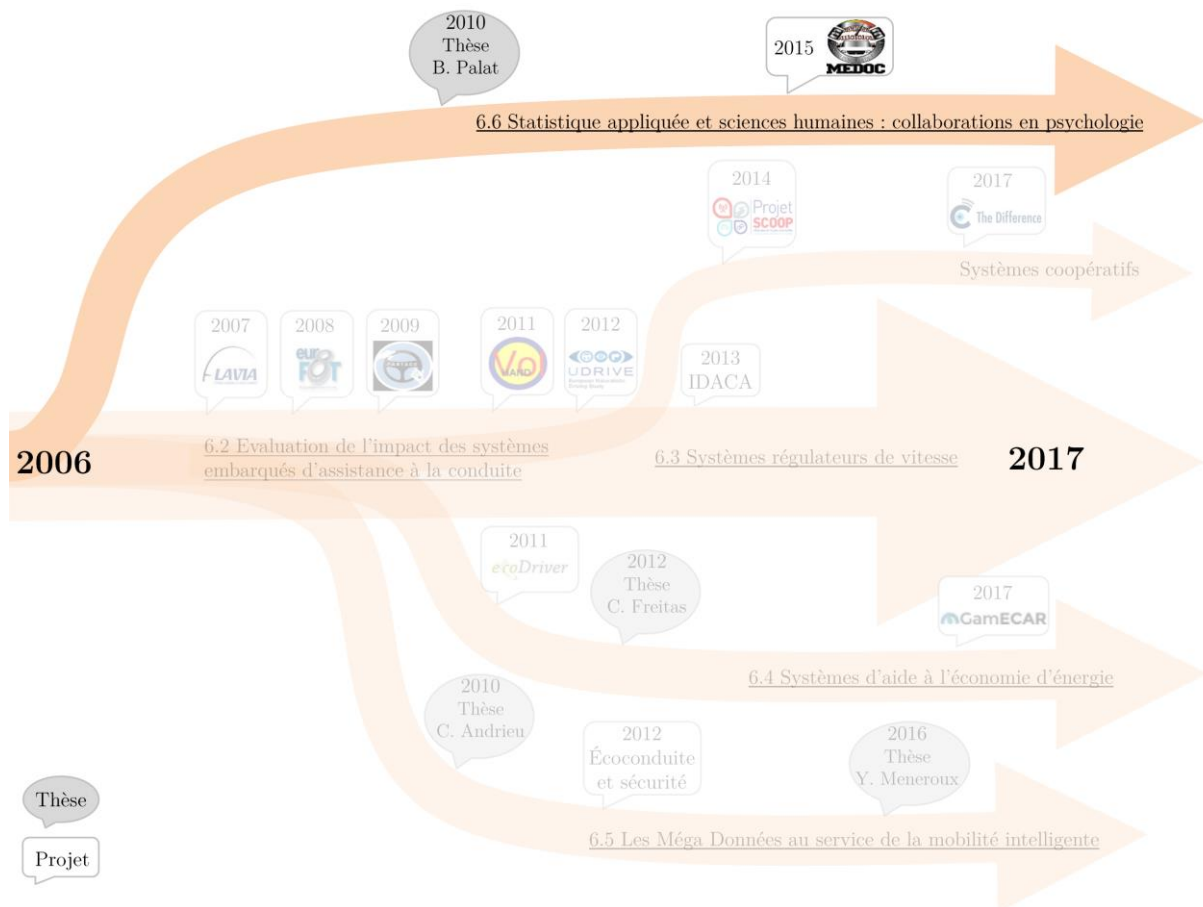


Figure 60 : Mise en évidence des axes de recherche développés dans cette section.

Comme déjà évoqué plus avant, cette expérience m'a permis de comprendre l'importance de la prise en compte de l'humain pour l'évaluation des systèmes embarqués. J'ai donc souhaité

collaborer avec mes collègues chercheurs en sciences humaines pour mieux comprendre comment intégrer cette dimension dans mes recherches, et leur apporter un soutien méthodologique pour leurs propres recherches.

6.6.2. Psychophysique

Lors de ce premier passage à l'IRCAM, j'ai pu me familiariser avec les différents descripteurs du son, mais j'ai aussi et surtout découvert la richesse et la complexité de certains protocoles d'observation utilisés en psychologie expérimentale. Ces protocoles utilisent des mesures répétées sur un faible nombre de sujets et ne respectent donc pas l'hypothèse d'indépendance des modèles statistiques usuels. Ils nécessitent la mise en œuvre de modèles adaptés, qui peuvent être soit les modèles linéaires généralisés mixtes (GLMM), soit les modèles d'équations d'estimations généralisées (GEE). Leur emploi et leur analyse sont complexes, et cela requiert une certaine expertise que j'ai su apporter à l'IRCAM.

Ce bon souvenir et ma connaissance du domaine de l'acoustique, m'a naturellement conduit à continuer à échanger autour de diverses questions liées à la perception des sons, notamment ceux d'un véhicule, l'équipe de perception et design sonores travaillant régulièrement sur ce sujet. Ainsi est née une collaboration qui dure depuis de nombreuses années et qui concerne la perception des sons selon leur intensité (loudness). Elle s'est concrétisé par des publications communes suite à différentes études menée par l'équipe « Perception et design sonore » de l'IRCAM.

La première étude analyse la façon dont la temporalité de l'intensité d'un son influence la manière dont il est perçu. Nous avons montré qu'un son montant, ou descendant, est perçu comme un son d'une intensité au maximum de sa variation. Un son montant ou descendant, est donc toujours perçu plus fort qu'un son de même intensité.

Dans la seconde étude, nous comparons deux protocoles expérimentaux différents pour évaluer la perception des sons de véhicules, et observons leur sensibilité à la variation de l'intensité sonore (même sons mais en appliquant une égalisation d'intensité). Un de ces protocoles utilisait des méthodes de type MDS (multidimensional scaling) sur des données de distances entre sons (similaires au cas étudié pendant mon post-doctorat à l'INRIA). L'effet de la non-égalisation du son surpasse les autres caractéristiques du son, et cela produit un artefact numérique assez rare dans les résultats d'une analyse par MDS : l'effet « fer à cheval ». Cet effet et l'analyse détaillée que nous menons prouve qu'il est nécessaire de procéder à une égalisation en sonie avant toute expérimentation sur les facteurs de la perception d'un son, et nous recommandons d'utiliser une approche par dissimilarité (pseudo-distances entre sons). Ces travaux ont été publiés dans Ponsot, Susini, Saint Pierre et Meunier (2013) ainsi que dans Susini, Houix et Saint Pierre (2015).

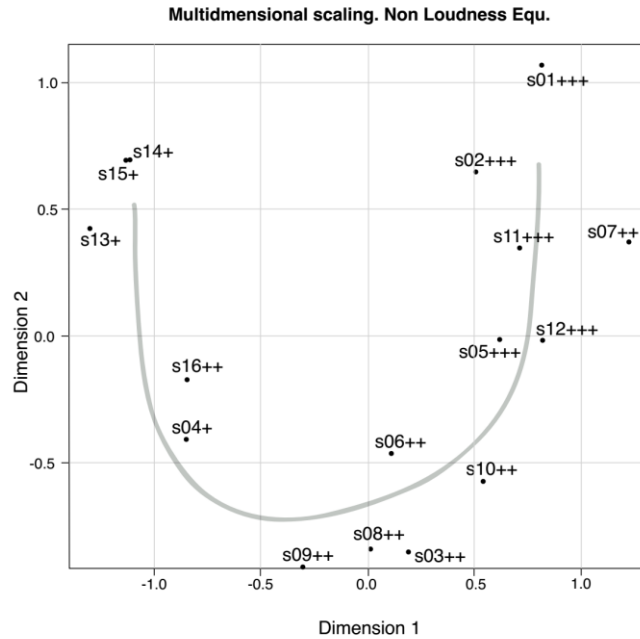


Figure 61 : Effet "fer à cheval" produit par une projection sur un espace à deux dimensions de sons identifiés par les « pseudo-distances » les séparant. Les sons de la catégorie « +++ » sont artificiellement proches des sons de la catégorie « + », et l'algorithme recommande l'ajout d'une dimension supplémentaire (non réellement présente) pour tenir compte de cet aspect. Ce phénomène est dû à l'absence de la propriété d'inégalité triangulaire de la pseudo-distance utilisée.

6.6.3. Psychologie du comportement et des mobilités

Suite à ma première expérience à l'IRCAM et à l'INRETS, je connaissais mieux les problématiques des psychologues et celles du transport. C'est naturellement, qu'une fois à l'Ifsttar, j'ai pu répondre à certaines attentes de mes collègues en psychologie de la conduite et étude du comportement. En effet, les protocoles employés par mes collègues utilisent aussi le principe des mesures répétées dans le cadre d'expériences sur simulateurs de conduite par exemple. Mon goût pour les questions concrètes posées par ces chercheurs m'a encouragé à collaborer et mettre à profit mes connaissances théoriques. Cet échange a été possible grâce au rapprochement géographique du laboratoire de psychologie du comportement et des mobilités (LPC, Ifsttar) de mon lieu d'affectation (Versailles-Satory). J'ai ainsi pu collaborer avec deux chercheurs de ce laboratoire : V. Cavallo (DR), et P. Delhomme (DR).

6.6.3.1. Perception des motocyclistes

J'ai utilisé les modèles d'équation d'estimation généralisées (GEE) dans le cadre de mes travaux avec V. Cavallo (DR, Ifsttar). La question concernait cette fois-ci la perception des motocyclistes dans l'environnement routier. En effet, avec la généralisation de l'allumage des feux de route pendant la journée, les motocyclistes ne sont plus différenciés. Cette étude réfléchit donc à de

nouvelles manières de signaler ces usagers de la route avec de nouvelles configurations de feux (sur les poignées, sur le casque, feux jaunes, cf. Figure 62), et expérimente leur perception par des sujets dans un contexte routier réaliste. Nous montrons que la différenciation, liée aux feux jaunes ou au feu casque, améliore la perception des motocyclistes (cf. Figure 63) et suggérons d'explorer cette voie d'amélioration de la sécurité.

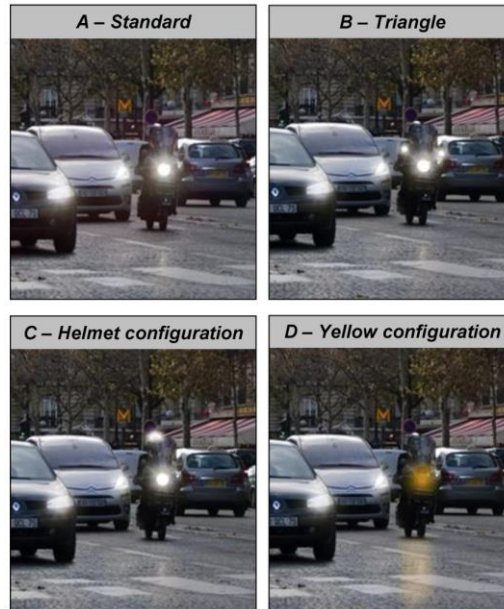


Figure 62 : Exemple des 4 configurations de signalisation lumineuse étudiées.

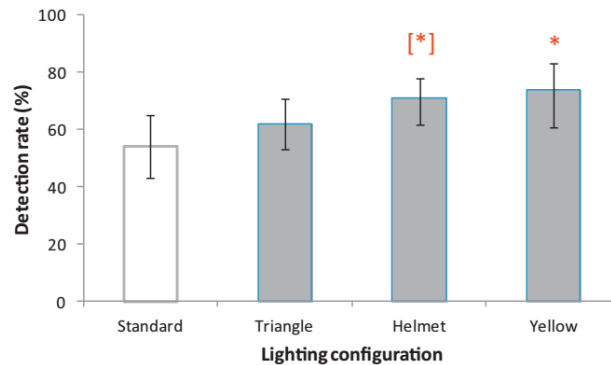


Figure 63 : Moyennes estimées par un modèle GEE et intervalles de confiance pour le taux de détection en fonction de la configuration lumineuse. La différence est significative entre la configuration standard et jaune (astérisque) et faiblement significative entre la configuration standard et celle sur le casque (astérisque entre crochets). La configuration « casque » est cependant mieux détectée que toutes les autres lorsque le motocycliste se situe à grande distance de l'observateur.

La collaboration avec Viola Cavallo se poursuit actuellement avec l'étude des facteurs influençant la prise de décision d'une traversée de chaussée. Un article de conférence est accepté dans la conférence RSS2017 et un article de revue en cours de préparation.

6.6.3.2. Perception et choix de trajet

Je collabore régulièrement avec P. Delhomme, dont les thèmes de recherche recoupent certaines de mes préoccupations : le comportement du conducteur et les économies de carburant. Cette collaboration a débuté par le co-encadrement d'un doctorant travaillant sur l'heuristique de numérosité dans la perception des feux tricolores et le choix de trajet. J'ai apporté mon expertise statistique pour l'analyse et l'interprétation des résultats d'une partie de ses travaux. Nous avons examiné quelles informations concernant les feux tricolores sont utilisées comme critère pour choisir un trajet en milieu urbain. L'information sur le nombre de feux tricolores et sur la durée des feux rouges et verts est donnée aux participants. Nous avons montré que l'information sur le nombre de feux tricolores est alors privilégiée comme critère de choix lorsque l'écart du nombre de feux entre deux trajets est grand. Néanmoins, les participants tendent à éviter de longs temps d'attente. L'utilisation de ces deux critères dans les choix ne traduit pas forcément la prise en compte du seul critère de temps, mais probablement aussi celui du confort. Le détail de ces travaux est publié dans Palat, Delhomme et Saint Pierre (2014).

6.6.3.3. Perception du risque (MEDOC)

La suite de cette collaboration a été fructueuse et s'est traduite par des études communes menées pour le compte du ministère de l'écologie, du développement durable et de l'énergie (MEDDE) concernant l'étude du comportement des écoconducteurs (cf. section 6.4). Cette collaboration est toujours très active, et nous coordonnons ensemble un projet de recherche financé par la fondation MAIF dont l'objectif est d'évaluer la pertinence d'une connaissance des zones où les « accidents évités » sont fréquents pour améliorer la sécurité routière (projet MEDOC).

Le projet MEDOC vise à apporter un éclairage sur les comportements de déplacement des automobilistes, et en particulier en ce qui concerne les situations de « défaillances » qui auraient pu mener à un accident plus grave. Ce que l'on dénomme les incidents, ou les « presque-accidents », est donc souvent caractérisée par la mise en œuvre d'une action correctrice de la part du conducteur. La détection du caractère potentiellement convergent vers le risque de ces situations est cependant complexe, car il est très difficile d'appréhender le contexte de conduite avec les seules données issues du véhicule. En effet, des actions de contre-mesures peuvent être entreprises pour différentes raisons telles que prendre un risque ou à l'inverse éviter un risque. Ce sont donc des stratégies non dénuées d'ambiguïté dont on ignore l'importance. En l'état actuel des connaissances, il n'existe pas de méthode fiable de détection des événements à risque de conduite.

Ce projet a donc pour objectif de comprendre les situations de conduite critiques susceptibles de converger vers l'accident, de faire le lien avec les données dynamiques de conduite et ensuite proposer des recommandations pour en diminuer l'occurrence. Ce travail peut s'effectuer à partir des informations suivantes :

- Historique de l'accidentologie routière.
- Mesures liées à la dynamique du véhicule/aux actions de conduite.
- Informations recueillies à partir des déclarations des participant(e)s sur les conditions dans lesquelles ces incidents surviennent (usagers impliqués et leurs comportements, infrastructure, responsabilité de l'incident, contexte personnel des participant(e)s) et leur gestion.

Le projet MEDOC essaye de mettre à profit la connaissance de ces différentes notions pour en comprendre les liens. En effet, il n'est pas encore avéré que les événements enregistrés de manière automatique grâce aux données du véhicule sont réellement précurseurs d'un accident. Mais si ce lien pouvait être mis en évidence, il serait alors possible de quantifier objectivement le risque d'un conducteur au moyen des informations de la dynamique de son véhicule.

Afin de progresser dans cette voie, j'ai proposé de mener une détection des incidents selon l'état de l'art international, à partir de l'enregistrement de la dynamique des véhicules d'une flotte de sociétaires MAIF (Objectif de 200 participants). Une fois détectés, ces incidents peuvent :

- Etre analysés individuellement pour comprendre les raisons de leur occurrence,
- Etre utilisés sous forme agrégée pour en déduire des indicateurs pouvant être associés à la mesure d'un risque individuel,
- Etre analysés conjointement avec les mesures auto-rapportées pour valider ou infirmer la présence d'un lien entre mesures objectives et subjectives.

Le principe de la détection et l'analyse des incidents est rappelé à la Figure 26, tandis que les définitions utilisées pour la détection des incidents ont été données au Tableau 5.

Dans le contexte du projet MEDOC, le financement ne permet pas d'analyser individuellement les quelques 10000 incidents détectés. J'ai donc procédé à l'extraction d'indicateurs agrégés (par exemple, nombre de freinages brusques au 100km pour chaque conducteur) que je cherche à mettre en lien avec les réponses des usagers quant à leurs caractéristiques ou leurs performances de conduite.

A l'heure de l'écriture de ce document, les résultats finaux ne sont pas encore disponibles, mais les premiers résultats tendent à prouver que les détections obtenues à partir du paramètre Jerk (dérivée de l'accélération) sont plus pertinentes et mieux associées au risque réel ou perçu que les détections fondées uniquement sur l'accélération ou le yaw rate.

6.6.4. Publications associées à cette section

Articles dans des revues à comité de lecture répertoriées (ACL)

- Susini P., Houix O., **Saint Pierre G.** (2015), « *Loudness effect on pairwise comparisons and sorting task* », J. Acoust. Soc. Am. 138, 1918, DOI:10.1121/1.4934034 (*JCR, Psycho-acoustic, IF=1.503*).
- Susini P., Houix O., **Saint Pierre G.** (2015), « *The Effect of Loudness on the Perceptual Representation of Sounds With Similar Timbre* », Acta Acustica United With Acustica Vol. 101, N. 6, pp. 1174-1184(11), DOI:10.1016/j.trc.2015.04.026 (*JCR, Psycho-acoustic, IF=0.783*).
- Palat B., Delhomme P., **Saint Pierre G.** (2014), « *Numerosity Heuristic in Route Choice Considering the Presence of Traffic Lights* », Transportation Research Part F: Traffic Psychology and Behaviour, vol 22, pp 104-112, DOI:10.1016/j.trf.2013.11.004 (*JCR, Transportation science, IF=1.473*).
- Pinto M., Cavallo V., **Saint Pierre G.**, (2013) « *Influence of Front Light Configuration on the Visual Conspicuity of Motorcycles* », Accident Analysis & Prevention, vol 62, pp 230-237, DOI:10.1016/j.aap.2013.09.026 (*JCR, Transportation science, IF=2.07*).
- Ponsot E., Susini P., **Saint Pierre G.**, Meunier S. (2013), « *Temporal weights of loudness for sounds with increasing and decreasing intensity profiles* », The Journal of acoustical society of America, 10/2013; 134(4):EL321-EL326. DOI:10.1121/1.4819184 (*JCR, Psycho-acoustic, IF=1.503*).

Communications avec actes dans congrès international (ACTI)

- Susini P., Houix O., **Saint Pierre G.**, 2015, « Loudness effect on pairwise comparisons and sorting », 170th ASA meeting, Jacksonville, ETATS-UNIS, 2015-11-02, 138, 1918

Communications avec actes dans congrès national (ACTN)

- Palat B., Delhomme P., **Saint Pierre G.** (2012), « *Heuristique de numérosité dans la perception des feux tricolores et le choix de trajet* », 54ième Congrès de la Société Française de Psychologie, Montpellier, 3-5 septembre 2012.

Communications invitées sans actes (INV)

- **Saint Pierre G.** (2015), « *Quantifying ecodriving performances: efficient feedback for efficient driving* », exposé invité au workshop « Human factors of green driving », TRB 94th Annual Meeting, Washington, 11 janvier 2015.
- **Saint Pierre G.** (2012), « *Statistical modeling of safety events in naturalistic driving data: An application to speed regulation systems* », conférencier invité à la "5th International Conference on Traffic and Transport Psychology", ICTTP 2012, Groningen, The Netherlands, August 29-31, 2012.

7. Projet de recherche

Le premier mars 2017, j'ai intégré le Cerema (Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement) de Toulouse afin de mener à bien un projet de vie familiale, en accord avec ma carrière de chercheur. J'ai la chance d'occuper un poste de chargé de recherche dans le domaine de l'évaluation des systèmes d'aide au déplacement, parfaitement en adéquation avec mon expertise et mes projets futurs. Je souhaite poursuivre les travaux entamés, et continuer mes collaborations dans le domaine des sciences humaines. Cela passera par la construction d'une équipe-projet-recherche interne au Cerema, associant l'équipe de Toulouse (Louahdi Khoudour, directeur de recherche en informatique, et moi-même) avec des chercheurs de Saint Briec travaillant sur le comportement de l'utilisateur en lien avec l'infrastructure (psychologie), et des chercheurs spécialisés dans la mobilité et la sécurité des déplacements en situations dégradées situés à Clermont-Ferrand.

La préfiguration de cette équipe a été validée par le Cerema, et les thématiques envisagées me permettront de poursuivre mes recherches tout en collaborant étroitement avec mes nouveaux collègues. Cette équipe de recherche implique 6 chercheurs, 5 ingénieurs, 6 techniciens supérieurs, ainsi que 6 doctorants et 2 post-doctorants. Elle se structure selon deux axes :

- Axe 1 : Réseaux de Transports et systèmes de transport intelligents
- Axe 2 : Vers l'automatisation de la conduite.

J'ai accepté la responsabilité de l'animation du premier axe de cette équipe qui vise à la production de connaissances associées aux impacts sur la sécurité, l'environnement, et l'efficacité des déplacements, des systèmes d'assistance aux déplacements actuellement déployés ou en passe de l'être. Ce cadre institutionnel me permettra de poursuivre sereinement mes recherches en collaboration étroite avec mes collègues chercheurs en sciences humaines, afin d'intégrer au mieux le comportement du conducteur dans l'évaluation des systèmes de transports intelligents.

En tant que chercheur du Cerema, je me dois d'articuler mes recherches avec le tissu universitaire et économique local. Le contexte toulousain est très favorable à une recherche appliquée au service de l'industrie et du citoyen. En effet, Toulouse héberge un des quatre fabricants mondiaux de navettes autonomes « Easy Mile », et l'entreprise « Continental » va y implanter un second centre de recherche sur les véhicules intelligents et connectés. En outre, les ambitions d'une « licorne » comme Sigfox dans le domaine de l'internet des objets (une voiture intelligente est aussi un objet connecté), promettent à la région toulousaine une place de premier rang dans l'innovation pour la mobilité intelligente et connectée. Dans le cadre du pôle de compétitivité Aerospace Valley, un de mes rôles de chercheur, sera de contribuer à l'innovation par le biais de projets de recherche impliquant les partenaires économiques locaux, mais aussi les écoles d'ingénieurs et universités locales. A ce titre, la présence de mon lieu de travail en plein cœur du campus scientifique est un atout et une motivation supplémentaire.

Sans pour autant abandonner l'évaluation des systèmes embarqués, et en restant en accord avec mes nouvelles missions, j'ai pour objectif de poursuivre mes recherches personnelles selon deux axes prometteurs :

1. Développer une méthodologie adaptée à l'évaluation des systèmes coopératifs, pour répondre aux enjeux de société.
2. Etudier la performance des algorithmes d'apprentissage statistique sur des masses de données véhicules pour en extraire de l'information susceptible d'améliorer l'expérience de la mobilité individuelle et de faciliter l'arrivée des véhicules autonomes.

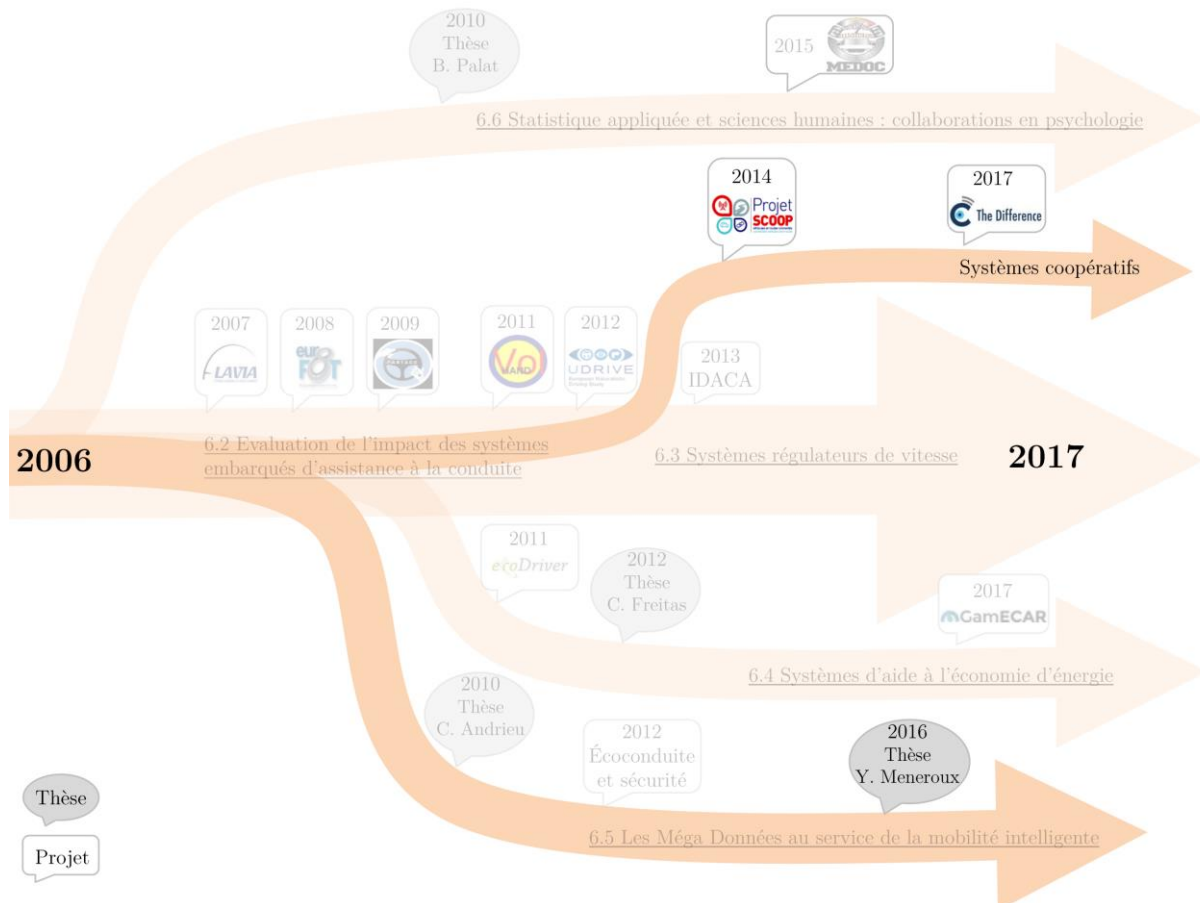


Figure 64 : Mise en évidence des axes de recherche que je souhaite développer dans le cadre de ma prise de poste au Cerema.

7.1. Impact des systèmes coopératifs de transports intelligents

L'évolution des systèmes de transport intelligents étant constante, l'enjeu s'est déplacé des systèmes embarqués vers les systèmes coopératifs, capables d'échanger des informations. Ces systèmes ne délivrent pas constamment des messages, mais seulement à certaines conditions (présence de travaux par exemple). En ce sens, ils se rapprochent des systèmes activés manuellement par le conducteur (régulateur de vitesse par exemple), à ceci près que ce n'est plus le conducteur qui contrôle leur activation. Ces spécificités posent des questions de recherche particulières, pour lesquelles la méthodologie des FOT n'est pas transposable directement.

En effet, l'évaluation des systèmes embarqués classiques considère que chaque système peut être testé indépendamment, et repose sur la comparaison entre une conduite normale (sans système) et une conduite avec le système. La question se complexifie lorsque plusieurs systèmes interagissent (limiteur de vitesse, affichage tête haute, régulateur adaptatif etc.), et constitue un véritable challenge lorsque ces systèmes communiquent entre eux. Mon implication forte dans le projet SCOOP@F, projet phare de l'Etat français, représente pour moi un engagement dans ce nouvel axe de recherche : comment évaluer les performances à grande échelle d'une flotte de véhicules communicants ?

En effet, les enjeux économiques poussent les acteurs industriels à déployer ces systèmes rapidement, sans procéder à une évaluation préalable en conditions contrôlées. Le protocole d'étude de la conduite en situation naturelle, est ici combiné avec un usage continu de systèmes embarqués s'activant dans certaines circonstances dépendant des messages reçus par les autres véhicules équipés ou l'infrastructure (cas d'un feu qui envoie le temps d'attente restant par exemple). L'enjeu est donc de proposer des méthodes statistiques pertinentes pour d'une part, procéder à une réduction des données afin d'en faciliter le traitement de type « big data », et d'autre part pour estimer et extrapoler les effets observés sur la sécurité, l'environnement et l'efficacité des déplacements. L'enjeu est très important, autant économique que social, car les attentes autour de ces systèmes sont parfois immenses. Il est essentiel que la recherche publique propose des méthodes fiables pour estimer au mieux ses effets, sans surestimer les bénéfices potentiels. Mon expérience de l'étude des systèmes embarqués non connectés (limiteur, régulateur, écoconduite) à grande échelle, et mon excellente connaissance du milieu industriel automobile, m'incitent à prendre à bras le corps cette question. Je me suis donc engagé récemment dans la phase 2 de SCOOP@F qui concerne justement l'évaluation, ainsi que dans le projet C-The-Difference cherchant à déployer et évaluer en conditions réelles, un système connectant les conducteurs à l'infrastructure grâce à leurs smartphones. Dans le cadre de ces projets je coordonnerai une partie des travaux d'évaluation.

Le véhicule connecté peut être considéré comme préfigurant les véhicules autonomes dont l'arrivée prochaine sur nos routes est annoncée par les constructeurs. A ce titre, les méthodes d'évaluation de l'impact des véhicules connectés sont susceptibles de pouvoir être appliquées à ce domaine. Des expérimentations de navettes autonomes ont déjà lieu sous la forme de démonstrateurs, et certains véhicules semi-autonomes, ou à conduite partagée, sont déjà sur nos routes. La puissance publique (notamment la direction de la sécurité et de la circulation routière) est en demande d'éléments objectifs et concrets sur l'impact et l'efficacité réelle de ces solutions de mobilité. En tant que chercheur du Cerema, et grâce à mon expérience dans ce domaine, je m'emploierai à proposer des solutions pour répondre à cette demande, en veillant à conserver la prise en compte de l'humain dans l'évaluation.

7.2. Méga données et apprentissage statistique

Nous vivons une période de rupture pour laquelle la numérisation de nos déplacements, et de leur contexte, rend possible l'étude de la mobilité à grande échelle, mais aussi dans toute sa précision. Durant mon parcours de recherche, j'ai acquis une expertise du traitement de masses de données issues de véhicules individuels. Après avoir poursuivi des objectifs précis et concrets (impacts environnementaux, sécurité etc...) dans le cadre des recherches appliquées de l'Ifsttar, j'ai compris que l'information contenue dans de telles masses de données pouvait être exploitée de nombreuses façons pour mieux connaître la mobilité individuelle et améliorer les conditions de sa mise en œuvre. Le contexte actuel de développement des méthodes dites d'apprentissage statistique, ou d'apprentissage profond (deep learning), combiné à l'avènement des « mega données » (big data), propose un cadre très adapté au traitement des masses d'information fournies par les véhicules modernes. Ainsi, le deuxième axe que je souhaite développer plus avant cherche à répondre à une question que je me pose depuis longtemps : Comment exploiter les traces numériques de véhicules pour en extraire de l'information utile pour améliorer les conditions de la mobilité ?

Cette information « utile » peut être par exemple, la position d'éléments fins de l'infrastructure contraignant la conduite (stops, feux rouges, dos d'ânes, zebras, etc.), des zones d'accumulation de comportements défaillants (freinages brusques etc.), la connaissance d'un profil de vitesse recommandée (le profil « moyen » obtenu grâce aux méthodes mises en œuvre dans la thèse de C. Andrieu), le temps d'attente moyen à un feu, ou le taux de piétons observés sur un axe spécifique. Cette liste d'exemples, bien entendu non exhaustive, permet de réaliser le potentiel énorme de l'utilisation de ces données. J'ai d'ores et déjà entamé les travaux sur ces questions en obtenant des financements de post-doctorants et doctorants. J'ai mis en place une collaboration avec l'IGN (institut géographique national) en leur proposant de travailler sur la question des cartes, de la pure création d'information provenant des traces numériques, à la validation des informations qu'elle contient grâce à ces traces. Le cumul d'information fourni par des passages répétés de véhicules, combiné à l'utilisation d'information spécifique à l'IGN (bâtiments, photos satellites, etc.), peut permettre de créer, et de valider avec une grande confiance, la position d'un élément d'infrastructure. Ce sujet de recherche a séduit le laboratoire COGIT de l'IGN, qui m'a proposé d'encadrer le doctorat d'un jeune ingénieur de l'ENSG (Ecole nationale supérieure de géomatique) en poste dans ce laboratoire. Ce travail d'encadrement a débuté en janvier 2016, et les résultats obtenus sont d'ores et déjà prometteurs.

Les méthodes que je mets en œuvre s'appuient sur les techniques mathématiques les plus récentes pour exploiter des collections de traces GPS. Les applications potentielles concernent les grandes métropoles qui ont un besoin croissant d'outils pour la gestion du réseau et l'optimisation de son usage. Le développement de la mobilité collaborative et le partage des informations entre usagers à l'échelle de la ville, en association avec des méthodes statistiques automatisées de création d'une information de haut niveau (cartographie détaillée, état et prévision du trafic, temps d'attente aux intersections, etc.), peuvent contribuer à faciliter l'émergence de villes augmentées

(« smart cities ») avec une mobilité maîtrisée et efficace. Ce type d'approche peut aussi servir aux pays en voie de développement dans lesquels les téléphones portables équipés de GPS sont très répandus. L'exploitation de ces traces pour la création d'un bien commun (cartographie par exemple) me semble utile et pertinent dans ce contexte.

A plus long terme, ce ne sont pas seulement les traces GPS qui seront exploitées, mais aussi l'ensemble des données issues des capteurs équipant les véhicules semi-autonomes ou entièrement autonomes à venir. On estime la quantité de données générée par ces véhicules à environ 1 téraoctet par heure, incluant des images vidéo de la scène routière, des images radars, des images 3D obtenue par lasers, et des données de capteurs de proximité (ultrasons). Cette masse d'information, si elle est correctement exploitée, peut contribuer à créer une image virtuelle très précise et détaillée, voire dynamique, de l'infrastructure de transport et de son usage. Ces idées émergent déjà chez les industriels comme Continental ou Daimler, qui sont en train de constituer de grands centres de recherche sur ces questions, et exploitent déjà certaines approches (Daimler exploite les données de ses camions pour alimenter la cartographie Here en informations routières comme la pente, la qualité de l'enrobé etc.).

Ainsi, je souhaite mettre mes compétences au service du développement de la mobilité de demain, pour contribuer à l'émergence de territoires où le déplacement est maîtrisé et efficace, en accord avec les exigences de sécurité et de respect de l'environnement.

7.3. Conclusion

Le lien entre ces deux axes de recherche passe par la notion de réseau. Le véhicule moderne sera plus ou moins autonome, mais il sera surtout connecté avec l'infrastructure, avec la ville, avec le réseau de transport en commun ou partagé, et aussi bien sûr, avec ses utilisateurs. Ce système complexe n'en est qu'à ses débuts, et les besoins sont énormes pour le construire dans le respect des grands enjeux de société.

Par exemple, il se peut que cette mobilité connectée et autonome devienne paradoxalement le signe d'une perte d'autonomie de l'individu au profit d'algorithmes puisant leur force dans d'immenses « lacs de données ». La question du contenu et du paramétrage de ces algorithmes relève dès lors d'un enjeu public. Je souhaite participer à cette aventure par le biais de mes recherches futures, en concevant des algorithmes ouverts et respectant les enjeux de protection des données personnelles, tout en rendant au citoyen la connaissance qu'il aura contribué à créer via le partage de ses données.

En termes plus concrets et plus spécifiques, je souhaite aborder pleinement la question de l'exploitation des traces numériques de flottes de véhicules ou d'objets mobiles, qui deviendront un jour disponible en très grandes quantités. Cela permettra de construire une connaissance « a priori » très détaillée du réseau routier et de son usage, voire de la ville dans son ensemble. Je

compte ainsi développer le concept de « véhicule bayésien », car je reste convaincu que ceux-ci ne deviendront réellement « intelligents » qu'à la condition de tenir compte de cet a priori. Des travaux récents de Stanislas Dehaene (Collège de France) suggèrent que notre cerveau fonctionne de manière bayésienne. Il me semble dès lors pertinent de considérer cette voie d'étude comme très prometteuse pour améliorer la fiabilité des véhicules, qui seront peut-être moins « autonomes » qu'on ne le pense. A plus long terme, ces concepts peuvent s'appliquer à la mobilité au sens large, en incluant par exemple les traces numériques de piétons ou d'objets mobiles (robots par exemple), afin de construire une cartographie augmentée de la ville, de ses lieux de vies, et de ses réseaux de transport.

J'ai la chance de travailler dans un domaine en pleine révolution, au cœur des préoccupations du citoyen et des collectivités. Pour un chercheur, il est rare et passionnant de se retrouver au cœur d'une telle effervescence suscitant autant d'attentes et de questions de la société. Stimulé par ce contexte, je m'investis avec force dans la recherche de solutions simples, efficaces, et respectueuses des enjeux de société tout autant que des libertés individuelles.

L'obtention du diplôme d'habilitation à diriger des recherches me permettrait d'assumer pleinement les missions qui me sont confiées au Cerema, en collaboration avec le tissu industriel et universitaire local. L'encadrement de doctorants et d'étudiants me permettra de contribuer à transmettre le message d'un monde mobile, ouvert et collaboratif, ainsi que le plaisir d'une recherche appliquée en phase avec les besoins de la société. J'ambitionne ainsi de contribuer activement à la création d'une équipe de recherche autour des systèmes de transport intelligents, et d'assumer le développement et le suivi à long terme des thèmes de recherche explorés dans cette section.

8. Bibliographie

- Andrieu, C. (2013) « *Modélisation fonctionnelle de profils de vitesse en lien avec l'infrastructure et méthodologie de construction d'un profil agrégé* ». PhD thesis, Université Toulouse 3 Paul Sabatier.
- Ahmed M., Karagiorgou S., Pfoser D., Wenk C., (2015) “*Map construction algorithms*”, Springer International Publishing, DOI: 10.1007/978-3-319-25166-0.
- Bagdadi, O., Várhelyi, A. (2011). “*Jerky driving - an indicator of accident proneness?*” Accident Analysis and Prevention, 43 (4), 1359–1363.
- Bagdadi, O. (2013). “*Assessing safety critical braking events in naturalistic driving studies*”. Transportation Research Part F (16) 117-126.
- Benmimoun, M., Fahrenkrog, F., Zlocki, A., Eckstein, L. (2011), “*Incident detection based on vehicle CAN-data within the large scale field operational test “euroFOT”*”, 22nd International Technical Conference on the Enhanced Safety of Vehicles, June 13 - 16, Washington, DC.
- Berlinet, A. and Thomas-Agnan, C. (2004), “*Reproducing kernel Hilbert spaces in probability and statistics*”. Springer.
- Besse P., Guillouet B., Loubes J.M., Royer F., (2016) « *Destination Prediction by Trajectory Distribution Based Model* », IEEE Transactions on Intelligent Transportation Systems, Vol 17, issue 11, DOI: 10.1109/TITS.2016.2547641.
- Besse P., Guillouet B., Loubes J.M., Royer F., (2016) « *Review and Perspective for Distance-Based Clustering of Vehicle Trajectories* », arXiv:1605.03027v1.
- Bigot, J. (2003), « *Recalage de signaux et analyse de variance fonctionnelle par ondelettes : application au domaine biomédical* ». PhD thesis, Université Joseph Fourier, Grenoble I.
- Bratt, H. et Ericsson, E. (1999). « *Estimating speed and acceleration profiles from measured data* ». Conference Proceedings of the 8th International Symposium “Transport and Air Pollution”, Graz, Austria.
- Carsten, O., Lai, F., Chorlton, K., Goodman, P., Carslaw, D. & Hess, S. (2008) “*Speed Limit Adherence and its Effect on Road Safety and Climate Change*”. Report for CfIT and the Motorists’ Forum.
- Chapnick, N. (2007). “*Hypermiling: quest for ultimate fuel economy*”, Edmunds.com, Retrieved from <http://www.edmunds.com/fuel-economy/hypermiling-quest-for-ultimate-fuel-economy.html>.
- Cuevas, A., Febrero, M., and Fraiman, R. (2006) “*On the use of the bootstrap for estimating functions with functional data*”. Computational statistics & data analysis, 51(2):1063–1074.
- Delhome R., Billot R., Nour-Eddin E.F., (2017), “*Travel time statistical modeling with the Halphen distribution family*”, Article in Journal of Intelligent Transportation Systems, DOI: 10.1080/15472450.2017.1326115
- Diggle P. J., Liang K. Y. et Zeger S. L. (1994). « *Analysis of longitudinal data*”. Oxford Press.
- Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J. , Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R.,

- Jermeland, J., and Knippling, R. R. (2006) *“The 100-Car Naturalistic Driving Study: Phase II—Results of the 100-Car Field Experiment”*. DOT HS 810 593. NHTSA, 2006.
- Driscoll, R., Page, Y., Lassarre, S., Ehrlich, J. (2007) *“LAVIA – an evaluation of the potential safety benefits of the French Intelligent Speed Adaptation project”*, AAAM.
 - Elvik, R., Christensen, P., and Amundsen, A. (2004), *“Speed and Road Accidents. An Evaluation of the Power Model.”* Oslo, Norway, Transportøkonomisk Institutt.
 - Elvik, R. (2013). *“A re-parameterisation of the Power Model of the relationship between the speed of traffic and the number of accidents and accident victims”*. Accident Analysis and Prevention, 50, 854-860.
 - FESTA consortium, (2014), FESTA Handbook version 5, available at http://wiki.fot-net.eu/index.php/FESTA_Handbook
 - Finch D J, Kompfner P, Lockwood C R and Maycock G (1994). *“Speed, speed limits and accidents”*. Project Report PR58. Crowthorne: TRL Limited.
 - Guo, F., Klauer, S., Hankey, J., Dingus, T., (2010), *“Near crashes as crash surrogate for naturalistic driving studies”*, Transportation Research Record: Journal of the Transportation Research Board, 2147, pp 66-74, Transportation Research Board of the National Academies.
 - Heinrich, Hw. (1931), *“Industrial accident prevention: a scientific approach”*, New York: McGraw-Hill.
 - Kerper, M. Wewetzer, C., and Mauve, M. (2012). *“Analyzing vehicle traces to find and exploit correlated traffic lights for efficient driving”*. In Intelligent Vehicles Symposium (IV), 2012 IEEE, pages 310–315.
 - Kloeden, C. N; Ponte, G.; McLean, A. J., (2001), *“Travelling speed and the risk of crash involvement on rural roads”*.
 - Kloeden, C. N.; McLean, J.; Glonek, G. F. V. (2002), *“Reanalysis of travelling speed and the risk of crash involvement in Adelaide South Australia”*, Australian Transport Safety Bureau.
 - Liang, K.Y. and Zeger, S. L. (1986) *“Longitudinal Data Analysis Using Generalized Linear Models”*, Biometrika, Vol. 73, No. 1 (Apr., 1986), pp. 13-22.
 - Ligterink, N. E., Van Zyl, S. and Stelwagen, U. (2014). *“CO2 -based response emission modelling”* Transport and Air Pollution Conference. Graz, Austria.
 - Ozuysal M, Fua P, Lepetit V (2007) *“Fast keypoint recognition in ten lines of code”*. In: Computer Vision and Pattern Recognition, CVPR’07. IEEE Conference, pp 1-8.
 - Rakha, H., Dion, F. et Sin, H. (2001). *« Using Global Positioning System Data for Field Evaluation of Energy and Emission Impact of Traffic Flow Improvement Projects: Issues and Proposed Solutions »*. Transportation Research Record: Journal of the Transportation Research Board. 1768, p. 210–223.
 - Ramsay, J. O. and Silverman, B. W. (2005) *“Functional Data Analysis”*, Second Edition. Springer-Verlag, New York.
 - Ramsay, J. O. and Li, X. (1998) *“Curve registration”*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 60(2):351–363.
 - Rakha Y., H. Ding (2003). *“Impact of stops on vehicle fuel consumption and emissions”*. Journal of transportation engineering, vol. 129, part 1. pages 23–32.
 - Romon S. (2006), *« Lavia : Traitement des données »*. Rapport technique

- Song, X., Kanasugi, H., Shibasaki, R., DeepTransport: “*Prediction and Simulation of Human Mobility and Transportation Mode at a Citywide Level*”. Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16), 2016.
- Stroup, W.W. (2012), “*Generalized Linear Mixed Models*”, CRC Press
- Sun, Y. et Genton, M. (2011). « *Functional boxplots* ». Journal of Computational and Graphical Statistics. 20(2)
- Taylor M C, Lynam D A and Baruya A (2000). “*The effects of drivers’ speed on the frequency of road accidents*”. TRL Report TRL421. Crowthorne: TRL Limited.
- Trotta M., (2017), « *Avons-nous encore besoin des données radar?* », Bruxelles, Belgique: Institut Belge pour la Sécurité Routière – Centre de Connaissance Sécurité Routière.
- Violette, E., Hublart, A., Subirats, P. and Louah, G. « *Estimation de la vitesse pratiquée sur un itinéraire - méthode et application* ». Technical report, CETE Normandie Centre, DITM, ESM, 2010.
- Wahba, G., “*Spline models for observational data*”. SIAM, 1990.